THE COMPARATIVE PSYCHOLOGY OF INTELLIGENCE: MACPHAIL REVISITED

EDITED BY: Michael Colombo, Damian Scarf and Thomas Zentall
PUBLISHED IN: Frontiers in Psychology

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88966-597-6
DOI 10.3389/978-2-88966-597-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE COMPARATIVE PSYCHOLOGY OF INTELLIGENCE: MACPHAIL REVISITED

Topic Editors:
Michael Colombo, University of Otago, New Zealand
Damian Scarf, University of Otago, New Zealand
Thomas Zentall, University of Kentucky, United States

Citation: Colombo, M., Scarf, D., Zentall, T., eds. (2021). The Comparative Psychology of Intelligence: Macphail Revisited. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-597-6

Table of Contents

frontiers in Psychology

EDITORIAL
published: 10 February 2021
doi: 10.3389/fpsyg.2021.648782

Editorial: The Comparative Psychology of Intelligence: Macphail Revisited

Michael Colombo[1]*, *Damian Scarf*[1] *and Tom Zentall*[2]

[1] *Department of Psychology, University of Otago, Dunedin, New Zealand,* [2] *Department of Psychology, University of Kentucky, Lexington, KY, United States*

Keywords: Macphail, null hypothesis, comparative cognition, intelligence, species differences

Editorial on the Research Topic

The Comparative Psychology of Intelligence: Macphail Revisited

In a series of papers in the 80s, Macphail (1982, 1985, 1987) put forth evidence in support of the Null Hypothesis for differences in intelligence across non-human species, stating in his 1985 paper that there are "no differences, either qualitative or quantitative, among vertebrates" (p. 46). He further claimed that association formation dominates intelligent behavior, that any differences in intelligence between species could be accounted for by the differential effects of contextual variables, that learning mechanisms are of general applicability and did not evolve as species-specific specializations, and that human intelligence differs from that of other animals in that only humans possess a species-specific language device.

The peer commentaries on Macphail's (1987) *Behavioral and Brain Sciences* paper were generally negative. The most scathing comment was from Goldman-Rakic and Preuss (1987) who suggested that "Macphail's 'null hypothesis' is merely the epitaph on the head stone of comparative cognition" (p. 667). Surprisingly, despite the negative tone, Macphail (1987) ended his response to the commentaries on an uplifting note, stating that "For my part, I remain an optimist, and prefer to see the failure to demonstrate differences as evidence not that our scientific procedures are weak but that the animal mind is not what we expected it to be. And after all, did we *really* expect that it would be?" (p. 688). Based on the growth of comparative cognition in the more than three decades since Macphail's (1987) paper, and the papers included in this Research Topic, it is clear Macphail was right to be optimistic. As Pepperberg notes, Macphail (1987) should be given credit for "…instigating a variety of controversies, stimulating the wide-ranging discussions, and generating the types of challenges that have led to many new avenues of research" (p. 10).

At the time, many of the remarkable abilities of non-human animals were unknown to Macphail. Abilities such as episodic memory, theory of mind, orthographic processing, planning for the future, fast mapping, and numerical competence, to name but a few, were yet to have their time in the limelight. With the wealth of comparative data collected over the past 30 years, we thought it was timely to review the status of Macphail's Null Hypothesis, and gauge how the current generation of comparative psychologists approach the inherent challenge that Macphail put forward.

The manuscripts we received ranged from empirical to theoretical, and covered research on a variety of different animals such as pigeons, fish, rats, humans, parrots, eels, crows, monkeys, marine mammals, and spiders. All the papers addressed Macphail's main claim that there are no qualitative or quantitative differences in intelligence across species. A subset of papers addressed Macphail's other claims that (1) contextual variables can explain all of the observed differences between species, (2) associative processes account for all non-human intelligence, and (3) the uniqueness of human intelligence is due to a species-specific language acquisition device.

Edited and reviewed by:
Sarah Till Boysen,
Independent Researcher, Sunbury,
United States

****Correspondence:***
Michael Colombo
colombo@psy.otago.ac.nz

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *01 January 2021*
Accepted: *19 January 2021*
Published: *10 February 2021*

Citation:
Colombo M, Scarf D and Zentall T
(2021) Editorial: The Comparative
Psychology of Intelligence: Macphail
Revisited. Front. Psychol. 12:648782.
doi: 10.3389/fpsyg.2021.648782

QUALITATIVE AND QUANTITATIVE DIFFERENCES

Issues of the definition of intelligence aside, virtually all of the manuscripts supported Macphail's view that there are no qualitative differences in intelligence across animals. Whether it is orthographic processing in pigeons (Scarf and Colombo) numerical competence in pigeons, crows, and fish (Scarf and Colombo; Nieder; Petrazzini et al.), counterintuitive features of skill learning in rats (Reid and Swafford), planning in humans (Martin-Ordas), syntactic abilities in parrots (Pepperberg), spatial learning in eels (Watanabe), working memory in crows (Hahn and Rose), the irrational choices of pigeons and humans (Stagner et al.; Zentall), equivalence relations (Colombo and Scarf; Zentall), a variety of other abilities in marine mammals (Bauer et al.) and birds (Bastos and Taylor; Zentall), tool use (Cabrera-Álvarez and Clayton), predatory strategies in spiders (Cross et al.), or complex sequential behavior in rats, mice, pigeons, and humans (Fountain et al.), not to mention a host of other sophisticated behaviors alluded to in many of these papers that have been conducted over the past three and a half decades, all vertebrates seem capable of displaying behaviors that were once considered the domain of only humans or, at most, non-human primates. On the issue of qualitative differences, what might easily be considered the core thrust of the Null Hypothesis, Macphail was nothing short of prescient.

While there was consensus regarding the absence of qualitative differences across species, the majority of the papers were in favor of rejecting Macphail's Null Hypothesis with respect to no quantitative differences across species. The issue of quantitative differences is difficult and depends, of course, on how one chooses to define "quantitative." Macphail (1985) defined a quantitative difference as "…one species used a mechanism or mechanisms common to both species more efficiently than the other, and this might be reflected in a faster rate of solution or better asymptotic performance…" (p. 38). If defined by asymptotic performance, the fact that the performance of animals on tasks of orthographic processing and numerical competence (Scarf and Colombo), syntactic abilities (Pepperberg), and tool use (Cabrera-Álvarez and Clayton) are comparable between distantly related species (namely birds and primates) lends supports to Macphail's view that there may not even be quantitative differences across species. Cross et al. take this one step further by suggesting that there are even comparable predatory strategies between vertebrate species and spiders (i.e., a non-vertebrate species). That said, most of the papers in the Research Topic accepted that quantitative differences do exist between species, especially if one takes Macphail's definition of "quantitative" as measured by faster rates of solution, or more efficient use of an ability, a vague term that might better be cast as greater flexibility in the use of an ability.

ROLE OF CONTEXTUAL VARIABLES

Most of the papers also agree that contextual variables can explain many of the observed differences in abilities across species. The case for the importance of contextual variables was made most strongly by Colombo and Scarf who showed that with respect to a number of tasks, when contextual variables are properly accounted for, qualitative differences that have been observed across different animals vanish. Schubiger et al. presented an enormous list of potential contextual variables, both subject-related and task-related, supporting the possibility that once considered, evidence of even quantitative differences may vanish as well.

THE ROLE OF ASSOCIATIONS AND THE UNIQUENESS OF HUMAN INTELLIGENCE

Although few, some papers also tackled Macphail's (1987) claims that association formation underlies intelligence and the uniqueness of human language. With respect to association formation, there were those that supported the idea that association processes underlie cognitive behavior (Scarf and Colombo; Colombo and Scarf), and those that disagreed that cognitive behavior is guided by associative processes (Bastos and Taylor). There seems little doubt that associative processes are universal. Whether Macphail was correct when he stated that association formation dominates intelligent behavior hinges, as so many commentaries in his 1987 paper raised, on how one defines "intelligence." Bauer et al. were correct when they stated that "In many ways, 'intelligence' seems to be a folk psychology term that maps poorly on natural psychological and biological processes, and therefore, lends itself to a wide range of often-inconsistent interpretations" (p 14). Whether behaviors extend beyond associative process is a complex topic. In his typically prescient manner, Macphail (1987) forestalled this issue when he stated that "The problem is that we do not understand what processes underlie…complex behavior" (p. 683).

On the topic of language, countering Macphail's (1987) claim of a species-specific language acquisition device that sets humans apart from other animals, Corballis elegantly highlights the fact that language is now recognized as an amalgam of several abilities (e.g., mental time travel, theory of mind, etc.), many which are present to varying degrees in non-human animals. Further, according to Corballis, it is unlikely that the hierarchical and generative aspects of thought are even unique to humans. Petrazzini et al. echoed similar sentiments in their paper. Thus, if anything, it would seem that Macphail erred on the side of caution with his Null Hypothesis, as it appears that even language may not be a dividing line between humans and non-humans.

THEORETICAL ISSUES

A number of papers also addressed a variety of theoretical issues around the notion of a Null Hypothesis. Petrazzini et al. called into question how we assess the presence vs. absence of differences between species. Traditionally, in keeping with its namesake, Macphail's Null Hypothesis is assessed by rejecting or failing to reject the null hypothesis based on a *p*-value. Petrazzini et al. argued for finer comparative methodologies such as a Bayesian approach, which would evaluate the relative strength of two competing hypotheses. Similarly, Bastos and Taylor also suggested a Bayesian framework to distinguish between support for the Null Hypothesis and a lack of statistical power. Taking this one step further, even when paired with null-hypothesis testing, Bayes factor could be made mandatory for comparative papers, allowing a measure of confidence in any null findings.

Finally, Hahn and Rose argue that working memory is a critical component of cognitive abilities, and that a better way to compare species is to use their working memory capacity and retention limits as a proxy for their cognitive abilities. "Differences and similarities in WM (e.g., in its capacity) may offer insights into why some animals may be (un-)able (*sic*) to display certain cognitive behaviors. Macphail's null-hypothesis can thus be investigated in the light of potentially qualitative, and quantitative differences of a fundamental trait of cognition" (p. 2).

CONCLUSIONS

In light of the results of research over the past 30 years, Macphail's hypothesis that all vertebrates have similar cognitive capacities may not be as implausible as it may have at appeared at the time. In order to conclude that there are qualitative or quantitative differences among species, however, one must first eliminate important differences in contextual variables concerned with perception, motor skills, and motivation. If nothing else, Macphail's proposition has served as encouragement and a valuable challenge to comparative researchers to conduct well designed tests of the abilities of a large number of animal species.

Fountain et al. are correct when they say that "…Macphail's claim continues to challenge all empiricists and theorists to consider the power of even simple neural systems to account for animals' ability to encode simplicity in terms of neural representation from the complexity of the surrounding environmental milieu" (p. 3). With respect to the other issues that Macphail raised, such as whether species differ quantitatively and whether association formation dominates intelligent behavior, Macphail (1987) provided a roadmap as to how the field of comparative cognition can advance our understanding of the human mind by stating that "I express the hope that workers seeking to disprove the null hypothesis will attempt to devise novel tasks for comparative work—tasks which associative devices could not solve" (p. 683).

There is much that the field of comparative cognition owes to Macphail. Corballis stated it perfectly when he said that "Macphail's writing set up the challenge, and attempting to answer it can only advance our knowledge of how animals think, and where humans fit into the overall scheme of things" (p. 8).

AUTHOR CONTRIBUTIONS

MC wrote the first draft. DS and TZ reviewed and edited the draft. All authors approved the work for publication.

REFERENCES

Goldman-Rakic, P. S., and Preuss, T. M. (1987). Wither comparative psychology? *Behav. Brain Sci.* 10, 666–667.

Macphail, E. (1982). *Brain and Intelligence in Vertebrates.* Clarendon Press.

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans.R. Soc. Lond. Ser. B Biol. Sci.* 308, 37–51.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

MINI REVIEW
published: 17 January 2020
doi: 10.3389/fpsyg.2019.03017

Columban Simulation Project 2.0: Numerical Competence and Orthographic Processing in Pigeons and Primates

Damian Scarf* and Michael Colombo

Department of Psychology, University of Otago, Dunedin, New Zealand

Thirty years ago Burrhus Frederic Skinner and Robert Epstein began what is known as the Columban Simulation Project. With pigeons as their subjects, they simulated a series of studies that purportedly demonstrated insight, self-recognition, and symbolic communication in chimpanzees. In each case, with the appropriate training, they demonstrated that pigeons performed in a comparable manner to chimpanzees. When discussing these studies in the context of his Null Hypothesis, Macphail paid little attention to how the pigeons and chimpanzees solved the tasks and simply assumed that successful performance on the tasks reflected a similar underlying mechanism. Here, following a similar process to the original Columban Simulation Project, we go beyond this success testing and employ the signature testing approach to assess whether pigeons and primates employ a similar mechanism on tasks that tap numerical competence and orthographic processing. Consistent with the Null Hypothesis, pigeons and primates successfully passed novel transfer tests and, critically, displayed comparable cognitive signatures. While these findings demonstrate the absence of a qualitative difference, the time taken to train pigeons on these tasks revealed a clear quantitative difference.

Keywords: Null Hypothesis, comparative cognition, numerical competence, orthographic processing, counting, reading

Edited by:
Christian Agrillo,
University of Padua, Italy

Reviewed by:
Michael Beran,
Georgia State University,
United States
Maria Elena Miletto Petrazzini,
Queen Mary University of London,
United Kingdom

****Correspondence:***
Damian Scarf
damian@psy.otago.ac.nz

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *10 November 2019*
Accepted: *19 December 2019*
Published: *17 January 2020*

Citation:
Scarf D and Colombo M (2020)
Columban Simulation Project 2.0:
Numerical Competence
and Orthographic Processing
in Pigeons and Primates.
Front. Psychol. 10:3017.
doi: 10.3389/fpsyg.2019.03017

INTRODUCTION

Thirty years ago Burrhus Frederic Skinner and Robert Epstein began what is known as the Columban Simulation Project (Epstein, 1981, 1986, 1991). First floated as the "Pigeon Simulation Project," "Pigeon" was switched out for "Columban" (derived from the taxonomic name for pigeons) because it sounded more "computer like" (Epstein, 1981). Rather than just a play on words, Skinner and Epstein drew a great deal on the computer simulation literature and their intention was to provide a true simulation, one that "faithfully reproduces all significant characteristics of some phenomenon" (Epstein, 1986, p. 132). What were they trying to simulate? With pigeons as their subjects, they were trying to simulate a series of studies that purportedly demonstrated insight (Kohler, 1925), self-recognition (Gallup, 1970), and symbolic communication (Savage-Rumbaugh et al., 1978) in chimpanzees.

The three studies followed a somewhat similar method (Epstein et al., 1980, 1981, 1984). Pigeons first went through a series of training phases and, following their successful completion, were transferred to the pivotal test that consisted of placing pigeons in a novel situation and observing their behavior. The study of insight provides perhaps the most fruitful example of their approach. The study was based on the work of Kohler (1925), who presented a group of chimpanzees with an intriguing problem. In a large enclosure, Kohler (1925) suspended a banana 2 to 3 m above the ground. Also in the enclosure was a small wooden box. As Kohler (1925) notes, after realizing the banana was out of reach, one chimpanzee "…suddenly stood still in front of the box, seized it, tipped it hastily straight toward the [banana]…began to climb upon it…and springing upward with all his force, tore down the banana" (p. 40–41). Simulating this behavior in pigeons, Epstein et al. (1984) made some basic assumptions about the behaviors that may have led up to the chimpanzee's behavior. Specifically, they reinforced pigeons to move a small box toward a target and to climb upon the box to reach a small toy banana suspended from the ceiling. In the critical test, they placed the box in one corner and suspended the banana in the other. Mimicking the chimpanzee described above, after initially pacing and looking perplexed, the pigeon pecked/pushed the box toward the banana, stopped underneath it, and then climbed upon the box and pecked the banana.

A discussion regarding whether Kohler's (1925) chimpanzees and Epstein et al.'s (1984) pigeons truly displayed insight is beyond the scope of the current review. Indeed, Epstein et al. (1981) noted that the concepts themselves, and discussions regarding them, "…impede the search for the controlling variables of the behavior they are said to produce" (p. 696). When discussing these studies in the context of his Null Hypothesis, Macphail (1985) noted that he was "…not concerned here to discuss the nature of the solutions of such problems, whether insight, for example, is a necessary or a useful concept, the key point of interest is the parallel between the chimpanzee and the pigeon performance. There is clearly every reason to suppose that the pigeons solved the problem in exactly the same way as the chimpanzee" (p. 47). While we agree with Macphail's (1985) first point about the utility, or lack thereof, of concepts such as insight, one could take issue with the second. The point is that similar looking behavior does not imply a similar underlying mechanism and, when one is arguing for the absence of cognitive differences between species, the variables that control the behavior matter. Gallup (1985), when discussing Epstein et al.'s (1981) simulation of his self-recognition study, similarly stated that "Simply because you can mimic the behavior of one species by reinforcing a series of successive approximations to what looks like the same routine in another, it does not follow that the behavior of the former species necessarily arose in the same way" (p. 633).

Mirroring Gallup's (1985) argument, one could argue that a major limitation of the Columban Simulation Project, and one that may limit its implications for the topics under study, was the focus on what is now termed success-testing (Taylor, 2014). That is, beyond the actual behavior observed (e.g., pecking a blue dot on their body), there were few, if any, additional measures that would allow a closer analysis and comparison of the chimpanzees' and pigeons' behavior. An approach that goes beyond mere success-testing is signature-testing, which holds that we "…search for the signatures of various cognitive mechanisms in terms of their errors, biases and limits, rather than a "success-testing" approach where experimenters simply examine whether a problem can be solved or not" (Taylor, 2014, p. 369).

COLUMBAN SIMULATION PROJECT 2.0

With the aim of testing the limits of Macphail's (1985) Null Hypothesis, and drawing inspiration from the Columban Simulation Project (Epstein, 1981, 1986, 1991), we set about comparing birds and primates using a signature-testing approach. We initially sought out corvids as experimental subjects. Indeed, work with corvids was rapidly growing at that time and Emery (2006) noted that corvids had displayed abilities that "…are qualitatively and quantitatively more sophisticated than have been demonstrated by other birds, and in many domains comparable to monkeys and apes" (p. 23). Unfortunately, New Zealand is home to only one corvid species (rooks) and their low numbers in the South Island made sourcing the birds extremely difficult. Consequently, much like Epstein (1986) noted when pondering the question of "why pigeons?" we simply went with the materials at hand, and that happened to be the humble pigeon. Constantly on the search for tasks, two high-profile studies presented themselves: first, Brannon and Terrace's (1998) study on numerical competence and second, Grainger et al.'s (2012) study on orthographic processing. Critically, these studies not only included a novel transfer test (for which we could test for success), but also a number of behavioral metrics that would allow us to compare the signatures/cognitive mechanisms that pigeons and monkeys applied to the tasks.

Numerical Competence in Monkeys and Pigeons

Numerical competence consists of three concepts, quantity (i.e., cardinality), rank (i.e., ordinality), and counting (i.e., nominal/labeling) (Nieder, 2005). Obviously, in the absence of language, counting is beyond the grasp of non-human animals. Quantity and rank, however, can easily be tested in non-human animals (Chen et al., 1997; Brannon and Terrace, 1998; Scarf and Colombo, 2011; Scarf et al., 2011). Brannon and Terrace's (1998) study in rhesus monkeys is one of the most powerful examples of cardinality. They trained two monkeys to order stimuli consisting of one, two, three, or four elements. Critically, to ensure the monkey's behavior was driven by the number of elements in each stimulus rather than other features of the stimuli (e.g., surface area), the elements varied in size, color, and shape. Monkeys were trained on 35 of these 4-item lists and then tested with novel pairs of numerical stimuli. The pairs were one of three types: familiar-familiar (F-F) pairs contained two numerosities drawn from the training range (i.e., 1–4), familiar-novel (F-N) pairs contained one numerosity from the training range (i.e., 1–4) and one numerosity drawn from the novel range of 5 to 9, and novel-novel (N-N) pairs contained two novel numerosities drawn from the 5 to 9 range. Following in Brannon and Terrace's (1998) footsteps, Scarf et al. (2011) trained four pigeons using an

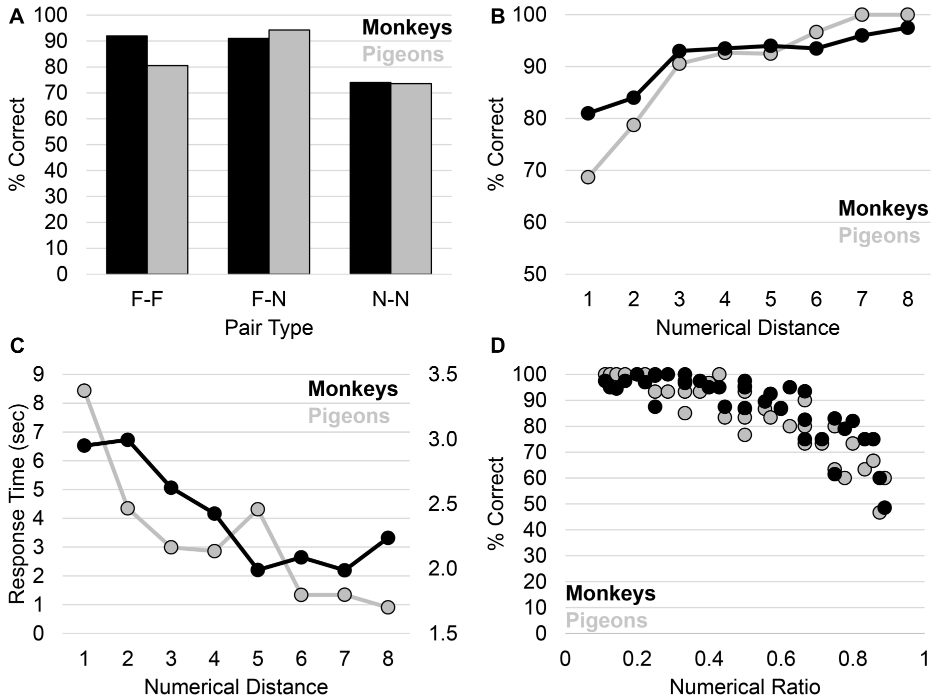


FIGURE 1 | (A) Test performance of subjects. **(B)** Test performance of subjects as a function of the numerical distance between the test pair. **(C)** Response time of subjects as a function of the numerical distance between the test pair. **(D)** Test performance of subjects as a function of the numeric ratio between the test pair.

identical paradigm, with the exception that pigeons were trained on 35 3-item, rather than 4-item, lists.

With respect to success testing, consistent with the view that both monkeys and pigeons acquired an abstract numerical rule during training, both performed above chance on the critical N-N pairs (**Figure 1A**). As one would expect, the monkeys and pigeons also performed well on their respective F-F and F-N pairs (**Figure 1A**). Following the signature approach, we delved deeper into the behavior of the monkeys and pigeons by assessing two aspects of their performance. First, we assessed the distance effect, the finding that as the distance between two numbers increases, accuracy increases and response time decreases (Moyer and Landauer, 1967; Buckley and Gillman, 1974). For example, subjects should be faster and more accurate with pair 1 vs. 9 (i.e., a distance of 8) than pair 2 vs. 4 (i.e., a distance of 2). Both monkeys and pigeons displayed a clear distance effect, with accuracy increasing (**Figure 1B**) and response time decreasing (**Figure 1C**) as the numerical distance between the two stimuli increased.

Second, we investigated whether the performance of the monkeys and pigeons was constrained by Weber's (1834) law. Weber's (1834) law reflects the fact that it is not only the distance between stimuli, but also their ratio, that influences discrimination performance. For example, although the distance between pair 1 vs. 2 and pair 8 vs. 9 is 1, the ratio between them is vastly different (0.5 vs. 0.89), thus we would expect performance on pair 1 vs. 2 to be higher than that on pair 8 vs. 9. Consistent with both monkeys and pigeons representing the stimuli in a similar way to humans, their performance was constrained by Weber's (1834) law in that performance decreased as the numeric ratio increased (**Figure 1D**).

Orthographic Processing in Baboons and Pigeons

Learning to read involves the acquisition of letter-sound relationships (i.e., decoding skills) and the ability to visually recognize words (i.e., orthographic knowledge). Much like counting, in the absence of language, decoding skills are human unique. In contrast, recent research and theory suggest that orthographic processing may derive from the exaptation or recycling of visual circuits that are shared by both human and non-human animals (Dehaene, 2009; Dehaene and Cohen, 2011). To test this theory, Grainger et al. (2012) trained six baboons to discriminate four-letter English words (e.g., DONE) from 7,832 four-letter non-words/gibberish (i.e., DMET). Word by word, the baboons acquired vocabularies of between 81 words and 308 words. Following Grainger et al. (2012), Scarf et al. (2016) trained four pigeons using an identical paradigm, with the pigeons acquiring vocabularies between 26 and 58 words. Following training, the success test consisted of presenting subjects with novel words. The baboons and pigeons displayed a similar level of performance with novel words (**Figure 2A**).

To assess whether the signature underlying their performance matched that displayed by humans, three aspects of the baboons' and pigeons' performance was assessed. First, the performance of baboons and pigeons on words increased as the bigram frequency of the words increased (Vinckier et al., 2011). That is, the more

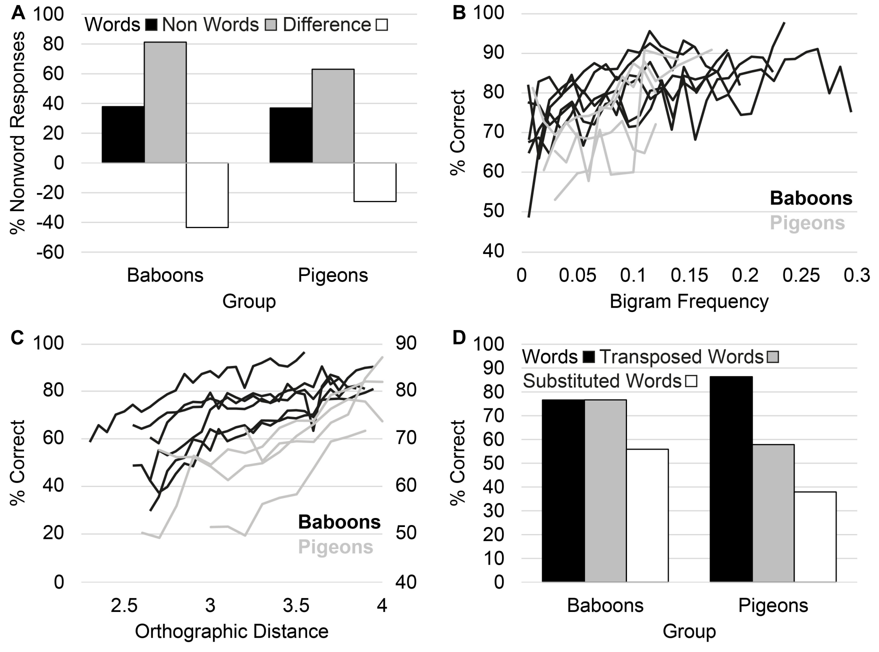


FIGURE 2 | (A) Test performance of subjects. **(B)** The performance of subjects as a function of the bigram frequency of words. **(C)** The performance of subjects on non-words as a function of their similarity to words. **(D)** The performance of subjects on the transposed word test.

frequent certain letter pairs were in the baboons' and pigeons' vocabulary, the more accurate they were in responding to them (**Figure 2B**). Second, the performance of baboons and pigeons on non-words increased as the orthographic similarity between non-words and words in their vocabulary decreased (**Figure 2C**). Orthographic similarity was measured by calculating each non-words' Levenshtein (1966) distance, which is the number of changes (e.g., substituting letters in the non-word) required to turn a non-word into a word. For example, to change the non-word DMET into the word DONE, would require substituting letters M, E, and T, for letters, O, N, and E, respectively (i.e., 3 substitutions). Finally, baboons and pigeons were presented with a transposed-letter test (Ziegler et al., 2013; Scarf et al., 2016). The test consisted of presenting subjects with words in which the order of the internal letters were transposed (e.g., "DONE" transposed to "DNOE"), essentially turning them into non-words. Similar to humans, baboons and pigeons showed a tendency to misclassify transposed non-words as words (**Figure 2D**).

TRUE SIMULATIONS OR CIRCUS TRICK?

As noted above, Epstein (1981, 1984, 1986, 1991) went to great lengths to explain that the intention of the Columban Simulation Project was to produce true, rather than adequate (i.e., reproduces only some characteristics) or dissimilar (i.e., reproduces no characteristics), simulations (Murphy, 1950). Moreover, Epstein (1986) made clear the simulations were not mere superficial circus tricks, such as a "…circus animal that wears glasses and turns the pages of a book appears to be a reader but does not do these things for the same reasons a person does" (p. 132). An important question is where on this spectrum, from circus trick to true simulation, do the current studies sit? The ability of subjects in both the numerical and orthographic studies to pass novel transfer tests demonstrates that their performance is no surface trick. In fact, we would argue that our simulations are true simulations, and perhaps even stronger simulations than those conducted by Epstein et al. (1980, 1981, 1984). For example, in the studies of numerical competence, monkeys and pigeons displayed two characteristics of human numerical processing, namely the distance effect (Moyer and Landauer, 1967; Buckley and Gillman, 1974) and Weber's (1834) law. Similarly, in the studies of orthographic processing, baboons and pigeons displayed three features that literate humans display when processing words. Specifically, they perform better on high bigram-frequency words (Grainger et al., 2012), perform better on non-words as their orthographic distance from words increased (Keuleers et al., 2012), and display a clear transposed-letter effect (Perea and Lupker, 2004; Duñabeitia et al., 2014; Tiffin-Richards and Schroeder, 2015).

IMPLICATIONS FOR MACPHAIL'S NULL HYPOTHESIS

Much like the initial set of studies in the Columban Simulation Project, our work on numerical competence and orthographic processing clearly demonstrates there are no qualitative differences between primates and pigeons on these tasks. Macphail's (1985, 1987) Null Hypothesis holds that there are also no quantitative differences between species. A quantitative difference is defined as "…one species used a mechanism or mechanisms common to both species more efficiently than the other, and this might be reflected in a faster rate of solution or better asymptotic performance level by one species in some task solved by both" (Macphail, 1985, p. 38). The answer to this question is somewhat more difficult. If our measure of asymptotic performance is based on performance on the novel transfer tests, than the current studies support the Null Hypothesis, with the pigeons performing comparable to the monkeys on the novel numerical pairs (Monkeys: 74% vs. Pigeons: 73.6%) and comparable to the baboons on the novel words (Baboons: 62.1% vs. Pigeons: 63%).

If we use training time as our measure of rate of solution, however, a clear quantitative difference emerges. For example, Brannon and Terrace's (1998) monkeys acquired their 35 4-item training lists in a matter of months, while Scarf et al.'s (2011) pigeons required well over a year to acquire their much simpler 35 3-item lists. Similarly, Grainger et al.'s (2012) baboons acquired their relatively larger vocabularies (81 to 308 words) in a mere month and a half, while Scarf et al.'s (2016) pigeons took upward of 2 years to acquire their much smaller vocabularies (26 to 58 words). Vast differences in the time required to train pigeons and primates on tasks is something we have observed across an array of tasks (Colombo et al., 2003; Scarf et al., 2018), and supports a clear quantitative difference across animals.

CONCLUSION

Macphail (1987) noted that he "…cannot claim strong support for the conclusion that there are no quantitative differences in intelligence" (p. 685). Although alternative training procedures have been shown to drastically impact or reverse differences between animals (Lucon-Xiccato and Bisazza, 2014), based on our extensive experience with pigeons and monkeys, we find it extremely unlikely that any change would eliminate the marked and consistent differences that appear to exist between these groups. While not fulfilling the quantitative component of Macphail's (1985, 1987) Null Hypothesis, the Columban Simulation Project 2.0 provides convincing evidence that there are no qualitative differences between pigeons and primates on the numerical or orthographic tasks we have studied. Critically, this conclusion holds at both the success and signature level. That is, the absence of qualitative differences holds when we look at the performance of pigeons and primates on the novel transfer tests and, going one step further, look at their respective cognitive signatures (a.k.a., the variables that control the behavior).

AUTHOR CONTRIBUTIONS

DS and MC conceptualized and wrote the manuscript.

REFERENCES

Brannon, E. M., and Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science* 282, 746–749. doi: 10.1126/science.282.5389.746

Buckley, P. B., and Gillman, C. B. (1974). Comparisons of digits and dot patterns. *J. Exp. Psychol.* 103, 1131–1136. doi: 10.1037/h0037361

Chen, S., Swartz, K. B., and Terrace, H. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychol. Sci.* 8, 80–86. doi: 10.1111/j.1467-9280.1997.tb00687.x

Colombo, M., Cottle, A., and Frost, N. (2003). Degree of representation of the matching concept in pigeons (*Columba livia*). *J. Comp. Psychol.* 117, 246–256. doi: 10.1037/0735-7036.117.3.246

Dehaene, S. (2009). *Reading in the Brain: The New Science of How We Read*. New York, NY: Viking Penguin.

Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15, 254–262. doi: 10.1016/j.tics.2011.04.003

Duñabeitia, J. A., Orihuela, K., and Carreiras, M. (2014). Orthographic coding in illiterates and literates. *Psychol. Sci.* 25, 1275–1280. doi: 10.1177/0956797614531026

Emery, N. J. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philos. Trans. R. Soc. Lond. B* 361, 23–43. doi: 10.1098/rstb.2005.1736

Epstein, R. (1981). On pigeons and people: a preliminary look at the Columban simulation project. *Behav. Analy.* 4, 43–55. doi: 10.1007/bf03391851

Epstein, R. (1984). Simulation research in the analysis of behavior. *Behaviorism* 12, 41–59.

Epstein, R. (1986). "Simulation research in the analysis of behavior," in *Research Methods in Applied Behavior Analysis*, eds A. Poling, and R. W. Fuqua, (New York, NY: Plenum Press), 127–155. doi: 10.1007/978-1-4684-8786-2_7

Epstein, R. (1991). Skinner, creativity, and the problem of spontaneous behavior. *Psychol. Sci.* 2, 362–370. doi: 10.1111/j.1467-9280.1991.tb00168.x

Epstein, R., Kirshnit, C. E., Lanza, R. P., and Rubin, L. C. (1984). 'Insight' in the pigeon: antecedents and determinants of an intelligent performance. *Nature* 308, 61–62. doi: 10.1038/308061a0

Epstein, R., Lanza, R. P., and Skinner, B. (1981). "Self-awareness" in the pigeon. *Science* 212, 695–696. doi: 10.1126/science.212.4495.695

Epstein, R., Lanza, R. P., and Skinner, B. F. (1980). Symbolic communication between two pigeons (*Columba livia* domestica). *Science* 207, 543–545. doi: 10.1126/science.207.4430.543

Gallup, G. G. (1970). Chimpanzees: self-recognition. *Science* 167, 86–87. doi: 10.1126/science.167.3914.86

Gallup, G. G. (1985). Do minds exist in species other than our own? *Neurosci. Biobehav. Rev.* 9, 631–641. doi: 10.1016/0149-7634(85)90010-7

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., and Fagot, J. (2012). Orthographic processing in baboons (Papio papio). *Science* 336, 245–248. doi: 10.1126/science.1218152

Keuleers, E., Lacey, P., Rastle, K., and Brysbaert, M. (2012). The british lexicon project: lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behav. Res. Methods* 44, 287–304. doi: 10.3758/s13428-011-0118-4

Kohler, W. (1925). *The Mentality of Apes*. London: Routledge & Kegan Paul.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet. Phys. Doklady* 10, 707–710.

Lucon-Xiccato, T., and Bisazza, A. (2014). Discrimination reversal learning reveals greater female behavioural flexibility in guppies. *Biol. Lett.* 10:20140206. doi: 10.1098/rsbl.2014.0206

Macphail, E. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 308, 37–51.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–656.

Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0

Murphy, G. (1950). *Similitude in Engineering*. New York, NY: Ronald.

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 6, 177–190. doi: 10.1038/nrn1626

Perea, M., and Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *J. Mem. Lang.* 51, 231–246. doi: 10.1027/1618-3169/a000029

Savage-Rumbaugh, E. S., Rumbaugh, D. M., and Boysen, S. (1978). Symbolic communication between two chimpanzees (Pan troglodytes). *Science* 201, 641–644. doi: 10.1126/science.675251

Scarf, D., Boy, K., Reinert, A. U., Devine, J., Güntürkün, O., and Colombo, M. (2016). Orthographic processing in pigeons (*Columba livia*). *Proc. Natl. Acad. Sci. U.S.A.* 113, 11272–11276.

Scarf, D., and Colombo, M. (2011). Knowledge of the ordinal position of list items in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 37, 483–487. doi: 10.1037/a0023695

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334, 1664–1664. doi: 10.1126/science.1213357

Scarf, D., Johnston, M., and Colombo, M. (2018). Pigeons (*Columba livia*) learn a four-item list by trial and error. *J. Comp. Psychol.* 132, 234–239. doi: 10.1037/com0000124

Taylor, A. H. (2014). Corvid Cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 361–372. doi: 10.1002/wcs.1286

Tiffin-Richards, S. P., and Schroeder, S. (2015). Children's and adults' parafoveal processes in German: phonological and orthographic effects. *J. Cogn. Psychol.* 27, 531–548. doi: 10.1080/20445911.2014.999076

Vinckier, F., Qiao, E., Pallier, C., Dehaene, S., and Cohen, L. (2011). The impact of letter spacing on reading: a test of the bigram coding hypothesis. *J. Vis.* 11, 1–21. doi: 10.1167/11.6.8

Weber, E. H. (1834). *De Tactu*. Leipzig: Koehler.

Ziegler, J. C., Hannagan, T., Dufau, S., Montant, M., Fagot, J., and Grainger, J. (2013). Transposed-letter effects reveal orthographic processing in baboons. *Psychol. Sci.* 24, 1609–1611. doi: 10.1177/0956797612474322

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 31 January 2020
doi: 10.3389/fpsyg.2020.00055

Searching for the Critical *p* of Macphail's Null Hypothesis: The Contribution of Numerical Abilities of Fish

Maria Elena Miletto Petrazzini[1], Alessandra Pecunioso[2], Marco Dadda[2] and Christian Agrillo[2,3]*

[1]School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom, [2]Department of General Psychology, University of Padova, Padua, Italy, [3]Padua Neuroscience Center, University of Padova, Padua, Italy

In 1985, Macphail argued that there are no differences among the intellects of non-human vertebrates and that humans display unique cognitive skills because of language. Mathematical abilities represent one of the most sophisticated cognitive skills. While it is unquestionable that humans exhibit impressive mathematical skills associated with language, a large body of experimental evidence suggests that Macphail hypothesis must be refined in this field. In particular, the evidence that also small-brained organisms, such as fish, are capable of processing numerical information challenges the idea that humans display unique cognitive skills. Like humans, fish may take advantage of using continuous quantities (such as the area occupied by the objects) as proxy of number to select the larger/smaller group. Fish and humans also showed interesting similarities in the strategy adopted to learn a numerical rule. Collective intelligence in numerical estimation has been also observed in humans and guppies. However, numerical acuity in humans is considerably higher than that reported in any fish species investigated, suggesting that quantitative but not qualitative differences do exist between humans and fish. Lastly, while it is clear that contextual factors play an important role in the performance of numerical tasks, inter-species variability can be found also when different fish species were tested in comparable conditions, a fact that does not align with the null hypothesis of vertebrate intelligence. Taken together, we believe that the recent evidence of numerical abilities in fish call for a deeper reflection of Macphail's hypothesis.

Keywords: fish, counting, non-symbolic numerical abilities, approximate number system, inter-species differences

Edited by:
Damian Scarf,
University of Otago,
New Zealand

Reviewed by:
Zhanna Reznikova,
Institute of Systematics and Ecology of Animals (RAS), Russia
Michael Colombo,
University of Otago, New Zealand

****Correspondence:***
Maria Elena Miletto Petrazzini
m.e.milettopetrazzini@qmul.ac.uk

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: 15 November 2019
Accepted: 09 January 2020
Published: 31 January 2020

Citation:
Miletto Petrazzini ME, Pecunioso A, Dadda M and Agrillo C (2020) Searching for the Critical p of Macphail's Null Hypothesis: The Contribution of Numerical Abilities of Fish. Front. Psychol. 11:55. doi: 10.3389/fpsyg.2020.00055

INTRODUCTION

The capacity to process numerical information represents one of the most sophisticated cognitive skills in our species. Studies on individuals living in non-Western societies with a limited vocabulary for numbers showed that an adult human brain *per se* is not enough to elaborate complex mathematical skills. Culture and language play a fundamental role in developing abstract numerical competence (Dehaene et al., 2008). For instance, native speakers of Mundurukú have a limited vocabulary for numbers (only for the numbers 1 through 5). This Amazonian indigenous group proved to have an exact arithmetic with numbers smaller than 5. However, they are also able to compare and add large numbers far beyond their naming range, showing the existence of "non-symbolic" numerical abilities that are approximate and independent from language and culture (Pica et al., 2004). Apart from cross-cultural studies, developmental (Izard et al., 2009) and cognitive (Revkin et al., 2008) psychology also showed the existence of non-symbolic numerical abilities. These cognitive skills are supposed to be shared with other vertebrates (Feigenson et al., 2004; Beran, 2008). Rudimentary numerical abilities in animals have been reported since the 1930s, mainly mammals and birds (Koehler, 1941; Hauser et al., 1996; Brannon et al., 2001; Agrillo and Bisazza, 2018). Especially, the capacity to discriminate the larger/smaller group of biologically relevant items is supposed to solve most of the quantitative problems encountered in nature (e.g., select the most advantageous group of food items, sexual mates, or social companions).

The first evidence of numerical abilities in cold-blooded vertebrates was provided by Uller et al. (2003), studying amphibians. Since then, we have witnessed an increase in the publications on this group of vertebrates, mainly represented by studies on fish (reviewed in Agrillo and Bisazza, 2018). The discovery that small-brained species that lack cortex, such as fish, display similar numerical abilities described in humans represents a true challenge to the hypothesis advanced by Macphail (1985). In his seminal paper, the author argued that *there are neither quantitative nor qualitative differences among intellects of non-human vertebrates* (p. 37). Also, he claimed that *man's intellectual superiority may be due solely to our possession of a species-specific language-acquisition device* (p. 37). Any evidence in support of a surprising similarity in numerical abilities of humans and fish would be an argumentation against the humans' superiority of cognitive skills advanced by Macphail (1985). Indeed, fish represent the vertebrate group more distantly related to humans, as fish and land vertebrates diverged approximately 450 million years ago. The structure of the brain is largely different in terms of size and neural circuits. Besides these aspects, the aquatic environment is clearly incomparable with the dry land occupied by primates (and most mammals in general), a fact that is likely to have differently impacted on the selective pressures that shaped cognitive skills. Lastly, fish represent approximately half of vertebrate species. Most of these species occupy very different ecological niches, ranging from dense environments of shallow waters of the rivers to empty environments in the deep waters of the oceans. In this sense, they represent the ideal vertebrate group to study the existence of interspecies differences, a fact that would contrast with the null hypothesis of vertebrate intelligence.

In this work, we review the literature of numerical abilities of fish analyzed under the prism of Macphail's argumentations. The first part of the work will be devoted to outlining the evidence against the null hypothesis; the second part will summarize why we should not reject this hypothesis. Lastly, we will suggest some future directions necessary to form a broader comprehension in this field.

REJECTING THE NULL HYPOTHESIS ($P < 0.05$)

In this section, we will split our argumentations in two main directions, starting from the statements made by Macphail in 1985: the absence of difference in cognitive abilities of animals and the superiority of humans' intellectual skills.

Neither Quantitative Nor Qualitative Differences Among Vertebrates

According to Macphail, similar cognitive abilities among the species are expected. However, data coming from numerical cognition studies in fish do not support this view. There is evidence that numerical acuity is different across the species. This is clear, for instance, in the different ability of teleost fish to select the larger shoal when exploring a novel and potentially dangerous environment. Such ability is supposed to be highly useful in nature to reduce the risks of being predated. It has been shown that the capacity to discriminate between a large and a small shoal varies as a function of the species: when the two shoals differ by one unit, angelfish seem to be able to find the larger shoal up to 3 units (2 vs. 3, Gómez-Laplaza and Gerlai, 2011), mosquitofish up to 4 (3 vs. 4, Agrillo et al., 2008), guppies up to 5 (4 vs. 5, Lucon-Xiccato et al., 2017), while stickleback seem to be able to discriminate even 6 from 7 conspecifics (Mehlis et al., 2015). As these species are highly social, it is unlikely that the variability here observed could be explained by the different degree of motivation in reaching social companions. Also, one may argue that such differences are the results of different stimuli and procedures. There is indeed evidence that the precision in numerical tasks is affected by the experimental procedure adopted (Gatto et al., 2017). For instance, the capacity to discriminate two vs. three social companions in goldbelly topminnows depends on the type of stimulus presentation (two shoals presented on the same side of the tank vs. two shoals presented on the opposite sides of the tank, Agrillo and Dadda, 2007). This is exactly what Macphail (1985) was referring to about the difficulty to establish if the different performance reported among vertebrates actually reflects true inter-species differences in cognitive skills or instead reflects the consequence of contextual variables, such as the type of methodology used.

For this reason, a fine comparative study of the numerical ability of animals should take into account this issue, reducing the methodological variability among the species. To tackle this problem, Agrillo et al. (2012a) tested numerical acuity of five different teleost fish using the same stimuli, apparatus, and procedure. Two sets of two-dimensional figures of different numerosities were presented at the opposite ends. Food was provided only near the stimulus to be reinforced. The proportion of time spent near the positive stimulus in probe trials without food reward was used as a dependent variable. This training procedure was applied to five different fish species: redtail splitfin (*Xenotoca eiseni*), guppies (*Poecilia reticulata*), zebrafish (*Danio rerio*), angelfish (*Pterophyllum scalare*), and Siamese fighting fish (*Betta splendens*). The same visual patterns were presented to all subjects. In one experiment, subjects were initially trained on a 0.5 ratio (5 vs. 10 and 6 vs. 12 figures). For instance, they were required to select the larger array to receive a food reward. Once they reached the learning criterion, they were presented with novel numerical contrasts with harder numerical ratios: 0.67 (8 vs. 12) and 0.75 (9 vs. 12). In another experiment, after reaching the learning criterion, they were observed in their capacity to generalize the numerical rule to very small (2 vs. 4) or very large (25 vs. 50) numerical contrasts. Overall, fish proved able to generalize the learned rule to harder numerical contrasts (0.67 ratio but not 0.75) and were able to generalize it to a smaller set of items (2 vs. 4 but not 25 vs. 50). However, a deeper analysis of fish performance suggested at least two main inter-specific differences: angelfish was not able to discriminate between 8 and 12 items, suggesting a lower numerical acuity. Similarly, the performance of zebrafish was lower in terms of proportion of individuals that reached the learning criterion. Although alternative explanations were also taken into account by the authors, these two results leave open the concrete possibility that quantitative differences exist in the cognitive processes underlying numerical estimation of fish. This hypothesis is further supported by a study on a blind cavefish (*Phreatichthys andruzzii*) that evolved for approximately 2 million years in the phreatic layer of the Somalia desert (Bisazza et al., 2014a). As they lack visual modality, the training procedure adopted in the previous inter-specific study was partially modified to include three-dimensional stimuli submerged in the tank (instead of two-dimensional figures). The subjects were trained to discriminate between two groups of sticks placed in opposite positions of the experimental tank in order to receive a food reward. Cavefish showed the ability to discriminate accurately two vs. four objects but not two vs. three. This indicates that the brains of fish species that live in a very peculiar ecological niche (dark caves with no predators) are still equipped with neural circuits that support numerical processing. However, it is worth noting that cavefish showed lower performance in terms of numerical acuity compared to the majority of fish species investigated (that commonly discriminate 0.67 ratio, e.g., Agrillo et al., 2012a,b). At least three main hypotheses have been advanced: Provided that cavefish cannot use a visual modality to solve these tasks, the most likely explanation is that they used the lateral line, a sense organ typical of fish, which is integral to detecting movement, vibration, and pressure gradients in water. It is possible that object representation through lateral line might be less precise. If so, cavefish might have the same numerical acuity of other species but exhibit a worse performance because of a general noise in detecting the items to be enumerated. Another possibility is that cognitive numerical systems might be more accurate in the visual modality. Tokita et al. (2013) found in human participants a different performance in numerosity judgments tested in visual and auditory conditions, advancing the idea of multiple numerical systems—with different degrees of precision—related to the different sensory modalities. Lastly, it is possible that the peculiarity of the ecological niche plays an important role in shaping numerical systems. This species evolved for millions of years in a homogeneous environment with a scarcity of food resources and without natural predators. Selective pressures might have acted reducing the cerebral mass in order to optimize the metabolic consumption of the brain, thus lessening also the neural circuits supporting cognitive functions not useful in a cave's life.

In sum, the comparative investigation of fish species tested with reduced methodological variability (similar apparatuses, stimuli, and procedures) provided enough experimental material to argue that the assumption of no inter-specific differences among the species can be hardly sustained, at least with respect to numerical cognition.

Man's Intellectual Superiority

Humans are clearly very precise in numerical tasks compared to fish (and presumably to all other animals, see Section "Failing to Reject the Null Hypothesis ($p > 0.05$)"). However, if a superiority does exist in absolute terms, it is expected to emerge also in issues other than numerical acuities, such as the cognitive mechanisms used to estimate quantities.

It is known than numerosity co-varies with several other physical attributes of the stimuli, also known as "continuous quantities," such as cumulative surface area (i.e., the sum of areas of the items to be enumerated), density, and convex hull (the overall space occupied by the most lateral items of the array). There is evidence that humans involved in non-symbolic numerical tasks can establish which group of objects is larger by using a combination of discrete (numerical) and continuous information (Gebuis and Reynvoet, 2012a,b; see Leibovich et al., 2017 for a review about). In short, when comparing two groups of three and four circles, we would extrapolate both the numerosity of items and the associated continuous quantities. The capacity to discriminate the larger/smaller group would be the result of this number-space interplay.

There is evidence that also fish can process both numerical and continuous quantities. A decade ago, Agrillo et al. (2009) provided the first evidence that fish can use numerical information also when all continuous quantities were controlled for. Mosquitofish (*Gambusia holbrooki*) were placed in an unfamiliar environment. To re-join their social companions, subjects were required to select one of two identical tunnels at opposite corners. The correct tunnel was associated with a specific number of items (either two or three) presented above the tunnels. The shapes and spatial arrangements of the figures were changed across the trials to prevent the fish from learning to recognize specific patterns. Furthermore, the items were controlled for continuous quantities so that the only discriminative cue was numerical information. Subjects proved able to solve the task, indicating the use of numerical information by fish. To date, we know that at least eight fish species can process numerical information in different experimental contexts (Agrillo and Bisazza, 2018).

Fish, however, can also use continuous quantities. Agrillo et al. (2009) set up an experiment in which the fish were trained to discriminate between two and three figures in a condition in which the number and continuous quantities were simultaneously available. For example, the larger group occupied also the larger area. In the test phase, researchers controlled for one continuous quantity at a time and observed the performance of mosquitofish: accuracy decreased when the stimuli were matched for the cumulative surface area or the convex hull, indicating that these cues had been used during the tasks. The combination of these continuous quantities is exactly what has been advanced as an important mechanism for human numerical estimation. According to the occupancy model (Allik and Tuulmets, 1991), numerosity estimation is linearly related to the total area occupied (occupancy) by virtual disks that circumscribe each dot. When dots are close to one other, the virtual disks overlap, leading to an underestimation of the dots; when the dots are more distant, the overall space occupied by these disks is larger, leading to an overestimation of the dots. Therefore, the combination of cumulative surface area and inter-item distance (a parameter that is linearly related to the convex hull) seems to influence numerical estimation of both humans and mosquitofish.

The use of discrete and continuous quantities has been reported not only in the presence of neutral laboratory stimuli (such as two-dimensional stimuli) but also with biologically relevant stimuli. For what concern discrete information, Dadda et al. (2009) found that mosquitofish can select the larger shoal also when stimulus fish were presented one at a time and hence are required to sum the number of fish contained in each shoal. Similar capacity was later observed in newborn guppies (Bisazza et al., 2010), suggesting the existence of inborn numerical abilities in fish similar to that described in human infants (Izard et al., 2009). Continuous quantity discrimination in a highly ecological context was studied by Lucon-Xiccato et al. (2015). The authors found that guppies are able to select the larger piece of food when the ratio between the smaller and larger piece is 0.75.

Clustering is another perceptual cue that affects non-symbolic numerical tasks. Humans tend to overestimate the number of items if they are arranged to form a single Gestalt. This is particularly evident in the Solitaire illusion studied by Frith and Frith (1972), a visual pattern in which items forming a single Gestalt is overestimated compared to the same number of items arranged in separate (smaller) clusters. Perception of the Solitaire illusion has been recently studied in fish. Guppies were trained to select an array containing a larger quantity of black dots in the presence of two arrays made by white and black dots. After reaching the learning criterion, subjects were presented with two illusory arrangements: One array presented 16 black dots centrally located to form a single Gestalt and 16 white dots on the perimeter to form 4 separate clusters; the other presented 16 white dots centrally located with 16 black dots on the perimeter. If the subjects perceived the illusion, they were expected to select the array in which the black dots were centrally located (as they appear to be larger to human observers). Although higher inter-individual variability was found in fish compared to humans (Agrillo et al., 2016; Pecunioso and Agrillo, 2019), guppies exhibited a human-like susceptibility to this numerosity illusion, suggesting that clustering of items is a further common mechanism used by both humans and fish to estimate the number of items in the visual scene.

It is important to clarify that humans appear to be equally able to use numerical information over continuous quantities (Hurewitz et al., 2006). One may argue that animals might find it more difficult to process numerical information than continuous quantities. This was indeed the idea advanced by different authors in the 1980s (Davis and Memmott, 1982; Davis and Perusse, 1988) that led to the hypothesis of numerical information as "last-resort strategy" used only when no other continuous quantity would permit an animal to discriminate which group is larger/smaller. A study by Agrillo et al. (2011) does not encourage to this view. Three groups of mosquitofish were trained in different conditions: In one condition, the mosquitofish could use only numerical information to distinguish between the quantities (2 vs. 3, "numerical" condition). In the second condition, fish could use only continuous quantities (1 vs. 1, the ratio between the areas was equal to two-thirds, the "continuous quantity" condition). In the third condition, both numerical and continuous information was available (2 vs. 3, with the larger group occupying more space, "number and continuous quantity" condition). If numerical information were more cognitively demanding, subjects were expected to need more trials to learn the task in the first condition than in the other two conditions. As expected, higher performance was found when fish could use both numerical and continuous quantities as the presence of multiple cues is supposed to represent the easiest (and the most ecological) condition (Gebuis and Reynvoet, 2012a,b). However, no difference was found between the numerical condition and the continuous quantity condition, suggesting that, at least for mosquitofish, processing numbers is not more complex than processing continuous quantities. After all, artificial neural networks suggest that numerosity estimation does not enroll a large neural network Hope et al. (2010), found that fewer than 25 units might be enough for a system to represent quantity with a performance comparable to that observed in fish (Agrillo et al., 2008). This is also supported by a more recent study (Stoianov and Zorzi, 2012), showing that as few as 35 hidden neurons were able to spontaneously extract numerical information in a visual scene. In this sense, it is not surprising that also a fish brain can apparently use number with the same cognitive effort used in continuous quantity discrimination.

The cognitive strategy used to learn a numerical rule is a further aspect that must be taken into consideration to establish similarities and differences between humans and fish. It is known that, when animals learn to select the larger of two arrays (e.g., 5 vs. 10), they might potentially use two alternative strategies. One strategy consists in learning to always select the array containing 10 items ("absolute numerical rule"). The other strategy consists of assessing which group is larger and smaller in order to "select the larger numerosity of each stimulus pair" ("relative numerical rule"). Because the behavioral output is the same, the exact cognitive strategy used by animals is often neglected. Miletto Petrazzini et al. (2016) dissociated the two hypotheses by training angelfish to discriminate between two arrays of figures differing in numerosity. One group of subjects was required to select 10 items in a 5 vs. 10 discrimination; the other group were required to select 10 items in the 10 vs. 20 discrimination. After reaching the learning criterion, the former group was presented with a 10 vs. 20 numerical contrast. If subjects had learned the task by using a relative numerical rule, they should have selected the novel larger numerosity (20); otherwise, if angelfish had used an absolute numerical rule, they were expected to select the numerosity previously reinforced (10). The other experimental group (10 vs. 20) was presented in test trials with 5 vs. 10 discrimination. Angelfish belonging to both groups spontaneously used a relative numerical rule, selecting the novel numerosity instead of the previously reinforced numerosity.

Interestingly, the authors also tested undergraduate students in the same task (Miletto Petrazzini et al., 2016). In order to observe the spontaneous use of a relative vs. absolute numerical rule, no verbal instructions were provided so that participants had to infer the numerical rule only by the feedback, exactly like fish. Humans used a relative numerosity rule too with very limited inter-individual variability. This implies that distantly related species share similar cognitive systems for making decisions about quantities, a fact that does not properly align with the idea of any kind of human's superiority in terms of qualitative differences.

The similarities between humans and fish are not confined to the performance of the two species individually tested in cognitive tasks. It is known that interacting people can generally achieve more accurate decisions than single individuals. Although this is an open debate (e.g., Cantlon et al., 2006), it was suggested to occur also in numerical tasks. In a study by Bahrami et al. (2013), pairs of participants made both individual and collective estimations of which group of dots was larger. In the "collective enumeration" condition, they could negotiate joint decisions *via* verbal communication and received feedback about accuracy at the end of each trial. Results showed that two individuals collectively estimate the number of dots better than either one alone. Collective intelligence in non-human animals has been reported in different fields (Krause et al., 2010). However, although several species showed impressive numerical skills, including invertebrates (e.g., eusocial ants: Reznikova and Ryabko, 2011; Reznikova, 2017; bees: Pahl et al., 2013), no evidence of an advantage in collective enumeration was reported in non-human animals before 2014. Bisazza et al. (2014b) investigated this issue in fish. Guppies were observed in their spontaneous preference of joining the larger shoal (exp. 1) and in their capacity to learn a numerical rule after operant conditioning (exp. 2). Subjects' performance was observed both when they were singly inserted in the experimental apparatus and when they were inserted in pairs. In both experiments, interacting guppies achieved a superior level of numerosity discrimination compared to the average ability of the isolated individual fish. Even though the reasons underlying the enhanced cognitive performance of interacting guppies are unknown, the result is intriguing as it suggests that the well-known collective intelligence that has been advanced in humans (Bahrami et al., 2013) can be traced also in a fish species.

In summary, we believe that all the above-mentioned studies provide a robust argumentation to say that the concept of *Man's intellectual superiority* need to be deeply revised.

FAILING TO REJECT THE NULL HYPOTHESIS (*P* > 0.05)

Here we will delineate why we believe that Macphail's argumentation still holds in numerical cognition.

Man's Intellectual Superiority

As said in the "Introduction" section, it is unquestionable that the capacity of humans to process numerical information represents a clear example of high cognitive functions related to a species-specific language-acquisition device. However, we also display numerical abilities that are not related to language, a cognitive skill particularly evident when we are forced to estimate which group of objects is more numerous without the possibility to see the two groups long enough to count the objects (e.g., only 150–250 ms; Halberda et al., 2008; Revkin et al., 2008).

The comparison of non-symbolic numerical abilities of humans and animals clearly indicates that humans are more precise even in this numerical skill. Humans can discriminate a 0.90 ratio (9 vs. 10, Halberda et al., 2008), while numerical acuity of other species is often more limited (Hauser et al., 1996; Uller et al., 2003; Rugani et al., 2008). It is interesting to note that, although the superiority in numerical acuity of humans supports the null hypothesis, it also contradicts one of the predictions related to the importance of language: *Humans without language would, according to this view, be no more intelligent than non-human vertebrates* (Macphail, 1985, p. 49). Indeed, because non-symbolic numerical tasks prevent the use of verbal counting, one should not expect a higher performance of humans in this task. To tell the true story, participants involved in the studies that showed impressive abilities in non-symbolic numerical tasks were teenagers or university students of Western societies (e.g., Halberda et al., 2008; Agrillo et al., 2014). Even though experimental strategies were taken to limit the use verbal language, we cannot exclude that language and education of subjects positively impacted on the cognitive skills necessary to support numerical estimation, thus improving their performance. In line with this hypothesis, when members of non-Western societies are tested (e.g., Mundurukù: Pica et al., 2004; Warlpiri, and Anindilyakwa: Butterworth et al., 2008), participants' performance in numerical estimation tasks is not far from that observed with several non-human species.

As said in "Rejecting the Null Hypothesis ($p < 0.05$)" Section, the observation of spontaneous behavior showed that sticklebacks can discriminate up to six vs. seven social companions (Mehlis et al., 2015). Other fish species, however, showed a lower performance when the groups to be compared differ by one unit (Agrillo et al., 2012c; Agrillo and Bisazza, 2018); therefore, the high level of performance exhibited by sticklebacks is not likely to reflect the average precision of fish in quantitative tasks. Trained fish guppies can also reach surprising performances (e.g., the capacity to discriminate up to 0.75 ratio, Bisazza et al., 2014c; Miletto Petrazzini et al., 2015a), but no study showed the capacity to discriminate up to a 0.90 ratio.

Humans' superiority in numerical tasks extends far beyond relative numerosity judgments. Ordinal abilities are the capacity to understand that "3″ is larger than "2″ and smaller than "4." This ability permits to solve several numerical tasks, including the capacity to locate an object on the basis of its position in a sequence of other objects. Unlike humans, fish showed a very limited ability to use ordinal information. Miletto Petrazzini et al. (2015b) trained guppies to select the third feeder in a row of eight alternative feeders placed perpendicularly in front of them. The inter-feeder distance was experimentally manipulated between trials to avoid the use of continuous quantities, such as the overall distance necessary to reach the correct feeder. The guppies solved the task, thereby providing the first evidence that ordinal abilities exist in fish species. However, in another experiment, researchers placed the correct feeder in the fifth position: In this case, the performance was no longer significant, showing a clear limit in using ordering information that does not exceed 3–4 units.

As said, the literature on fish in this section is just an example, since humans outperform mammals, birds, amphibians, reptiles, and fish in non-symbolic numerical tasks. All this literature clearly indicates that, although animals have recently shown evidence of impressive numerical abilities to demonstrate no qualitative differences, quantitative differences seem to exist between humans and animals.

CONCLUSIONS AND FUTURE REMARKS

We reviewed the literature on fish numerical cognition as a tool to shed light on the modernity of Macphail's argumentations. We believe that most of the evidence collected in numerical tasks of fish call for a deep reframing of the null hypothesis of vertebrate intelligence. At least two different bodies of experimental evidence support our dissertation: (1) it is not true that numerical abilities in fish did not differ among the species and (2) most of the literature speaks in favor of qualitative similarities between humans and fish. Evidence supporting the first claim comes from studies in which numerical abilities of fish were compared with the same experimental material/procedure. These studies showed that, although similarities are greater than differences, inter-species differences exist among fish. The latter claim is supported by studies showing a similar use of discrete and continuous quantities in human and fish, by the observation of comparable cognitive strategy to learn a numerical rule and by the evidence of an enhanced numerical performance when multiple individuals are involved in the numerical task.

However, rejecting this hypothesis might be precocious at this stage. The null hypothesis of Macphail still holds if one considers a crucial aspect of numerical abilities, the precision of numerical estimation. Although in the last two decades several studies showed impressive numerical abilities in fish (and in animals in general), the higher performance is repeatedly reported in humans, even in tasks in which they are prevented to use verbal counting. Therefore, if any clear difference exists between human and fish, such difference is quantitative but not qualitative.

That said, it is important to specify that the procedure in human and animal studies often differs for a fundamental aspect: The presence/absence of verbal instructions. As known, animals have to infer the rule trial by trial, while most of human studies are often introduced by verbal instructions (Halberda et al., 2008; Revkin et al., 2008; Price et al., 2012). This permits participants to focus on the most relevant aspects of the experiment since the beginning, providing a potential advantage that might be misinterpreted as higher numerical abilities. Only recently researchers have begun to take into account this potential confound and present human participants with tasks with no verbal instructions (Beran, 2006; Miletto Petrazzini et al., 2016; Parrish et al., 2019).

Some important issues need to be investigated. To better understand the similarities between humans and fish it would be important to assess whether fish display an abstract concept of number. We know that humans can compare quantities of objects presented in different sensory modalities (e.g., three lights and three sounds). The capacity to transfer numerical information from the visual to the acoustic modality is important evidence of an abstract concept of number. To date, existing studies in fish reported the capacity to generalize the numerical rule to novel stimuli presented through the same sensory modality (e.g., visual stimuli; reviewed in Agrillo et al., 2017). No study has established whether fish can transfer numerical information from one sensory modality to another, a fact that prevents to understand whether the complexity of abstract numerical representation is similar or not in human and fish. Also, the investigation of continuous quantities used by fish is limited to a few species (Agrillo et al., 2009, 2010, 2011; Gómez-Laplaza and Gerlai, 2013; Miletto Petrazzini et al., 2018). In order to understand whether the cognitive mechanisms used by human and fish are similar, we need to enlarge the number of species under investigation. Lastly, the spatial representation of numbers is another important aspect that should be considered to comprehend whether fish have human-like mechanisms of number processing. It is known that most humans represent numbers aligned from left to right, the so-called "mental number line" (Galton, 1880; Zorzi et al., 2002). There is evidence that also birds have a similar spatial representation of numbers (Rugani et al., 2015), but this issue has never been investigated in fish.

In 1985 Macphail said, *In common with all scientific hypotheses, this null hypothesis cannot be proved, only disproved; support for the hypothesis will grow as the number of failures to disprove it increases* (p. 46). After more than three decades, it is still difficult to reach a verdict on the hypothesis advanced by Macphail. We believe, however, that Macphail adopted a questionable statistical approach: when he introduced the idea of a "null hypothesis," he indirectly assumed an all-or-none approach to this issue, like the frequentist *p* approach based on rejecting/failing to reject the null hypothesis. However, this statistical approach can barely grasp all the shades of this issue. For instance, how much do fish differ from other vertebrates, humans included, in numerical skills? If we look at numerical acuity of humans and fish, we would be inclined to assume that the null hypothesis is correct; if we look at qualitative similarities among the species we would be tempted to reject this null hypothesis. Is the null hypothesis corroborated or not? Instead of assuming that a dichotomic response may exist in this issue, we believe that the Bayesian approach would be more appropriate. Bayes factors actually enable researchers to estimate the relative strength of the evidence for two competing hypotheses (Dienes, 2014). Even supposing that the next decade will be characterized by the development of finer comparative methodologies, we believe that researchers, at best, could try to establish how likely is the null hypothesis of vertebrate intelligence over the alternative one.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research was funded by PRIN grant no. 2015′ (prot.: 2015FFATB7) from "Ministero dell'Istruzione, Universita' e Ricerca" (MIUR, Italy) to CA and by Marie Skłodowska-Curie Individual Fellowship (750200) from the European Union's Horizon 2020 to MP. The present work was carried out within the scope of the research program Dipartimenti di Eccellenza (art.1, commi 314-337 legge 232/2016), which was supported by a grant from MIUR to the Department of General Psychology, University of Padua.

REFERENCES

Agrillo, C., and Bisazza, A. (2018). Understanding the origin of number sense: a review of fish studies. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 373:20160511. doi: 10.1098/rstb.2016.0511

Agrillo, C., and Dadda, M. (2007). Discrimination of the larger shoal in the poeciliid fish *Girardinus falcatus*. *Ethol. Ecol. Evol.* 19, 145–157. doi: 10.1080/08927014.2007.9522574

Agrillo, C., Dadda, M., Serena, G., and Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Anim. Cogn.* 11, 495–503. doi: 10.1007/s10071-008-0140-9

Agrillo, C., Dadda, M., Serena, G., and Bisazza, A. (2009). Use of number by fish. *PLoS One* 4:e4786. doi: 10.1371/journal.pone.0004786

Agrillo, C., Miletto Petrazzini, M. E., and Bisazza, A. (2017). Numerical abilities in fish: a methodological review. *Behav. Process.* 141, 161–171. doi: 10.1016/j.beproc.2017.02.001

Agrillo, C., Miletto Petrazzini, M. E., Piffer, L., Dadda, M., and Bisazza, A. (2012b). A new training procedure for studying discrimination learning in fishes. *Behav. Brain Res.* 230, 343–348. doi: 10.1016/j.bbr.2012.02.025

Agrillo, C., Miletto Petrazzini, M. E., Tagliapietra, C., and Bisazza, A. (2012a). Inter-specific differences in numerical abilities among teleost fish. *Front. Psychol.* 3:483. doi: 10.3389/fpsyg.2012.00483

Agrillo, C., Parrish, A. E., and Beran, M. J. (2014). Do primates see the solitaire illusion differently? A comparative assessment of humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), rhesus monkeys (*Macaca mulatta*), and capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 128, 402–413. doi: 10.1037/a0037499

Agrillo, C., Parrish, A. E., and Beran, M. J. (2016). How illusory is the solitaire illusion? Assessing the degree of misperception of numerosity in adult humans. *Front. Psychol.* 7:1663. doi: 10.3389/fpsyg.2016.01663

Agrillo, C., Piffer, L., and Bisazza, A. (2010). Large number discrimination by fish. *PLoS One* 5:e15232. doi: 10.1371/journal.pone.0015232

Agrillo, C., Piffer, L., and Bisazza, A. (2011). Number versus continuous quantity in numerosity judgments by fish. *Cognition* 119, 281–287. doi: 10.1016/j.cognition.2010.10.022

Agrillo, C., Piffer, L., Bisazza, A., and Butterworth, B. (2012c). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One* 7:e31923. doi: 10.1371/journal.pone.0031923

Allik, J., and Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Percept. Psychophys.* 49, 303–314. doi: 10.3758/BF03205986

Bahrami, B., Didino, D., Frith, C., Butterworth, B., and Rees, G. (2013). Collective enumeration. *J. Exp. Psychol. Hum. Percept. Perf.* 39, 338–347. doi: 10.1037/a0029717

Beran, M. J. (2006). Quantity perception by adult humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*) and rhesus macaques (*Macaca mulatta*) as a function of stimulus organization. *Int. J. Comp. Psychol.* 19, 386–397. https://escholarship.org/uc/item/76n9h06b

Beran, M. J. (2008). The evolutionary and developmental foundations of mathematics. *PLoS Biol.* 6:e19. doi: 10.1371/journal.pbio.0060019

Bisazza, A., Butterworth, B., Piffer, L., Bahrami, B., Miletto Petrazzini, M. E., and Agrillo, C. (2014b). Collective enhancement of numerical acuity by meritocratic leadership in fish. *Sci. Rep.* 4:4560. doi: 10.1038/srep04560

Bisazza, A., Agrillo, C., Lucon-Xiccato, T. (2014c). Extensive trainingextends numerical abilities of guppies. *Anim Cogn.* 17, 1413–1419. doi: 10.1007/s10071-014-0759-7

Bisazza, A., Serena, G., Piffer, L., and Agrillo, C. (2010). Ontogeny of numerical abilities in guppies. *PLoS One* 5:e15516. doi: 10.1371/journal.pone.0015516

Bisazza, A., Tagliapietra, C., Bertolucci, C., Foà, A., and Agrillo, C. (2014a). Non-visual numerical discrimination in a blind cavefish (*Phreatichthys andruzzii*). *J. Exp. Biol.* 217, 1902–1909. doi: 10.1242/jeb.101683

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., and Gibbon, J. (2001). Numerical subtraction in the pigeon: evidence for a linear subjective number scale. *Psychol. Sci.* 12, 238–243. doi: 10.1111/1467-9280.00342

Butterworth, B., Reeve, R., Reynolds, F., and Lloyd, D. (2008). Numerical thought with and without words: evidence from indigenous Australian children. *Proc. Natl. Acad. Sci. USA* 105, 13179–13184. doi: 10.1073/pnas.0806045105

Cantlon, J. F., Brannon, E. M., Carter, E. J., and Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biol.* 4:e125. doi: 10.1371/journal.pbio.0040125

Dadda, M., Piffer, L., Agrillo, C., and Bisazza, A. (2009). Spontaneous number representation in mosquitofish. *Cognition* 112, 343–348. doi: 10.1016/j.cognition.2009.05.009

Davis, H., and Memmott, J. (1982). Counting behavior in animals: a critical evaluation. *Psychol. Bull.* 92, 547–571. doi: 10.1037/0033-2909.92.3.547

Davis, H., and Perusse, R. (1988). Numerical competence in animals: definitional issues, current evidence and a new research agenda. *Behav. Brain Sci.* 11, 561–579. doi: 10.1017/S0140525X00053437

Dehaene, S., Izard, V., Spelke, E., and Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science* 320, 1217–1220. doi: 10.1126/science.1156540

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Front. Psychol.* 5:781. doi: 10.3389/fpsyg.2014.00781

Feigenson, L., Dehaene, S., and Spelke, E. S. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

Frith, C. D., and Frith, U. (1972). The solitaire illusion: an illusion of numerosity. *Percept. Psychophys.* 11, 409–410. doi: 10.3758/BF03206279

Galton, F. (1880). Visualised numerals. *Nature* 21, 252–256. doi: 10.1038/021252a0

Gatto, E., Lucon-Xiccato, T., Savasci, B. B., Dadda, M., and Bisazza, A. (2017). Experimental setting affects the performance of guppies in a numerical discrimination task. *Anim. Cogn.* 20, 187–198. doi: 10.1007/s10071-016-1037-7

Gebuis, T., and Reynvoet, B. (2012a). The role of visual information in numerosity estimation. *PLoS One* 7:e37426. doi: 10.1371/journal.pone.0037426

Gebuis, T., and Reynvoet, B. (2012b). The interplay between nonsymbolic number and its continuous visual properties. *J. Exp. Psych. Gen.* 141, 642–648. doi: 10.1037/a0026218

Gómez-Laplaza, L. M., and Gerlai, R. (2011). Spontaneous discrimination of small quantities: shoaling preferences in angelfish (*Pterophyllum scalare*). *Anim. Cogn.* 14, 565–574. doi: 10.1007/s10071-011-0392-7

Gómez-Laplaza, L. M., and Gerlai, R. (2013). The role of body surface area in quantity discrimination in angelfish (*Pterophyllum scalare*). *PLoS One* 8:e83880. doi: 10.1371/journal.pone.0083880

Halberda, J., Mazzocco, M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246

Hauser, M. D., MacNeilage, P., and Ware, M. (1996). Numerical representations in primates. *Proc. Natl. Acad. Sci. USA* 93, 1514–1517.

Hope, T., Stoianov, I., and Zorzi, M. (2010). Through neural stimulation to behavior manipulation: a novel method for analyzing dynamical cognitive models. *Cogn. Sci.* 34, 406–433. doi: 10.1111/j.1551-6709.2009.01079.x

Hurewitz, F., Gelman, R., and Schnitzer, B. (2006). Sometimes area counts more than number. *Proc. Natl. Acad. Sci. USA* 103, 19599–19604. doi: 10.1073/pnas.0609485103

Izard, V., Sann, C., Spelke, E. S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc. Natl. Acad. Sci. USA* 106, 10382–10385. doi: 10.1073/pnas.0812142106

Koehler, O. (1941). Vom Erlernen unbenannter Anzahlen bei Vögeln. *Naturwissenschaften* 29, 201–218. doi: 10.1007/BF01481755

Krause, J., Ruxton, G. D., and Krause, S. (2010). Swarm intelligence in animals and humans. *Trends Ecol. Evol.* 25, 28–34. doi: 10.1016/j.tree.2009.06.016

Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From 'sense of number' to 'sense ofmagnitude': the role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* 40:E164. doi: 10.1017/S0140525X16000960

Lucon-Xiccato, T., Dadda, M., Gatto, E., and Bisazza, A. (2017). Development and testing of a rapid method for measuring shoal size discrimination. *Anim. Cogn.* 20, 149–157. doi: 10.1007/s10071-016-1050-x

Lucon-Xiccato, T., Miletto Petrazzini, M. E., Agrillo, C., and Bisazza, A. (2015). Guppies (*Poecilia reticulata*) discriminate among two quantities of food items but prioritise item size over total amount. *Anim. Behav.* 107, 183–191. doi: 10.1016/j.anbehav.2015.06.019

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B* 308, 37–51.

Mehlis, M., Thünken, T., Bakker, T. C., and Frommen, J. G. (2015). Quantification acuity in spontaneous shoaling decisions of three-spined sticklebacks. *Anim. Cogn.* 18, 1125–1131. doi: 10.1007/s10071-015-0884-y

Miletto Petrazzini, M. E., Agrillo, C., Izard, V., and Bisazza, A. (2015a). Relative versus absolute numericalrepresentation in fish: can guppies represent'fourness. *Anim. Cogn.* 18, 1007–1017. doi: 10.1007/s10071-015-0868-y

Miletto Petrazzini, M. E., Lucon-Xiccato, T., Agrillo, C., and Bisazza, A. (2015b). Use of ordinal information by fish. *Sci.Rep.* 5, 15497. doi: 10.1038/srep15497

Miletto Petrazzini, M. E., Agrillo, C., Izard, V., and Bisazza, A. (2016). Do humans (*Homo sapiens*) and fish (*Pterophyllum scalare*) make similar numerosity judgments? *J. Comp. Psychol.* 130, 380–390. doi: 10.1037/com0000045

Miletto Petrazzini, M. E., Parrish, A. E., Beran, M. J., and Agrillo, C. (2018). Exploring the solitaire illusion in guppies (*Poecilia reticulata*). *J. Comp. Psychol.* 132, 48–57. doi: 10.1037/com0000092

Pahl, M., Si, A., and Zhang, S. (2013). Numerical cognition in bees and other insects. *Front. Psychol.* 4:162. doi: 10.3389/fpsyg.2013.00162

Parrish, A. E., Beran, M. J., and Agrillo, C. (2019). Linear numerosity illusions in capuchin monkeys (*Sapajus apella*), rhesus macaques (*Macaca mulatta*), and humans (*Homo sapiens*). *Anim. Cogn.* 22, 883–889. doi: 10.1007/s10071-019-01288-9

Pecunioso, A., and Agrillo, C. (2019). Do musicians perceive numerosity illusions differently? *Psychol. Music.* doi: 10.1177/0305735619888804

Pica, P., Lemer, C., Izard, V., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi: 10.1126/science.1102085

Price, G. R., Palmer, D., Battista, C., and Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychol.* 140, 50–57. doi: 10.1016/j.actpsy.2012.02.008

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., and Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychol. Sci.* 19, 607–614. doi: 10.1111/j.1467-9280.2008.02130.x

Reznikova, Z. (2017). *Studying animal languages without translation: An insight from ants*. Cham, Switzerland: Springer International Publishing.

Reznikova, Z., and Ryabko, B. (2011). Numerical competence in animals, with an insight from ants. *Behaviour* 148, 405–434. doi: 10.1163/000579511X568562

Rugani, R., Regolin, L., and Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *J. Exp. Psychol. Anim. B* 34, 388–399. doi: 10.1037/0097-7403.34.3.388

Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379

Stoianov, I., and Zorzi, M. (2012). Emergence of a 'visual number sense' in hierarchical generative models. *Nat. Neurosci.* 15, 194–196. doi: 10.1038/nn.2996

Tokita, M., Ashitani, Y., and Ishiguchi, A. (2013). Is approximate numerical judgment truly modality-independent? Visual, auditory, and cross-modal comparisons. *Atten. Percept. Psychophysiol.* 75, 1852–1861. doi: 10.3758/s13414-013-0526-x

Uller, C., Jaeger, R., Guidry, G., and Martin, C. (2003). Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Anim. Cogn.* 6, 105–112. doi: 10.1007/s10071-003-0167-x

Zorzi, M., Priftis, K., and Umiltá, C. (2002). Neglect disrupts the mental number line. *Nature* 417:138. doi: 10.1038/417138a

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

CORRECTION
published: 06 August 2020
doi: 10.3389/fpsyg.2020.01768

Corrigendum: Searching for the Critical *p* of Macphail's Null Hypothesis: The Contribution of Numerical Abilities of Fish

Maria Elena Miletto Petrazzini[1]*, *Alessandra Pecunioso*[2], *Marco Dadda*[2] and *Christian Agrillo*[2,3]

[1] School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom, [2] Department of General Psychology, University of Padova, Padua, Italy, [3] Padua Neuroscience Center, University of Padova, Padua, Italy

Keywords: fish, counting, non-symbolic numerical abilities, approximate number system, inter-species differences

Approved by:
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

***Correspondence:**
Maria Elena Miletto Petrazzini
m.e.milettopetrazzini@qmul.ac.uk

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *24 June 2020*
Accepted: *26 June 2020*
Published: *06 August 2020*

Citation:
Miletto Petrazzini ME, Pecunioso A, Dadda M and Agrillo C (2020) Corrigendum: Searching for the Critical p of Macphail's Null Hypothesis: The Contribution of Numerical Abilities of Fish. Front. Psychol. 11:1768. doi: 10.3389/fpsyg.2020.01768

A Corrigendum on

Searching for the Critical *p* of Macphail's Null Hypothesis: The Contribution of Numerical Abilities of Fish
by Miletto Petrazzini, M. E., Pecunioso, A., Dadda, M., and Agrillo, C. (2020). Front. Psychol. 11:55. doi: 10.3389/fpsyg.2020.00055

There is an error in the Funding statement. The correct number for Marie Skłodowska-Curie Individual Fellowship from the European Union's Horizon 2020 is 750200.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

frontiers in Psychology

ORIGINAL RESEARCH
published: 03 April 2020
doi: 10.3389/fpsyg.2020.00635

What Human Planning Can Tell Us About Animal Planning: An Empirical Case

*Gema Martin-Ordas**

Department of Psychology, University of Stirling, Stirling, United Kingdom

The ability to think about and plan for the future is a critical cognitive skill for our daily life. There is ongoing debate about whether other animals possess future thinking. Part of the difficulty in resolving this debate is that there is not a definite methodology that allow us to conclude that animals (and human children) are truly thinking about a future event. Research with humans—both children and adults- will benefit the field of comparative psychology by providing information about the range of humans' responses when they are faced with problems similar to those presented to other animals. Inspired by a problem that chimpanzees experienced in the wild, children of 4 and 5 years of age and young adults were presented with a situation in which they were expected to select two tools in order to obtain a reward. More older children than 4 years old successfully obtained the reward. Adults also succeeded at solving the problem. However, both children and adults struggled to select the two correct tools *before* any tool-use action was executed. While children's performance is discussed in the context of temporal components required to envisage future events, adults' performance is interpreted in the context of cognitive effort. These findings link developmental and adult cognition with comparative psychology.

Keywords: planning, tool use, sequence, preschoolers, adults

OPEN ACCESS

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Michael Charles Corballis,
The University of Auckland,
New Zealand
Jonathan Redshaw,
The University of Queensland,
Australia

****Correspondence:***
Gema Martin-Ordas
gema.martin-ordas@stir.ac.uk

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *10 January 2020*
Accepted: *17 March 2020*
Published: *03 April 2020*

Citation:
Martin-Ordas G (2020) What Human Planning Can Tell Us About Animal Planning: An Empirical Case. Front. Psychol. 11:635. doi: 10.3389/fpsyg.2020.00635

INTRODUCTION

There is no doubt that humans can think and plan for the future (e.g., Suddendorf and Corballis, 2007). In fact, we spend an important part of our time mind-wandering about the future (e.g., Smallwood et al., 2009; Baird et al., 2011). Developmental research has shown that the ability to think about the future develops between ages 4 and 5 (e.g., Russell et al., 2010; Suddendorf et al., 2011; Atance and Sommerville, 2014). The experimental approach to study future thinking in children has mainly relied on the use of the Spoon test (Tulving, 2005). This test is based on the following scenario: A young girl dreams that she is at a party where all the guests are being served chocolate pudding. To eat the pudding, the young girl needs a spoon, but she does not have one. That night, she falls asleep while holding a spoon. Bringing the spoon represents an instance of future thinking because it implies *envisioning* a need that will occur in the future.

Suddendorf et al. (2011) adapted Tulving's idea by presenting children with a problem (e.g., locked box with no key) in room A and a set of items (including a key) in room B. Their study showed that 4 but not 3 year olds choose the correct item to take back to room A (for similar results see: Atance and Meltzoff, 2005; Russell et al., 2010; Redshaw and Suddendorf, 2013; Scarf et al., 2013; Atance and Sommerville, 2014; Atance et al., 2015; Cuevas et al., 2015; Dickerson et al., 2018). Overall, these studies typically show an age-related improvement in future thinking between ages 3 to 5. The Spoon test has also been successfully implemented in studies with great apes. For example, Mulcahy and Call (2006) presented orangutans and bonobos with an out-of-reach reward and with a set of useful and useless tools, which they could take into a waiting room. To obtain the reward, subjects had to return to the room where the out-of-reach reward was placed, carrying the useful tool, either 1 or 24 h after having seen the reward. Mulcahy and Call showed that great apes did select and save the correct tool for a future use (see Osvath and Osvath, 2008 for similar results).

However, an important concern with the just described Spoon tests is that it is unclear whether thinking about a *future event* is needed when making successful choices. In these tasks, selecting the correct item may only indicate that subjects know that, for example, the key is useful for unlocking the box *now* without having to represent its use in a future event (e.g., Martin-Ordas, 2018; Hoerl and McCormack, 2019). To address this issue, it is required to demonstrate that individuals have some understanding of what the future might entail. Including a temporal component [i.e., before-and-after relationships; henceforth "temporal reasoning" (McCormack and Hoerl, 2011; Hoerl and McCormack, 2019)] will help assessing when in development the ability to envision the future emerges. Recently, Martin-Ordas (2018) addressed this issue by presenting 3-, 4- and 5-year-olds with a task in which, to secure a future need (e.g., play with a marble run game), children *first* had to obtain a key that allowed them *next* to access the marbles. By the age of four children selected the key; however, it is only by the age of 5 that children reasoned about the temporal sequence of future events *and* selected the key. Thus, this study highlighted the importance of assessing the temporal component of future thinking.

Interestingly, chimpanzees at the Goualougo Triangle (Republic of Congo) have been described to use two tools in sequence—a puncturing stick first and fishing probes next-when trying to access the termites from subterranean nests. Chimpanzees usually arrive at the nests with the two tools and, crucially, they have never been observed to only transport the puncturing stick—alone it would not be effective (Sanz et al., 2004). This study nicely illustrates how planning (e.g., Hayes-Roth and Hayes-Roth, 1979; Miller et al., 1960) might entail a temporal component since transporting *both tools*—as opposed to only bringing one tool regardless of its function—indicates envisioning the *two steps* of the termite-extracting problem (Byrne et al., 2013).

Inspired by Sanz et al.'s (2004) study, a termite-extracting problem was adapted to determine whether children 4 and 5 years of age and adults can plan for a future event that involves selecting two tools. In the current studies, participants were presented with a task that needed a "puncturing" tool to first make a hole on the top a cylinder and a "hook" tool to subsequently pull a reward through the hole (see Weir et al., 2002; Beck et al., 2011; for a similar task). In order to succeed, participants had to envision the two steps of the problem and select the correct two tools. There was a "Spatial-displacement group" (i.e., the task and four items—the two necessary tools and two other functionless items- were placed in two different rooms) and a "No-spatial-displacement group" (i.e., items and task were placed in the same room). For the Spatial-displacement group, successful performance required selecting the tools while holding a memory of the task and envisioning the correct sequence of tool-actions. By comparison for the No-spatial-displacement group, succeeding entailed selecting the tools while only envisioning the correct sequence of actions as the task was in plain view. The human ability to think about the future is unquestioned, thus older children and adult humans should be able to envision the steps of the task and plan accordingly. Note that participants were not trained in the task nor they were given demonstrations on how to solve a functionally equivalent task. In addition, single-trial methods were used in the present studies (e.g., Suddendorf and Corballis, 2007; Suddendorf et al., 2011). As a result, these experiments serve as a potentially interesting test of human planning under the criteria previously defined for animals. In this regard, these findings will contribute to the comparative research not only by offering insights on the range of responses that can be performed in planning tasks but also by identifying under which conditions humans produce those responses.

EXPERIMENT 1: CHILDREN (I)

Materials and Methods

Participants

A total of 60 children were recruited, with 1 participant being excluded due to experimental error, resulting in a final sample of 59 participants (28 females; 31 males) aged 4 (M = 53.13 months, SD = 2.97, n = 30) and 5 (M = 65.38, SD = 3.44, n = 29). All participants were predominantly White, middle class, and fluent in English. Children were tested individually at the Center for Life in Newcastle (United Kingdom). The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee (Project name: Future thinking in children and adults). Parents provided written informed consent for their children's participation, and children also provided their verbal assent.

Procedure

The experiment took place in two different areas: Room 1 and Room 2. First, participants were presented in Room 1 with a long narrow transparent container (16 cm length × 4.5 cm diameter) so that they could not use their hands to reach its bottom (e.g., 12). A 18 cm pipe cleaner with a hook made at one end ("hook tool"), a 10 cm long × 4 mm width paper blender stump ("puncturing tool"), a 10 cm long × 4 mm width strip of paper ("short tool") and a 22 cm long × 4 mm width strip of paper ("long tool") were used as tools. A bucket containing a reward (e.g., 3 stickers) was placed at the bottom of the container. The opening of the container was covered with extra-strong foil paper and children were explained that the foil was glued to the sides of the container (see **Figure 1**). Note that the puncturing tool could only function as a tool to puncture the foil paper and the hook tool as a tool to lift the bucket (the two ends of the hook were made soft so they could not pierce the foil paper).

FIGURE 1 | Task and tools used in the present experiment. From left to right: Puncturing tool, hook tool, short tool, and long tool.

The experimenter (E) said "*If you can get the stickers, you get to keep them.*" Each participant was then randomly assigned to one of two following groups:

(1) *Spatial-displacement group:* For this group, the task and the 4 tools were placed in different rooms. E and participant went to Room 2 (i.e., Tool room). From this area participants did not have visual access to Room 1 (i.e., Task room). E presented participants with the tools and said "*You can use some of these things to help you. Can you think what you will need to get the stickers?*" Children were told that they should get the stickers without turning the container. For each tool-choice opportunity, there were no explicit instructions about the number of tools children could choose—they could choose as many tools as they considered necessary. Likewise, E did not inform about the number of opportunities that participants had to choose tools. Once children made their choice, E and child went back to the Task room. Children were allowed to manipulate the tools so they could learn about the properties of the tools before making their choices. The procedure continued as follows depending on children's choices:

   1.1. Children chose the 4 tools. E allowed them to try to use the tool on the task. If participants tried to use first any of the incorrect tools, E said: "*Oh no, it does not work because we cannot get through the paper.*" Likewise, if children used the correct tool first and tried to use any of the incorrect tools next, E said: "*Oh no, it does not work because we cannot get the bucket.*" Pilot data suggested that children started to get frustrated if they tried to solve the problem more than four times. For that reason, children were given a maximum of 4 tool-use attempts to obtain the reward. At that point, if the child had not chosen and used the puncturing and hook tools in the correct order, E proceeded to get the bucket out of the container using the correct tools and gave the reward to the participant.

   1.2. Children chose one or more of the following tools: hook, long stick or short stick. E allowed them to try the tool(s) on the container. After each tool-attempt, E said: "*Oh no, it does not work because we cannot get through the paper. Let's go back to the other room and see if there is something else that could help you get the stickers.*" This procedure was repeated a maximum of 4 times. At that point, if the participant had not chosen and used the puncturing and hook tools in the correct order, E proceeded to get the bucket out of the container using the correct tools and gave the reward to the participant.

   1.3. Participants chose only the puncturing tool. As before, participants were allowed to use the tool on the container. Then, E said: "*What do we do next? Can you think what else you need to get the stickers?*" If children did not spontaneously suggest to go back to the Tool room, E said: "*Let's go back to the other room and see if there is something else that could help you get the stickers.*" If children chose the hook, s/he was allowed to use it to obtain the reward. If participant chose any of the other tools (e.g., short stick, long stick), E followed the procedure described in the previous sections.

   1.4. Participants chose both the puncturing tool and the hook. E allowed them to use both tools on the apparatus to obtain the reward.

(2) *No-spatial-displacement group*: For this group, the container and 4 tools were placed in plain view in the same room. The same procedure as for the Spatial-displacement group was used. Likewise, for participants' choices E followed the exact same procedure as above, except that she omitted "*Let's go back to the other room.*" The rationale for having this condition was to assess whether participants could solve the problem (1) when all the elements of problem were presented in the same room and (2) when the presentation of the task was immediately followed by the presentation of the tools.

Data Scoring and Statistical Analyses

Sessions were video-recorded. Participants received a score of 1 if they selected only the two correct tools before using the selected tools for the task (i.e., two-step planning). Any other response (e.g., selecting only 1 tool) received a score of 0. For those participants who only selected 1 tool, which tool was chosen on the first tool-choice opportunity was also scored. Participants were considered to have solved the task (i.e., success = 1) if they obtained the reward by themselves in a maximum of 4 tool use attempts and to fail the task if the E helped them to obtain the reward after 4 attempts (i.e., fail = 0). In addition, the total number of tool-use attempts required to obtain the reward was scored. For example, a child could choose 2 tools (puncturing tool and hook) but scored 3 tool-use attempts (e.g., participant first used hook, then, puncturing tool, then hook). Forty percent of the data was coded by a second rater. Cohen's k for planning and first chosen tool was perfect ($k = 1.000$), and excellent for solving the task ($k = 0.82$).

Pearson chi-square tests were used to analyze the effect of condition and age in planning, task success, and tool chosen first. Kruskal-Wallis tests were used to analyze the effect of age for the total number of tool-use attempts and Mann-Whitney tests were carried out to assess *post hoc* age effects. Cramer's V, r, η^2, and ϕ were used to report effect sizes for significant effects. Statistical tests were two-tailed, and results were considered significant if $p < 0.05$.

Results

Two-Step Planning

Overall, age and condition did not have an effect on children's responses ($\chi^2 = 2.66$, $df = 1$, $p = 0.266$; see **Figure 2**). That is, children did not choose the two correct tools in their first tool-choice opportunity either when the tools and task were in the same room (*No-spatial displacement* condition) or when the tools and task were in different rooms (*Spatial displacement* condition). See **Table 1** for the percentage of children selecting 1, 2, 3, or 4 tools in both the *No-spatial displacement* condition and the *Spatial Displacement* condition.

First Tool-Choice

Children's first tool-choice was dependent on age and condition ($\chi^2 = 7.30$, $df = 1$, $p = 0.014$, *Cramer's V* = 0.35). Further analyses revealed that for the *Spatial-displacement* condition, selecting the puncturing tool first was not determined by age ($\chi^2 = 1.00$, $df = 1$, $p = 0.316$). In contrast, age did have an effect for the *No-spatial-displacement* condition ($\chi^2 = 7.98$, $df = 1$, $p = 0.008$, *Cramer's V* = 0.52; **Figure 3**). In this case, more 5 years old selected the puncturing tool first compared to 4 year olds—suggesting that whereas older children might be thinking about the correct sequence in which the problem had to be solved, 4 years old might only be focusing on the last step of the sequence. In fact, from the 5 years old

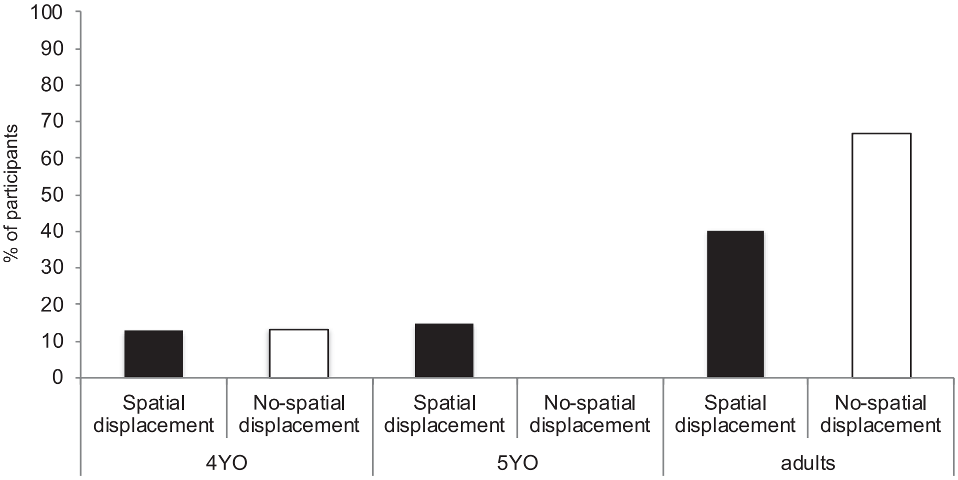


FIGURE 2 | Percentage of 4-year-olds (4YO), 5-year-olds (5YO) (Experiment 1) and adults (Experiment 2) selecting the two correct tools in their first tool-choice opportunity grouped by Spatial-displacement condition and No-spatial-displacement condition.

TABLE 1 | Percentage of children and adults selecting 1, 2, 3, or 4 tools in their first tool-choice opportunity in both the *No-spatial displacement* condition and the *Spatial Displacement* condition.

	Experiments 1 and 2							
	Spatial Displacement				**No-Spatial Displacement**			
	1 tool	**2 tools**	**3 tools**	**4 tools**	**1 tool**	**2 tools**	**3 tools**	**4 tools**
4YO	100%	–	–	–	88%	6%	6%	–
5YO	87%	13%	–	–	100%	–	–	–
Adults	60%	27%	–	13%	33%	67%	–	–
	Experiments 3 and 4							
	Spatial Displacement				**No-Spatial Displacement**			
	1 tool	**2 tools**	**3 tools**	**4 tools**	**1 tool**	**2 tools**	**3 tools**	**4 tools**
4YO	100%	–	–	–	88%	6%	–	–
5YO	88%	12%	–	–	100%	–	–	–
Adults	20%	47%	20%	13%	20%	74%	–	16%

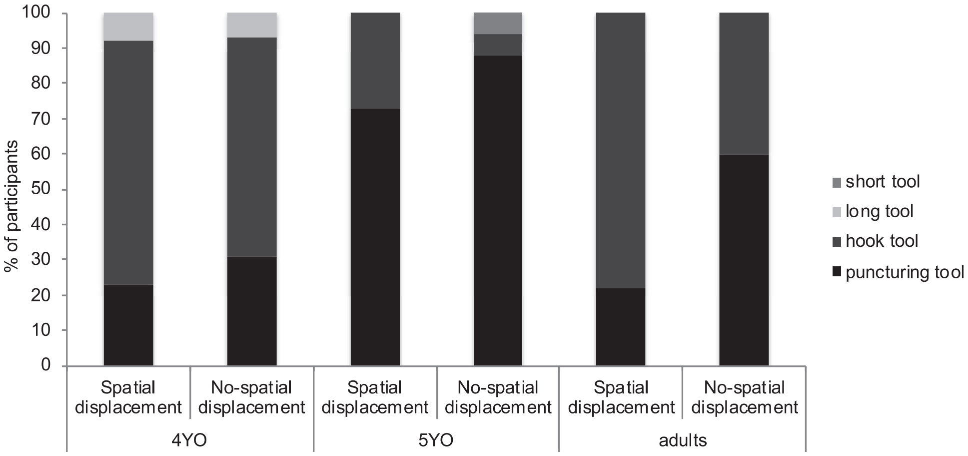


FIGURE 3 | Percentage of 4-year-olds (4YO), 5-year-olds (5YO) (Experiment 1) and adults (Experiment 2) selecting each of the possible tools (i.e., short tool, long tool, hook tool and puncturing tool) in those instances in which they only selected one tool in their first tool-choice opportunity. Data is grouped by Spatial-displacement condition and No-spatial-displacement condition.

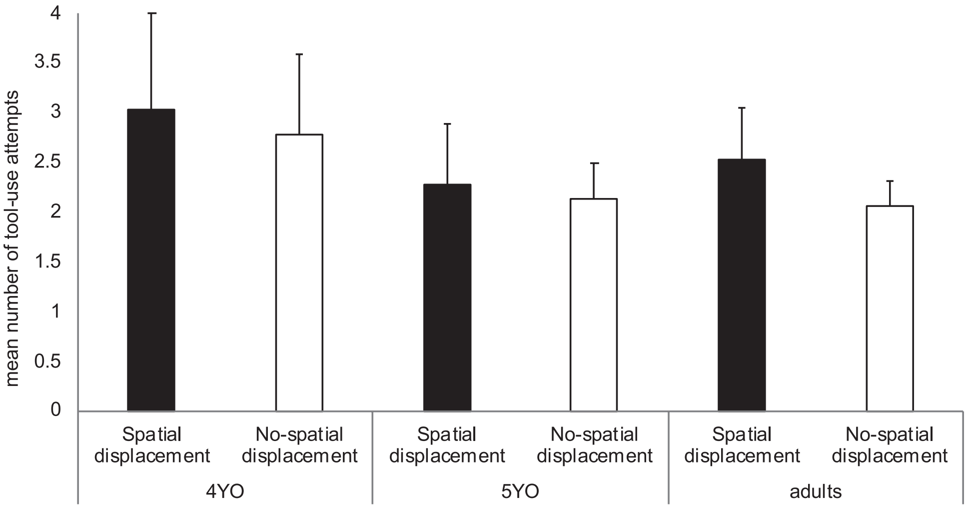


FIGURE 4 | Mean number of tool-use attempts that 4 year-olds (4YO), 5-year-olds (5YO) (Experiment 1) and adults (Experiment 2) needed to obtain the reward grouped by Spatial-displacement condition (black bars) and No-spatial-displacement condition (white bars). Error bars represent the SD.

who selected the puncturing tool first, 92% of them selected the hook second.

Task Success

Could the above findings be explained by a failure to solve the problem (i.e., obtain the reward)? Age and condition significantly affected participants' task success ($\chi^2 = 17.67$, $df = 1$, $p < 0.001$, *Cramer's V* = 0.55). Fewer 4 years old solved the task compared to 5 years old in both groups (*Spatial-displacement condition*: $\chi^2 = 9.94$, $df = 1$, $p = 0.002$, $\phi = 0.58$; *No-spatial-displacement condition*: $\chi^2 = 7.74$, $df = 1$, $p = 0.005$, $\phi = 0.51$). In fact, whereas 67% of 4 years old in the *Spatial-displacement condition* and 64% in the *No-spatial-displacement condition* obtained the reward, all 5 years old obtained the reward in both the *Spatial-displacement* and *No-spatial-displacement conditions*.

Age was also found to have a significant effect in the number of tool-use attempts required to obtain the reward. Particularly, 4 years old needed more tool-use attempts to solve the task than 5 years old (*Spatial-displacement condition*: Mann-Whitney: $U = 63.50$, $n = 29$, $p = 0.032$, $r = 0.54$; *No-spatial-displacement condition*: $U = 64.50$, $n = 27$, $p = 0.044$, $r = 0.57$; **Figure 4**). These results established that poor problem-solving abilities are at play in younger preschoolers' performance but not in 5 years old.

Discussion

Pre-schoolers in both *Spatial-displacement* and *No-spatial-displacement* conditions failed to select the two correct tools before performing any tool-use action. More 5-year-old children compared to 4-year-old children solved the task and they did so in fewer tool-attempts than younger children.

Poor problem-solving skills—i.e., failure at sequencing the order in which two tools had to be used- can account for 4-year-old children's performance. The current findings replicate previous research showing that children's temporal reasoning abilities are not fully developed before the age of 5 (McColgan and McCormack, 2008; McCormack and Hanley, 2011) and extend them to tool-use tasks involving reasoning about future goals.

There is no question that adults are better at planning than young children. Thus, if the present planning task is still challenging for children because it encompasses envisioning a sequence of actions, then adults would be expected to perform better than children. This possibility was investigated by presenting adults with the same task that children received in Experiment 1.

EXPERIMENT 2: ADULTS (I)

Materials and Methods

Participants

Thirty-one young adults were recruited and 1 was excluded due to malfunctioning of the apparatus, resulting in a final sample of 30 participants (27 females; 3 males) aged between 18 and 35 years. All participants were predominantly White, middle class, and fluent in English. Participants were tested individually in the lab facilities at the Institute of Neuroscience. The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee (Project name: Future thinking in children and adults). Adult participants provided written informed consent.

Procedure

The exact same procedure as in Experiment 1 was followed. Chocolates were used as rewards.

Data Scoring and Statistical Analyses

Data coding and statistical analyses were the same was in Experiment 1.

Results

Two-Step Planning

Condition had an effect on participants' responses ($\chi^2 = 4.82$, $df = 1$, $p = 0.028$, $\phi = 0.40$)—with more participants in the *No-spatial-displacement* condition (67%) choosing the two correct tools before using them than in the *Spatial-displacement* condition (27%) (**Figure 2**; see also **Table 1** for the percentage of participants selecting 1, 2, 3, or 4 tools in both the *No-spatial displacement* condition and the *Spatial Displacement* condition).

First Tool-Choice

Participants' first tool-choice was dependent on condition ($\chi^2 = 5.00$, $df = 1$, $p = 0.025$, $\phi = 0.40$; **Figure 3**). In this case, more adults selected the puncturing tool first in the *No-spatial-displacement* condition compared to *Spatial-displacement* condition—indicating that having the problem in participants' view might have facilitated thinking about the sequence in which the problem had to be solved. From those participants who chose the puncturing tool first, 100% chose the hook second in both conditions—indicating that participants might have envisioned the correct sequence of actions.

Task Success

Condition did not have a significant effect in the number of tool-use attempts required to obtain the reward (Mann-Whitney: $U = 112.50$, $n = 30$, $p = 1$; **Figure 4**). Note that all adults succeeded at obtaining the reward in both the Spatial-displacement and No-spatial-displacement groups.

Discussion

More adults in the *No-spatial-displacement* group compared to the *Spatial-displacement* group successfully chose the correct two tools before performing in the apparatus. All participants in both conditions successfully solved the task.

Certainly, the cognitive mechanisms involved in planning are fully matured in adults. However, adults' performance in the *Spatial-displacement* group did not select the two tools required to obtain the reward. Motivation or differences in procedure cannot account for these differences because reward and script were the same for both groups. One possibility is that participants in the *No-spatial-displacement* and *Spatial-displacement* groups used different strategies. Decision-making and problem-solving research has shown that adults select among different decision strategies by making a trade-off between the possibility of making correct decisions and the possibility of minimizing effort (Payne, 1976; Johnson and Meyer, 1984; Gigerenzer and Gaissmaier, 2011). For example, when facing a maze problem people usually choose what seems the most direct path to the goal at each step—even though this choice might be incorrect. Similarly, adults in the *Spatial-displacement* group might have traded off accuracy for cognitive effort by selecting the tool that seemingly could have two functions—piercing and extracting[1]. However, in the *No-spatial-displacement* condition such cognitive effort was lessened because task and tools were in plain sight—i.e., not having to recall the task might have facilitated a more effective planning strategy.

In order to investigate this possibility, we presented children (Experiment 3) and adults (Experiment 4) with the same task as before with the difference that now participants were limited to *one* opportunity to choose the tools that they needed to obtain the reward. While it is true that in Experiment 1 children did not show planning behaviors that clearly indicated that they envisioned the two-step sequence, limiting the number of tool-choice opportunities might still prompt them to choose the two correct tools—at least, in older children since they were able to solve the problem. The same could apply to adults. However, if participants are minimizing their cognitive effort, then it would be expected that by limiting to one the opportunity to choose

---

[1] All adults who chose the hook first attempted to use it to obtain the reward. Additionally, more adults in the *Spatial-displacement* group used the hook first compared to the *No-spatial-displacement* group ($\chi^2 = 7.77$, $df = 1$, $p = 0.005$).

tools, then, at least, adults would select the 4 available tools in the *Spatial-displacement* condition.

EXPERIMENT 3: CHILDREN (II)

Materials and Methods

Participants

A total of 66 children were recruited, with three participants being excluded due to experimental error, resulting in a final sample of 63 participants (25 females; 38 males) aged 4 (M = 55.13 months, SD = 3.02, n = 32) and 5 (M = 64.68, SD = 3.03, n = 31). All participants were predominantly White, middle class, and fluent in English. Children were tested individually at the Center for Life in Newcastle (United Kingdom). The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee (Project name: Future thinking in children and adults). Parents provided written informed consent for their children's participation, and children also provided their verbal assent.

Procedure

The same materials as in Experiment 1 were used for Experiment 3. The procedure was also the same as in Experiment 1 with exception that when presented with the tools, participants were explicitly told that they could only make a choice: "*Maybe you can use some of these things to help you. But you have to think carefully because you can only choose once, ok? Once you decide what you will need to get the stickers, I will put the things that you did not choose away.*" Once children selected the tool/s, E removed the remaining ones and let the children use the selected ones in the task. For those children who did not choose the correct tools, the E put the non-selected tools back on the table and asked them to obtain the reward by using any of the available tools. As in Experiment 1, children had 4 attempts to obtain the reward. This was done to assess children's problem-solving abilities. Also as in Experiment 1, 50% of the participants were presented with the task and tools in different rooms (*Spatial-displacement* group) and the other 50% were presented with the task and tools in the same room (*No spatial-displacement* group).

Data Scoring and Statistical Analyses

Sessions were video-recorded. Data were coded and analyzed in exact the same way as in Experiment 1. Forty percent of the data was coded by a second rater. Cohen's k for planning and first chosen tool was perfect (k = 1.000), and excellent for solving the task (k = 0.90). Cramer's V, r, η^2, and ϕ were used to report effect sizes for significant effects. Statistical tests were two-tailed, and results were considered significant if $p < 0.05$.

Results and Discussion

Two-Step Planning

Neither age nor condition had an effect on children's planning behavior ($\chi^2 = 0.31$, $df = 1$, $p = 0.573$). As in Experiment 1, children did not choose the two correct tools in advance when both the tools and task were in the same room (*No-spatial displacement condition:* 6% of 4YO and none of the 5YO) nor when the tools and task were in different rooms (*Spatial displacement condition:* none of 4YO and 12% of the 5YO; see **Table 1**). These results suggest that limiting the number of tool-choice opportunities did not improve children's abilities to select the two tools required to solve the problem.

First Chosen Tool

The total number of tools selected by the children in their only choice was not affected by age (Mann-Whitney: U = 492.5, p = 0.953, n = 63) or condition (Mann-Whitney U = 470.5, p = 0.669, n = 63). From all the children who only chose one tool, neither age nor condition were found to affect the selection of the correct first tool ($\chi^2 = 2.41$, $df = 1$, $p = 0.121$). These results suggest that forcing children to only make one choice did not improve their accuracy at selecting the tools they needed to solve the problem.

Task Success

Recall that for all children who did not select the two correct tools in their only tool-choice, E put the remaining objects back on the table. When task success was analyzed, neither age nor condition were found to affect children's performance ($\chi^2 = 0.005$, $df = 1$, $p = 0.941$). However, whereas 80% of 5 years old, obtained the reward in both the *Spatial-displacement* and the *No-spatial-displacement* groups, 69% of 4 year-olds did so in both conditions. The number of tool-use attempts was not determined by age (*No-spatial displacement*: Mann-Whitney U = 96, p = 0.438, n = 32; *Spatial displacement*: Mann-Whitney U = 103.5, p = 0.457, n = 31).

Overall, these results replicated the findings from Experiment 1. Younger children's performance can be explained by their difficulty to solve the problem. In contrast, older children were able to solve the problem but failed to anticipate that they needed two tools to obtain the reward. Next, adults' performance was examined.

EXPERIMENT 4: ADULTS (II)

Materials and Methods

Participants

Thirty-one young adults were recruited (18 females; 13 males) with one participant being excluded due to experimental error, resulting in a final sample of 30 participants aged between 18 and 35 years. All participants were predominantly White, middle class, and fluent in English. Participants were tested individually in the lab facilities at the Institute of Neuroscience. The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee (Project name: Future thinking in children and adults). Adult participants provided written informed consent.

Procedure

The same materials as in previous Experiments were used for Experiment 4. The procedure was also the same as in Experiment 3.

Data Scoring and Statistical Analyses

Sessions were video-recorded. Data were coded and analyzed in exact the same way as in Experiment 3.

Results and Discussion

Two-Step Planning

In this case, participants' planning and ability to envision the two-steps sequence were determined by condition ($\chi^2 = 4.88$, $df = 1$, $p = 0.050$). In the *Spatial displacement* condition, 46% of the participants selected the two correct tools in their only tool-choice opportunity and 85% did so in the *No-spatial displacement* condition. Thus, compared to Experiment 2, the number of participants selecting the 2 correct tools increased in this Experiment.

First Chosen Tool

The total number of tools selected by the participants in their only choice opportunity was not affected by condition (Mann-Whitney $U = 28.50$, $p = 1$, $n = 30$). The idea behind this manipulation was to investigate whether adults were minimizing the cognitive effort by selecting the tool that looked like it could have two functions (e.g., puncturing and lifting). If this were the case, then more participants should have selected the four tools in the *Spatial-displacement* condition compared to the *No-spatial displacement* condition. Note that 20% of the participants chose one tool in both *No-spatial displacement* condition and *Spatial-displacement* condition. In all these cases, participants chose the poking tool first. Moreover, in the *Spatial-displacement* condition 20% of the participants selected three tools and 13% selected the four tools. In the *No-spatial displacement* condition, 16% of the participants selected the four tools and none selected three tools (see **Table 1**).

Thus, limiting participants to only one tool-choice opportunity increased their tool selectivity although not enough to help them select the two correct tools in the *Spatial-displacement* condition.

Task Success

All adults obtained the reward in both the Spatial-displacement and the No-spatial-displacement groups. And the number of tool attempts did not differ between conditions (Mann-Whitney $U = 97$, $p = 0.508$, $n = 30$). As in Experiment 2, these findings demonstrate that performance was not determined by participants' problem-solving skills.

GENERAL DISCUSSION

The current studies showed that older children and adults were able to *use* two tools in sequence to obtain a reward. Fewer 4-year-old children—compared to older children- did so. Crucially, participants—both adults and children- struggled to anticipate the number of tools required to solve the problem in their first tool-choice opportunity. Although limiting to one the number of tool-choice opportunities improved adults' performance, children's responses were not affected by this manipulation. Adults in the *No-spatial-displacement* group successfully selected the two correct tools for the two-step sequence required to obtain the reward, but those in the *Spatial-displacement* group failed to anticipate the two correct tools.

By the age of 2 children have been shown to select an adequate tool based on properties such as length or rigidity (Bates, 1979; Willatts, 1985, 1999; Brown, 1990; Chen and Siegler, 2000; Gredlein and Bjorklund, 2005; see Martin-Ordas et al., 2014 for a study showing that by age 3 children can select a correct tool based on its diameter). In a similar task to the one presented here, results showed that it is only between ages 5 and 8 that children can make a tool suitable to get the bucket out of the tube (Beck et al., 2011; Cutting et al., 2011). Crucially, if children were given a choice between a straight pipe cleaner and a premade hook, by the age of 4 children could select the hook to get the bucket out of the tube (Beck et al., 2011). Thus, children seemed to find difficult the "innovation" aspect of the task (i.e., making the tool), but they understood what properties the tool should have in order for them to obtain the reward (Beck et al., 2011). Importantly, the studies described so far involve using *one* tool to solve a problem. This is in contrast to studies presented here—in which children had to use two tools in a correct sequence of actions to solve the problem. Thus, it is possible that younger children found the current task more difficult than older children did because they lack the ability to sequence the two tool-use actions. This is similar to previous studies showing that it is only by the age of 5 that children can incorporate temporal reasoning to their decision making (e.g., McCormack and Hoerl, 2005; McColgan and McCormack, 2008).

The ability to plan for a future event has been reported to develop between ages 4 and 5 (e.g., Suddendorf et al., 2011; Redshaw and Suddendorf, 2013; Atance and Sommerville, 2014; Atance et al., 2019). In the current experiment, both 4- and 5-year-old children struggled to anticipate the number of tools required to solve the current task. However, their tool selection indicated that children might have been planning for the future event since both age groups tended to select one of the two correct tools (see **Figure 3**). As mentioned above, lacking the temporal reasoning abilities could account for younger children's performance to select both tools. However, 5-year-olds did use two tools in sequence to obtain the reward. These findings are in contrast to a previous study showing that by the age of 5 children succeeded in a planning task that required envisioning the order of two future events (Martin-Ordas, 2018). Why did 5-year-old children fail to anticipate the number of tools required to solve the current problem?

There are two crucial differences between Martin-Ordas' (2018) study and the present ones. First of all, it is possible that whereas in the former the elements of the problem might have been semantically associated (e.g., keys open locks), in the task presented here such semantic association did not exist (e.g., pipe cleaners shaped as hooks might not necessary always be used to lift buckets). Second, whereas in Martin-Ordas (2018) children had to select *one* tool and decide the order in which two future events should happen (e.g., select the key, then visit the marble room to get the marbles and, next, go to the marble room), in the current task children had to envisage the two future actions in order to select the two correct items—with each action being associated to a particular tool. This aspect might have posed more cognitive demands to solve the problem—which, as a consequence, might have increased the difficulty of the task [see Burns and Russell, 2016 for a study showing only children over 5 years of age were able to anticipate a future event when the cognitive demands of the task were high (e.g., spatio-temporal predictions based on someone else's point of view)].

Adults can plan, envisage the future and think about temporal sequences, so why did they struggle in the current task? The studies presented here indicated that when limiting to one the number of tool-choices opportunities, more participants selected the correct two tools *before* performing any action on the task—although only in the *No-spatial-displacement* condition. These results suggest that adults might be selective planners—that is, even though they can plan, adults might only make use of this ability under particular circumstances. Recent developmental studies have highlighted that performance in the Spoon test is drastically affected when children are asked to *spontaneously* generate the solution to a problem rather than selecting a tool from a number of options (e.g., Moffett et al., 2018; Atance et al., 2019). These studies indicate that it is only by the age of 5 that children start to generate the solutions for a future problem. Along the same lines, the results presented here suggest that limiting participants' choices to one—and consequently, increasing the costs of making errors- affected their tool selectivity, at least, when the problem was in plain sight. This is similar to what previous tool-use studies with humans (Silva and Silva, 2010, 2012) and great apes (Mulcahy et al., 2005; Martin-Ordas et al., 2012) have already shown.

The constellation of results presented here suggest that, at least, adult humans' planning responses varied depending on whether the problem was in plain sight compared to when the problem was out of sight. It would be difficult to argue that adults in these experiments did not understand the critical features of the tasks that they had to solve—otherwise the differences in performance between the *Spatial-displacement* and *No-spatial displacement* conditions would not have been found. However, this explanation remains as a possibility for children's performance. Still, children's responses in the present tasks do not necessarily indicate an inability to plan. As mentioned earlier, children are not randomly choosing one of the four tools; and their first tool choices seem to indicate that there is a representation of the future event—although, they seem to have difficulties to envision the two steps of the problem.

These limitations should not undermine the value of the present studies. The results presented here still have crucial implications for the field of animal future thinking and planning. In the current studies, participants were presented with an unfamiliar tool-use task. It could have been possible that both children and adults performed better if presented with a more familiar problem (i.e., a task in which a strong semantic association between tools and task existed). However, not all planning situations require dealing with familiar contexts or objects. As such, it is also insightful to study this ability and its flexibility in less accustomed contexts. Additionally, Suddendorf et al. (2011) suggested that tasks aiming to test future thinking should involve using novel problems in order to avoid (associative) learning. This is an important factor to understand future thinking in animals, since, in most cases, subjects are presented with unfamiliar situations that require the use of training and multi-trial methods. In the last 20 years, comparative psychologists have provided empirical evidence that other animals possess some type of future thinking abilities (e.g., Mulcahy and Call, 2006; Osvath and Osvath, 2008; Kabadayi and Osvath, 2017). These findings have been the focus of arduous debates—with some claiming that future thinking abilities in some animals are similar to those in humans (e.g., Martin-Ordas et al., 2014) and others arguing that even the strongest pieces evidence of future thinking in animals can be acknowledged to be no more than (associative) learning achievements (e.g., Suddendorf and Corballis, 2007). Accordingly, if providing subjects with more than one trial (i.e., repeated exposure to the same stimulus-reward relationship) would entail that associative learning—rather than future thinking- could account for their performance, then a first response preferably without training should be considered the standard to show future thinking. Nonetheless, the artificiality of this situation might undermine performance, as one could argue for the present studies—recall that a single-trial method was used to test participants in the present studies. Thus, these studies with children and adults highlight conceptual and methodological issues in the criteria described to asses future thinking (e.g., Anderson, 2001; Silva et al., 2005; for similar arguments on tool-use tasks). The studies also provide the set of responses that humans display under some of the conditions required to test future thinking in animals.

To conclude, more older children and adults compared to younger children succeeded at using two tools in sequence to obtain a reward. Whereas children did not select the 2 tools required to solve the problem in their first tool-choice opportunity, adults were able to do so when the task was in plain view. Human performance in the present tasks highlights important points for comparative research. First, the issue of how to measure future thinking seems to not be completely solved if we are to focus on the novelty of the problems and the lack of training in order to rule out associative learning as the mechanism driving performance in these tasks. Thus, criteria that can equally be applied to humans and animals and that allow us to draw irrefutable comparisons across species are needed. Second, including groups of children and adults in comparative studies will offer reliability to the results and will be informative comparison groups for behavioral tests of these capacities in animals (e.g., Anderson, 2001; Silva et al., 2005; Silva and Silva, 2006). Examining what humans can do will provide us with critical information to be able to identify shortcomings in the study of the comparative research of future thinking and also to provide a context in which to interpret animals' responses.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The experiment received ethical approval from the Newcastle University's Faculty of Medical Sciences Ethics Committee (Project name: Future thinking in children and adults). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

GM-O designed the study, collected the data, and wrote the manuscript.

FUNDING

This work was supported by a Marie Skłodowska-Curie individual fellowship project number: 657505, Horizon 2020.

ACKNOWLEDGMENTS

I thank Quoc Vuong for insightful comments on earlier versions of this manuscript. And also thank Omar Jouel Miah for his help with data collection and coding. Special thanks to all the children, parents and adults who volunteered their participation.

REFERENCES

Anderson, J. R. (2001). Review of the book Folk physics for apes: the chimpanzee's theory of how the world works. *Anim. Behav.* 61, 1042–1043.

Atance, C. M., Louw, A., and Clayton, N. S. (2015). Thinking ahead about where something is needed: new insights about episodic foresight in preschoolers. *J. Exp. Child Psychol.* 129, 98–109. doi: 10.1016/j.jecp.2014.09.001

Atance, C. M., and Meltzoff, A. N. (2005). My future self: young children's ability to anticipate and explain future states. *Cogn. Dev.* 20, 341–361. doi: 10.1016/j.cogdev.2005.05.001

Atance, C. M., Nur Celebi, S., Mitchinson, S., and Mahy, C. E. V. (2019). Thinking about the future: comparing children's forced-choice versus "generative" responses in the "spoon test.". *J. Exp. Child Psychol.* 181, 1–16. doi: 10.1016/j.jecp.2018.12.006

Atance, C. M., and Sommerville, J. A. (2014). Assessing the role of memory in preschoolers' performance on episodic foresight tasks. *Memory* 22, 118–128. doi: 10.1080/09658211.2013.820324

Baird, B., Smallwood, J., and Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Conscious. Cogn.* 20, 1604–1611. doi: 10.1016/j.concog.2011.08.007

Bates, E. (1979). "The biology of symbols: some concluding thoughts," in *The Emergence of Symbols: Cognition and Communication in Infancy*, eds E. Bates, L. Benigni, I. Bretherton, L. Camaioni, and V. Volterra (New York, NY: Academic Press), 315–370. doi: 10.1016/b978-0-12-081540-1.50013-2

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., and Cutting, N. (2011). Making tools isn't child's play. *Cognition* 119, 301–306. doi: 10.1016/j.cognition.2011.01.003

Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cogn. Sci.* 14, 107–133. doi: 10.1207/s15516709cog1401_6

Burns, P., and Russell, J. (2016). Children's predictions of future perceptual experiences: temporal reasoning and phenomenology. *Dev. Psychol.* 52, 1820–1831. doi: 10.1037/dev0000197

Byrne, R. W., Sanz, C. M., and Morgan, D. B. (2013). "Chimpanzees plan their tool use," in *Tool Use in Animals. Cognition and Ecology*, eds C. M. Sanz, J. Call, and C. Boesch (Cambridge: Cambridge University Press), 48–64. doi: 10.1017/cbo9780511894800.004

Chen, Z., and Siegler, R. S. (2000). Across the great divide: Bridging the gap between understanding of toddlers' and older children's thinking. *Monogr. Soc. Res. Child Dev.* 65, 1–96.

Cuevas, K., Rajan, V., Morasch, K. C., and Bell, M. A. (2015). Episodic memory and future thinking during early childhood: Linking the past and future. *Dev. Psychobiol.* 57, 552–565. doi: 10.1002/dev.21307

Cutting, N., Apperly, I. A., and Beck, S. R. (2011). Why do children lack the flexibility to innovate tools? *J. Exp. Child Psychol.* 109, 497–511. doi: 10.1016/j.jecp.2011.02.012

Dickerson, K. L., Ainge, J. A., and Seed, A. M. (2018). The role of association in pre-schoolers' solutions to "spoon tests" of future planning. *Curr. Biol.* 28, 2309–2313.

Gigerenzer, G., and Gaissmaier, W. (2011). Heuristic decision making. *Annu. Rev. Psychol.* 62, 451–482. doi: 10.1146/annurev-psych-120709-145346

Gredlein, J. M., and Bjorklund, D. F. (2005). Sex differences in young children's use of tools in a problem-solving task. *Hum. Nat.* 16, 211–232. doi: 10.1007/s12110-005-1004-5

Hayes-Roth, B., and Hayes-Roth, F. (1979). A cognitive model of planning. *Cogn. Sci.* 3, 275–310. doi: 10.1016/s0364-0213(79)80010-5

Hoerl, C., and McCormack, T. (2019). Thinking in and about time: a dual systems perspective on temporal cognition. *Behav. Brain Sci.* 42:e244.

Johnson, E. J., and Meyer, R. J. (1984). Compensatory choice models of noncompensatory processes: the effect of varying context. *J. Consum. Res.* 11, 528–541.

Kabadayi, C., and Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science* 357, 202–204. doi: 10.1126/science.aam8138

Martin-Ordas, G. (2018). "First, I will get the marbles". Children's foresight abilities in a modified spoon task. *Cogn. Dev.* 45, 152–161. doi: 10.1016/j.jpeds.2016.10.001

Martin-Ordas, G., Atance, C. M., and Call, J. (2014). Remembering in tool-use tasks in chidren and apes: the role of information at encoding. *Memory* 22, 129–144. doi: 10.1080/09658211.2013.806553

Martin-Ordas, G., Schumacher, L., and Call, J. (2012). Sequential tool use in great apes. *PLoS One* 7:e52074. doi: 10.1371/journal.pone.0052074

McColgan, K. L., and McCormack, T. (2008). Searching and planning: young children's reasoning about past and future event sequences. *Child Dev.* 79, 1477–1497. doi: 10.1111/j.1467-8624.2008.01200.x

McCormack, T., and Hanley, M. (2011). Children's reasoning about the temporal order of past and future events. *Cogn. Dev.* 26, 299–314. doi: 10.1016/j.cogdev.2011.10.001

McCormack, T., and Hoerl, C. (2005). Children's reasoning about the causal significance of the temporal order of events. *Dev. Psychol.* 41, 54–63. doi: 10.1037/0012-1649.41.1.54

McCormack, T., and Hoerl, C. (2011). "Tool use, planning, and future thinking in animals and children," in *Tool Use and Causal Cognition*, eds T. McCormack, C. Hoerl, and S. Butterfill (Oxford: Oxford University Press), 129–147. doi: 10.1093/acprof:oso/9780199571154.003.0007

Miller, G. A., Galanter, E., and Pribram, K. H. (1960). *Plans and the Structure of Behavior*. New York, NY: Holt, Rinehart & Winston.

Moffett, L., Moll, H., and Fitz Gibbon, L. (2018). Future planning in preschool children. *Dev. Psychol.* 54, 866–874. doi: 10.1037/dev0000484

Mulcahy, N. J., and Call, J. (2006). Apes save tools for future use. *Science* 312, 1038–1040. doi: 10.1126/science.1125456

Mulcahy, N. J., Call, J., and Dunbar, R. I. M. (2005). Gorillas and orangutans encode relevant problem features in a tool-using task. *J Comp. Psychol.* 119, 23–32. doi: 10.1037/0735-7036.119.1.23

Osvath, M., and Osvath, H. (2008). Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: Self-control and pre-experience in the face of future tool use. *Anim. Cogn.* 11, 661–674. doi: 10.1007/s10071-008-0157-0

Payne, J. (1976). Task complexity and contingent processing in decision making: an information search and protocol analysis. *Organ. Behav. Hum. Perform.* 16, 366–387. doi: 10.1016/0030-5073(76)90022-2

Redshaw, J., and Suddendorf, T. (2013). Foresight beyond the very next event: Four-year-olds can link past and deferred future episodes. *Front. Psychol.* 4:404. doi: 10.3389/fpsyg.2013.00404

Russell, J., Alexis, D., and Clayton, N. S. (2010). Episodic future thinking in 3- to 5-year-old children: the ability to think of what will be needed from a different point of view. *Cognition* 114, 56–71. doi: 10.1016/j.cognition.2009.08.013

Sanz, C., Morgan, D., and Gulick, S. (2004). New insights into chimpanzees, tools and termites from the Congo Basin. *Am. Nat.* 164, 567–581. doi: 10.1086/424803

Scarf, D., Gross, J., Colombo, M., and Hayne, H. (2013). To have and to hold: Episodic memory in 3- and 4-year-old children. *Dev. Psychobiol.* 55, 125–132. doi: 10.1002/dev.21004

Silva, F. J., Page, D. M., and Silva, K. M. (2005). Methodological-conceptual problems in the study of chimpanzees' folk physics: how studies with adult humans can help. *Learn. Behav.* 33, 47–58. doi: 10.3758/bf03196049

Silva, F. J., and Silva, K. M. (2006). Humans' folk physics is not enough to explain variations in their tool-using behavior. *Psychon. Bull. Rev.* 13, 689–693. doi: 10.3758/bf03193982

Silva, F. J., and Silva, K. M. (2010). How do adult humans compare with New Caledonian crows in tool selectivity? *Learn. Behav.* 38, 87–95. doi: 10.3758/LB.38.1.87

Silva, F. J., and Silva, K. M. (2012). More but not less uncertainty makes adult humans' tool selections more similar to those reported with crows. *Learn. Behav.* 40, 494–506. doi: 10.3758/s13420-012-0069-3

Smallwood, J., Nind, L., and O'Connor, R. C. (2009). When's your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Conscious. Cogn.* 18, 118–125. doi: 10.1016/j.concog.2008.11.004

Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: what is mental time travel and is it unique to humans? *Behav. Brain Sci.* 30, 299–313. doi: 10.1017/s0140525x07001975

Suddendorf, T., Nielsen, M., and von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Dev. Sci.* 14, 26–33. doi: 10.1111/j.1467-7687.2010.00950.x

Tulving, E. (2005). "Episodic memory and autonoesis: uniquely human?," in *The Missing Link in Cognition*, eds H. S. Terrace and J. Metcalfe (Oxford: Oxford University Press), 3–56. doi: 10.1093/acprof:oso/9780195161564.003.0001

Weir, A. A. S., Chappell, J., and Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science* 297:981. doi: 10.1126/science.1073433

Willatts, P. (1985). Adjustment of means-ends coordination and the representation of spatial relations in the production of search errors by infants. *Br. J. Dev. Psychol.* 3, 259–272. doi: 10.1111/j.2044-835X.1985.tb00978.x

Willatts, P. (1999). Development of means-end behavior in young infants: pulling a support to retrieve a distant object. *Dev. Psychol.* 35, 651–667. doi: 10.1037/0012-1649.35.3.651

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 21 April 2020
doi: 10.3389/fpsyg.2020.00653

Crossing the Rubicon: Behaviorism, Language, and Evolutionary Continuity

Michael C. Corballis*

Faculty of Science, School of Psychology, The University of Auckland, Auckland, New Zealand

Euan Macphail's work and ideas captured a pivotal time in the late 20th century when behavioral laws were considered to apply equally across vertebrates, implying equal intelligence, but it was also a time when behaviorism was challenged by the view that language was unique to humans, and bestowed a superior mental status. Subsequent work suggests greater continuity between humans and their forebears, challenging the Chomskyan assumption that language evolved in a single step ("the great leap forward") in humans. Language is now understood to be based on an amalgam of cognitive functions, including mental time travel, theory of mind, and what may be more broadly defined as imagination. These functions probably evolved gradually in hominin evolution and are present in varying degrees in non-human species. The blending of language into cognition provides for both interspecies differences in mental function, and continuity between humans and other species. What does seem to be special to humans is the ability to communicate the contents of imagination, although even this is not absolute, and is perhaps less adaptive than we like to think.

Keywords: behaviorism, cognition, evolution, imagination, language

Edited by:
Thomas Zentall,
University of Kentucky, United States

Reviewed by:
Stephen V. Shepherd,
Independent Researcher, Richmond, VA, United States
Zhanna Reznikova,
Institute of Systematics and Ecology of Animals (RAS), Russia

Correspondence:
Michael C. Corballis
m.corballis@auckland.ac.nz

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *09 January 2020*
Accepted: *18 March 2020*
Published: *21 April 2020*

Citation:
Corballis MC (2020) Crossing the Rubicon: Behaviorism, Language, and Evolutionary Continuity. Front. Psychol. 11:653. doi: 10.3389/fpsyg.2020.00653

INTRODUCTION

> Language is our Rubicon, and no brute will dare cross it.
> — Müller (1861, p. 403)

In the days when behaviorism was the dominant paradigm in psychology, it was widely assumed not to matter which species you studied. In the late 1950s, when I first began to study psychology, rats were the species of choice, and a few years later they were more or less seamlessly replaced by pigeons. There was perhaps a limit as to how far one could go along the branches of the evolutionary tree, and in the 1950s the comparative psychologist James V. McConnell made a semi-serious attempt to introduce worms. *The Worm Runners' Digest* was established in 1959 and rather surprisingly lasted 20 years, mainly as a stimulus for countless school projects, and research reports often designed for humor rather than serious science.

With one exception, vertebrates were assumed to conform to the same universal behavioral laws. The British psychologist Euan Macphail went so far as to suggest that all were of equal intelligence, and for behaviorists this left little room for comparative psychology, or for ethology, which was left to the zoologists. Know one species, and you know them all, at least as far as intellect was concerned.

That one exception, though, was *Homo sapiens*, blessed with the faculty of language, raising human intelligence to a different level. Macphail held this position for over 20 years:

> It may be ironic that we suppose (e.g., Macphail, 1982) that the key to our intelligence lies in the appearance in humans alone of the capacity for language, and that the exploitation of linguistic skills (and other related symbolic skills, such as those of mathematics) results in our unique intelligence (Macphail and Bolhuis, 2001, p. 361).

This continued a tradition from at least Biblical times, and extending through most of western philosophy, of placing humans on a pedestal, somewhere between apes and angels. Hobbes (1651), in his essay *Of Man*, wrote that "the first author of speech was God himself," and the 17th-century philosopher Descartes (1984) supposed that language was the property that elevated humans above other animals, and bestowed the capacity for free will. Müller (1861) was Professor of Philology at Oxford University, and the quote that opened this article was his response to Darwin's (1859) book *On the Origin of Species*.

Skinner (1957) had been aware of the challenge posed by language, and made a monumental effort to account for language in behavioral terms in his book *Verbal Behavior*. Unfortunately, it was published in the same year as *Syntactic Structures* by a young Noam Chomsky (1957), who was part of a new generation launching what came to be called the "cognitive revolution." This was followed by Chomsky's (1959) influential but damning review of Skinner's book, which to many signaled the end of behaviorism as the dominant paradigm of psychology. Rats and pigeons were replaced by undergraduate students, eager for course credit, being tested for cognitive skills, and tapping at keys with the same dedication as their pigeon forebears.

The cognitive revolution also shifted attention away from learning toward structures generally assumed to be innate. In spite of the fact that some 7,000 different languages exist in the world, each incomprehensible to nearly all of the others, language was taken to depend on an innate endowment, and differing from animal communication in that it was based on computational rules applying universally but exclusively to humans. These rules were once called "deep structure" but later "universal grammar" (Chomsky, 1995). Universal grammar is considered a recursive system permitting an infinite variety of utterances, or "the infinite use of finite means" (von Humboldt, 1836/1999). Chomsky proposed, moreover, that universal grammar emerged fully-fledged in our own species in a single step, the "great leap forward," and perhaps even in a single individual, whom Chomsky (2010) whimsically named Prometheus. This runs counter to evolutionary theory; as Darwin (1859) himself put it, "natural selection can act only by taking advantage of slight successive variations; she can never take a leap, but must advance by the shortest and slowest steps" (p. 194).

At the end of the second decade of the 21st century, we are perhaps moving away from the Chomskyan era and the rigid insistence on innate structures underlying cognition. Given the often vast differences between languages, the very concept of universal grammar has been disputed. Evans and Levinson (2009), for example, concluded that "the emperor of Universal Grammar has no clothes" (p. 438) and Tomasello (2009) remarked similarly that "Universal grammar is dead" (p. 470). Some languages, such as those of the Pirahä of Brazil (Everett, 2005) or the Iatmul of New Guinea (Evans, 2009), do not seem to display the recursive structure of universal grammar. Viewed as a communication system, at least, language seems not to have a universal basis.

LANGUAGE AS THOUGHT

These objections have been partly finessed by Chomsky himself, who has insisted that universal grammar is part of what he calls *I-language*, which is not fundamentally concerned with communicative language at all. Rather, it is a component of thought. Actual languages, whether spoken, signed, or written, are part of a process of *externalization*, the transforming of internal thoughts into communicable form. This seems partly to resolve the problem of why there are so many languages in the world, and why they are so diverse. As Chomsky (2015b) put it: "It is a familiar fact (*sic*) that that the complexity and variety of language appears to be localized overwhelmingly—and perhaps completely—in externalization (p. xi)." This renders actual languages peripheral to the understanding of universal grammar itself.

The idea that the essence of language lies in thought rather than in communication should not be taken to mean that we think in words, as many have long claimed. Plato, for example, through the mouth of Sophocles, wrote that "the soul when thinking appears to me to be just talking—asking questions of herself and answering them" (Jowett, 1892, p. 190), and in 1798 Immanuel Kant wrote that "Thinking is speaking to ourselves" (quoted in translation by Butts, 1988, p. 278). The founder of behaviorism, John B. Watson, similarly equated thought with subvocal speech (Watson, 1928). For Chomsky, though, thought is primary, and words merely provide the means for articulating those thoughts, in agreement with the 17th century philosopher John Locke (1690/2017):

> We should have a great many fewer disputes in the world if words were taken for what they are, the signs of our ideas only, and not for the things themselves (p. 185).

Although the structure of I-language is not easily discerned in the thickets of languages themselves, Chomsky's views have become progressively simplified, culminating in *The Minimalist Program* (Chomsky, 1995, 2015b), which seemingly reduces the gap between humans and other species. Universal grammar is reduced to a single operation called *Unbounded Merge* (or simply *Merge*) in which elements are combined in recursive fashion to form new entities, which can themselves be merged, and so on. It is the basic mechanism for the construction of hierarchies, with language as the most overt example, but there are other examples, as we shall see later. In Chomsky's view, Merge applies not directly to communicative languages themselves, but rather to the abstract elements of I-language. Nevertheless, the operations of Merge might be manifest indirectly in spoken language, say, as successive merges of phonemes to form syllables, syllables to form words, words to form phrases, phrases to form sentences, and so on. But Merge itself seems a simple operation, leading Berwick and Chomsky (2016) to remark that, "we simply don't have as much to explain, reducing the Darwinian paradox" (p. 11). These authors nevertheless continue to insist that universal grammar is unique to humans, as reflected in the very title of their book, *Why Only Us?*

These developments blur the distinctions proposed by Macphail. Language may not after all be independent of animal intelligence, or of thought itself. Chomsky's Merge is presented as a highly specialized process, unique to humans, but generative processing may actually be a general aspect of thought, and evident in non-human animals, as I suggest later. One example is what has been termed *mental time travel*, the capacity to "replay" past events or imagine future ones (Suddendorf and Corballis, 1997, 2007)—also dubbed *episodic memory* and *episodic projection*, respectively. Indeed mental time travel may capture at least some of the properties of Merge itself, combining various elements into episodes, and accounting for much of the generativity of language itself. It underlies the linguistic property of *displacement*, defined by Hockett (1960) as the ability "to talk about things that are remote in space or time (or both) from where the talking is going on." Expressive language may have evolved precisely to enable us to communicate about the non-present (Corballis, 2009, 2017; Gärdenfors and Osvath, 2010; Bickerton, 2014).

One brain structure critically involved in mental time travel is the hippocampus. In one of our own brain-imaging studies (Martin et al., 2011), participants were asked to describe 110 episodes in their lives, and then construct 110 possible future episodes based on scrambled aspects of the remembered ones—tasks seemingly easily accomplished. While recalling or imagining these events, the hippocampus was active—more anteriorly for future events than past ones. The hippocampus also features in animal studies of mental time travel, discussed below.

The idea of mental time travel can be broadened to encompass imagination, which can include purely imaginary events and stories that are not necessarily grounded in reality. Indeed episodic projection necessarily includes the imagining of events that have not already occurred, and even episodic memories are often distortions of what actually occurred in the past, or are fabrications. We delight in making things up. Language, then, may be the externalization of imagination, involving richer and more experience-based constructions than suggested by the concept of Merge (Corballis, 2017). Along the same lines, Dor (2015) describes language as "the instruction of imagination."

Zuberbühler (2019) suggests that grammar itself derives from the perceived structure of events and evolved before communicative language itself. This compositional structure is made up of components such as actors, agents, patients, predication, and so on, and marking them with communicative signals. These components are merged in multiple ways to make up the real or imagined events, or episodes, of our lives. The linguist Paul Deane (1992) earlier argued that grammar is based on our understanding of space, and what happens in it. The way in which experience is structured in space and time may well be universal, and provide a basis for Chomsky's universal grammar. We all live in a spatiotemporal world, inhabited by things, people and artifacts of our own making. As humans, we are similarly size-scaled, differing from ants or elephants. We view the world at roughly the same angle and move about at roughly the same speed—at least until fast cars and airplanes disturbed our leisured pace. The mind no doubt adapted to these kinds of parameters, creating a near-universality of understanding. As I summarized earlier: "What Chomsky called universal grammar may therefore depend more on how long we and our forebears have inhabited the world and reacted to it than on some new internal program called unbounded Merge" (Corballis, 2017, p. 190).

But it is more than physical. We have also evolved as social animals, understanding how people function. This includes is *theory of mind*, which underlies the capacity to attribute mental states to other. This is recursive: I may know that you know something (level 1 theory of mind), but also know that you know that I know this (level 2). We can proceed, perhaps with effort, to higher levels of recursion, often in the interests of Machiavellian intrigue. Dunbar (2004) suggests level 5 recursion underlies the mutual understanding of gods: "I suppose that you think that I believe there are gods who intend to influence our future because they understand our desires (p. 185)." Grice (1989) proposed that at least level 2 theory of mind is necessary for communicative language:

> He said that P: he could not have done this unless he thought that Q; he knows (and knows that I know that he knows) that I will realize that it is necessary to suppose that Q; he has done nothing to stop me thinking that Q; so he intends me to think, or is at least willing for me to think that Q (pp. 30–31).

Or to put it in more everyday terms, communicative language requires that the speaker knows what's in the recipient's mind, and knows that the recipient knows that she know this!

The importance of theory of mind to natural language is elaborated by Sperber and Wilson (2002) and by Scott-Phillips (2015), who show that mutual understanding between speakers can often reduce the need for words themselves. As Scott-Phillips puts it, language is *underdetermined*. Even Chomsky (2011) seems to agree. He writes: "Communication relies on largely shared cognoscitive powers, and succeeds insofar as similar mental constructs, background, concerns, and presuppositions allow for similar perspectives to be reached" (p. 10). He does go on, however, to assert that these features are not present in animal communication.

Friederici (2019) raises the question of whether there is a single network in the brain underlying hierarchy processing and concludes, based on brain imaging, that there is not. The hierarchical aspect of language appears to be grounded neurologically in Area 44, part of Broca's area, in the left hemisphere. Theory of mind, in contrast, seems to depend on a bilateral fiber tract involving temporal and parietal lobes and anterior parts of the dorsal fiber tract, but does not include Broca's area. The harmonic and melodic structure of music are also hierarchical in structure, and also seem to activate Broca's area, along with its right-hemisphere homolog. Mathematics is also structured hierarchically, but functions independently of Broca's area (Varley et al., 2005). Friederici concludes "Broca's area in the left hemisphere is crucial for the processing of hierarchy in language, but not for hierarchy processing in other higher-order cognitive domains, and can thus not be viewed as domain-general (p. 6)."

The structure of language, then, is based on the structure of thought, whether in the construction of episodes, the understanding of other people, the invention of music or mathematics, or sheer imagination. These various mental events are hierarchical, and lend their structure to their expression. As Pinker and Jackendoff (2005) remark, "the only reason language needs to be recursive is because its function is to express recursive thoughts (p. 230)."

WHAT OF OTHER SPECIES?

These considerations take the focus away from language as a communication system, and raises questions about the nature of thought itself and about the status of humans relative to other species. Are the hierarchical and generative aspects of thought truly uniquely human, and Berwick and Chomsky (2016) continue to insist, or can we find evidence for them prior to the emergence of humans, or indeed in contemporary non-human species?

For a start, it seems highly unlikely that they could have emerged in a single step in the promiscuous Prometheus a mere 100,000 or so years ago, as Chomsky has repeatedly claimed. Some 6–8 million years spanned the interval between modern humans and our common ancestry with apes (Langergraber et al., 2012), which is a more reasonable time period for the progressive evolution of hierarchical thinking. With respect to language itself, there are suggestions that at least some large-brained prehuman hominins, such as the Neanderthals, were fully verbally competent (e.g., Dediu and Levinson, 2013; Johansson, 2013; Hoffmann et al., 2018). But in the search for generative processes more generally, we can probably go back much further in evolution. Knott (2012) goes so far as to suggest that Chomsky's Minimalist Program and the concept of Merge can be applied recursively to simple sensorimotor actions, such as grasping an object and bringing it to the mouth, activities common to primates species and seemingly intentional. This raises the question of whether Merge itself may have origins long predating language itself. Other aspects of animal action and thought also appear to exhibit at least a level of generativity comparable in kind to that in humans.

Many species, especially birds, do combine different signals. For example, Japanese tits have more than 10 different notes, and combine them to produce different warning signals (Suzuki et al., 2018). The three-note sequence ABC is a warning about predators, while another, D, is a call to attract conspecifics. The combined sequence, ABC-D, is a signal to recruit conspecifics to mob a stationary predator. There has been controversy as to whether this is genuinely combinatorial, retaining the meaning of each constituent, or whether each sequence is treated as a holistic unit (e.g., Bolhuis et al., 2010). At one level, at least, the combining of signals is an instance of Merge, but Merge itself can be considered to have different levels of recursion. Suzuki and Zuberbühler (2018) suggest four levels: 0, Merge, with no combination of element; 1, Merge, with combinations of elements but no recursion; 2, Merge, with merging of elements with previously merged combinations; and 3, Merge; with merging of different merged combinations. Recursive merging of the third type can generate unbounded hierarchical structures, or what Chomsky (1988) called "discrete infinity." There is so far no evidence that any non-human species is capable of this level of Merge, or indeed of creating the vast number of meaningful utterances evident in human language. Even so, the idea of different levels of Merge suggests a degree of continuity rather than an abrupt saltation restricted to humans. And as suggested above, this level of generativity is better regarded as a property of thought, rather than of language itself.

MENTAL TIME TRAVEL

Although it is commonly asserted that mental time travel itself is uniquely human (Tulving, 1985, 2002; Suddendorf and Corballis, 1997, 2007; Bulley et al., 2019), the evidence increasingly suggests that this is not true. It has long been known that honey bees perform waggle dances to indicate the location of food sources (von Frisch, 1967); even Hockett (1960) understood this to be an example of displacement, and it occurs not only in space but also in time (Plath and Barron, 2017). Evidence now also suggests that many vertebrate species have the capacity both to "replay" past events and imagine future ones (Corballis, 2013; but see Suddendorf, 2013). Some of the evidence is behavioral and comes from species as diverse as great apes (e.g., Martin-Ordas et al., 2010; Janmaat et al., 2014), birds (e.g., Clayton et al., 2003), rats (Wilson et al., 2013)., and even cuttlefish (Jozet-Alves et al., 2013).

In general, these studies suggest little of the generativity or expanse of human mental time travel. To be sure, there are prodigious feats of memory itself. The Clark's nutcracker is said to cache some 33,000 seeds in around 7,000 locations every fall and relies on spatial memory to recover them over the winter (Kamil and Balda, 1985). The giant tortoise may not be a creative animal, but seems to have explicit memories lasting up to 9 years (Gutnik et al., 2019). Perhaps because non-human species have no expressive language and therefore cannot verbally describe their memories, we are apt to underestimate their capabilities. Whether these memories are genuinely episodic is perhaps open to question. The Clark's nutcracker may simply remember where the seed are located, without any episodic memory of the act of caching itself.

Some recent studies of memory in vertebrates may offer more compelling evidence of human-like episodic memory for multiple events. In one study, rats remembered many different episodes over intervals of up to 45 min without any evidence of decline in performance (Panoz-Brown et al., 2016). Panoz-Brown et al. (2018) later showed that rats could remember different ordered sequences of odors associated with different contexts, implying memory for structured episodes; to rats, odors appear to be as distinctive and memorable as visual images are to humans. Accuracy was little affected by a delay of 60 min between encoding and testing, or by inserting an unrelated task, implying long-term episodic memory.

Other evidence for mental time travel comes from neurophysiology. Sequences of firing in "place cells" in the rat hippocampus not only track changes in location of the animal in a confined territory, such as a maze, but also track trajectories from past episodes in the maze, as well as possible future trajectories, or even purely exploratory ones. This hippocampal activity also records non-spatial associations tied to past events; the ability of rats to recall past sequences of odors was impaired following chemical suppression of hippocampal activity (Panoz-Brown et al., 2018). These observations appear to have the hallmarks of mental time travel, as though mentally "replaying" the past or imagining the future (Corballis, 2013; Moser et al., 2015).

There is perhaps still some doubt as to whether putative examples of mental time travel in non-human animals have the recursive structure of human imagination. Bulley et al. (2019) suggest that humans go beyond episodic projection to what they term *metaforesight*, the capacity to monitor, control and augment imagined futures, and argue that this superordinate capacity is unique to humans. Metaforesight is analogous to *metamemor*y, the comparable ability to monitor and control recollections of the past (e.g., Cavanaugh, 1982). Hence mental time travel, in humans, at least, may itself be under superordinate control, and hierarchically organized. Bulley et al. (2019) relate the emergence of Acheulian hand axes to dawning metaforesight from some 1.76 million years ago. Even so, we should perhaps not discount the seemingly free trajectories implied by hippocampal recordings in the rat as a form of controlled future planning. Pastalkova et al. (2008) showed that hippocampal recording could predict which way a rat plans to turn on the next trial in a maze, and suggest, "the neuronal algorithms, having evolved for the computation of distances, can also support the episodic recall of events and the planning of action sequences and goals (p. 1327)." Lewis et al. (2019) review evidence that great apes' ability to spontaneously recall past events after long intervals is at least comparable to that in humans, again questioning human uniqueness, and implying a degree of superordinate control. The chimpanzee Panzee communicates with cards and keyboards, and in a typical study watched while researchers hid dozens of objects—fruits, toys, balloons, paper shapes—outside of his enclosure, and when later shown the symbol for each could guide a keeper to where it was hidden. Performance was accurate after 16 h (Menzel, 2005).

It may be true that mental travels, and imagination generally, are more profuse, flexible and "generative" in humans than in other animals. In the study by Martin et al. (2011), mentioned earlier, our human participants had little difficulty recapturing multiple past memories and imagining new scenarios. Such flexibility might also explain the human disposition for storytelling, once we evolved the capacity to externalize. In spite of our ability to bore listeners with seemingly endless exploits from the past, we probably actually remember only a small proportion of the vast number of events that punctuate our lives. It has been suggested that memory capacity may be partly sacrificed for flexibility itself (Tello-Ramos et al., 2019). For example, food-caching animals require extensive long-term memory for later retrieval of caches but show greater proactive interference, suggesting decreased flexibility, than in non-caching animals. In contrast, nomadic animals that move constantly to different environments may show great flexibility in acquiring new information, but also poorer retention of older information that is no longer relevant. On this analysis, humans may lie toward the extreme of high flexibility but relatively poor long-term retention.

This possibility is further elaborated by the historian Fernandez-Armesto (2019), who suggests that it underlies human creativity, and is yet another basis for human uniqueness, albeit at the expense of memory capacity:

> The degree to which humans are, as far as we know, uniquely creative seems vast by comparison with any of the other ways in which we have traditionally been said to excel other animals (p. 3).

Claims of human uniqueness, though, run the risk of what has been called the "human superiority complex" (Villa and Roebroeks, 2014, p. 1) and the safer conclusion is that there are interspecies difference in the deployment of imagination. Claims of human uniqueness seem to progressively dwindle in the face of growing evidence for constructive thinking in non-human animals.

THEORY OF MIND

Just over 40 years ago, Premack and Woodruff (1978) raised the question of whether our closest non-human relative were capable adopting the mental perspective of others. Thirty years later, opinion was still sharply divided. Penn et al. (2008) argued that the failure to recognize the mental discontinuity between animals and humans was "Darwin's mistake," while Call and Tomasello (2008) concluded that chimpanzees do have an understanding of the goals, intentions, perceptions, and knowledge of others, but no understanding of others' beliefs or desires. Even ravens may show an understanding of what unseen birds can see Bugnyar et al. (2016). With respect to language, though, the critical question is whether an individual can understand what another individual *believes*. Over the succeeding decade, there has been some indication that great apes, at least, do have some understanding of what others believe.

The gold standard for assessing theory of mind at the level of belief is what has been termed the Sally-Anne test, designed to assess whether an individual understands that another individual has a false belief. In the original version, a child is shown two dolls, one called Sally and the other called Anne. Sally has a basket and Anne has a box. Sally puts a marble in her basket and leaves, and Anne then puts the marble in her box. Sally then comes back and the child is then asked where she will look for the marble. Autistic children say she will look in the box, but 4–6 year-old children understand that Sally has a false belief and say she will look in the basket (Baron-Cohen et al., 1985). Children aged four understand false belief but children aged three do not (Grosse et al., 2017).

Krupenye et al. (2016) tested three species of great apes (chimpanzees, bonobos and orangutans) on a version of the Sally-Anne test. A person hid an object and left, and the object was then moved. When the person came back, the animals look toward the original location, as though expecting the person to look there. That is, they behaved as though understanding that the person had a false belief. It is possible, though, that the apes had simply learned that people tend to look for things where they last saw them, and were not considering what the person believed. Kano et al. (2019) offer a more exacting test in which the apes saw a video in which an actor saw an object hidden under one of two boxes. The actor then moved behind a barrier that was either translucent or opaque, and the object was shifted to the other box. The eye movements of the apes were recorded, and only apes that had previously experienced the barrier as opaque visually anticipated that the actor would mistakenly look under the box where the object was originally hidden. That is, they were able to judge the actor's belief based on their own past experience.

It may still be the case that theory of mind in apes is not at the level required for human language. An ape may know what another ape is thinking, but may not know that the other ape knows this, which Grice believed was required for meaningful discourse between the two. Again, though, we should be wary of the "human superiority complex" which seems to denigrate all attempts to demonstrate human-like intelligence in other species. And even if theory of mind is at a lower level in apes than in humans, we might still agree with Darwin (1871) that "[T]he difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind" (p. 126).

SO WHAT IS SPECIAL ABOUT HUMANS?

In spite of increasing evidence of cognitive continuity between humans and other species, we humans do seem exceptional in the ability to communicate, manufacture objects, and desecrate the planet. This difference may indeed relate to language, but perhaps less in the cognitive component than in the power of communication itself. This is counter to Chomsky's proposal. To him, the communicative aspect is relatively trivial and uninteresting: "… externalization (hence a fortiori communication) is an ancillary aspect of language, peripheral to its core nature" (Chomsky, 2015a, p. 101).

I suggest here that it is more the communicative than the cognitive aspect that is largely responsible for our dominance on the planet. Imaginative thinking, whether directed to past or future events, or simply to the invention of scenarios, is no doubt adaptive to the individual, and may be common to many other species. What makes it special to humans is the ability to share it. This may be relatively trivial biologically, as Chomsky implies, but hugely important in our capacity to adapt to life on earth.

Most obviously, it simply increases the amount of information. The mental travels of others are incorporated into one's own, albeit with some loss of precision and personal relevance, but with vast increase in scope. The contents of our mental lives are derived as much from others as from personal experience, probably more so. Large portions of our lives are spent in the imaginations of others. It is stories, whether in the form of fiction, soap operas, tales around the campfire, or gossip, that prompted Niles (2010) to rename our species *Homo narrans*—the storytellers.

Through sharing, communities know much more than the individuals within them, and can make that information accessible to all. Preliterate communities told stories that were repeated down the generations, accompanies by song and dance. With the invention of writing, storage of information could become more lasting and accurate, and shared even more widely. Through books, computers, and the internet, communicative sharing had progressively fewer bounds, either geographically or in storage capacity. My own memory and communicative reach now seem lodged mainly in my laptop and i-phone. This is due much more to cultural invention than to any innate disposition, although it remains variable across cultures. There remain indigenous peoples without these technological facilities who are fully endowed cognitively, and may well have retained cognitive abilities in excess of those who live in the modern industrialized state.

THE PRODUCTION PROBLEMS

These advantages, though, raise the question of why other intelligent species have not evolved a comparable capacity to share. At least part of the answer has to do with the mechanisms of externalization, the production of signals that can meet the complexity of generative thought. Hippocampal recordings from the rat brain imply mental traveling well in excess of any signaling, vocal or otherwise, and it is now clear that many non-human species can comprehend much more than they can transmit. Bottlenosed dolphins easily learned human gestural signals instructing them to repeat up to 36 different behaviors, some of them complex and novel, but are themselves unable to make such gestures (Mercado et al., 1998). Two border collies appear to have receptive vocabularies in the hundreds (Kaminski et al., 2004; Pilley and Reid, 2011), but virtually no ability to articulate. Domestic dogs also recognize familiar speakers from their voice quality (Root-Gutteridge et al., 2019). Savage-Rumbaugh et al. (1998) reported that Kanzi, a bonobo, was able to follow simple spoken instructions, made up of several words, at a level comparable to that of 2.5-year-old child. Kanzi is now said to understand some 3,000 spoken words (Raffaele, 2006), but has virtually no ability to speak. Fischer and Hammerschmidt (2019) note that the structure of chimpanzee calls is largely innate, with only limited evidence for modification or conventionalization, while in contrast "comprehension learning may be extremely rapid and open-ended" (p. 1).

Vocal signaling seems especially constrained, especially in non-human primates. Control of the laryngeal muscles in the premotor cortex is only indirect (Simonyan and Horwitz, 2011; Koda et al., 2018), but well developed in humans. Koda et al. (2018) note, though, that even this is not sufficient for articulate speech, which also requires fine motor control of jaws, lips, tongue, and diaphragm—all of which constitute a "unique form of systems integration" (p. 11).

These transformative changes presumably occurred sometime in the course of hominin evolution, well after the split from the great apes.

Vocal production, though, is not the only avenue for externalization. Manual action offers equivalent flexibility and intentionality, as is evident from the signed languages invented by deaf communities, and probably allows greater evolutionary continuity. In contrast to their poor voluntary control of vocalization, non-human primates are well adapted for intentional manual activity, whether in climbing, picking berries, grooming, or play. Attempts to teach great apes to talk have largely failed, but greater success has been achieved using simplified forms of sign language (Patterson and Gordon, 2001) or keyboards containing arrays of word-like symbols (Savage-Rumbaugh et al., 1998). Byrne et al. (2017) list 84 different communicative gestures arising from the studies of great apes' gestures, and note that they are goal directed and intentional, unlike most primate calls. To be sure, there is little evidence for sentence-like structure, but chimpanzee gestures suggest a natural platform for more complex sequences.

These considerations have led many, including myself, to propose that productive language may have originated in manual gestures (e.g., Hewes, 1973; Fano, 1992; Armstrong et al., 1995; Rizzolatti and Arbib, 1998; Corballis, 2002, 2019; Arbib, 2005; Tomasello, 2008). In a recent study, children were asked to communicate about a picture in dyadic pairs, but denied the opportunity to speak. In less than 30 s, 4-to-6-year-old children developed systems of communicating in gestures, which rapidly became abstract and conventionalized. In 6- to 8-year old children, the gestures also showed evidence of grammatical structure. Gesture, then, seems as "natural" as speech (Bohn et al., 2019).

In evolutionary terms, gestures may have developed into pantomime, with increasing sequential properties. Although there is some limited evidence for pantomime in great apes (e.g., Boesch, 1993; Russon and Andrews, 2001), the critical period may have been the Pleistocene, dating from around 2.9 million to about 12,000 years ago, and heralding the emergence of the genus *Homo*. This era saw a tripling of brain size, obligate bipedalism, and the making of stone tools, and is also widely recognized as the period in which hominins established what has been termed the "cognitive niche" (Tooby and DeVore, 1987), establishing social bonding and enhanced communication for survival in the more exposed and dangerous environment of the African savanna. These developments probably established the setting for the emergence of pantomime as a dominant mode of intentional communication, enabling sharing of episodes or plans, perhaps resembling the modern game of charades. In the interests of efficiency, pantomime would become less iconic and more conventionalized by custom, perhaps to resemble modern sign language. Tomasello (2008), for example, writes of the possibility "that the human capacity evolved quite a long way in the service of gestural communication alone, and the vocal capacity is actually a very recent overlay (p. 246)."

The transition from gesture to speech was itself likely to have been gradual, with facial gestures accompanying manual ones (Corballis, 2017). Primates have intentional control over facial movements (Dobson and Sherwood, 2011) and Shepherd and Freiwald (2018) show that facial movements, such as the lip smack, act as visual signals in marmosets in so-called second-person social settings, involving interaction between signaler and audience. Production of these movements recruits areas homologous to Broca's area in humans. Facial movements are also an important component of the sign languages, and people engaged in sign language watch face as much as they watch the hands, sometimes more so (Muir and Richardson, 2005).

In any event, speech itself is primarily gestural, based on movements of the lips, tongue, and larynx, but largely invisible to the viewer, so that sound was added to make them accessible. Indeed the retreat of gestures into the mouth can be regarded as part of the conventionalization process, and an early example of miniaturization (Corballis, 2017), although manual gestures remains an integral accompaniment to speech, even in the blind (Iverson and Goldin-Meadow, 1998). They can improve the speaker's lexical access and fluency (Rauscher et al., 1996), and even reduce the speaker's working memory load (Goldin-Meadow et al., 2001; Wagner et al., 2004).

The critical change that led to the development and vast expansion of communication options may have been the shift to obligate bipedalism. This freed the hands from locomotion, allowing them to play new roles in tool manufacture and communication; it may have been these growing demands that drove the shift to bipedalism itself. In anatomical terms, the change may have been relatively minor, but its consequences were immense. It is perhaps difficult to think of comparable changes with such dramatic consequences, but one possible example might be flight: One small step for bird, one giant leap for birdkind.

Communicative language therefore does seem to have a dramatic influence in human evolution, even if not in the way envisaged by Chomsky. But we may still exaggerate its benefits. As a sharing device, language is far from perfect; many complex thoughts or emotions seem to defy description. Albert Einstein is said to have developed the theory of relativity by imagining himself traveling on a light beam, and only with difficulty rendered it in mathematical terms. Often, too, it is more adaptive not to keep secrets and not have thoughts shared. In most non-human species, vocalization is largely involuntary, and acts as an "honest signal," whereas language allows for deception, through lying and the dissemination of fake news. The 7,000 languages of the world are also testimony to the use of language as a moat, enabling sharing within groups but acting as a barrier between them—language seems to operate as much to prevent communication as to enable it. In many respects, then, language is an exclusionary and even destructive force. We should remember that at least 20 different species identified as hominins, but only humans survive, and have done so for only a few hundred thousand years.

SOME CONCLUSION

Macphail's work captured something of the dilemma facing psychology in the latter part of the 20th century. Behavioral psychology was largely built on the commonalities of learning principles across different species, whereas the cognitive revolution was built on computational principles, with language as a primary exemplar supposedly unique to humans. The Rubicon seemed as impenetrable as it had seemed to Müller a century earlier.

In his later writing, though, Macphail was well aware that the science of animal behavior had moved on from behaviorism. Macphail and Bolhuis (2001), for example, wrote, "The behaviorist domination of experimental psychology had many unfortunate consequences, amongst them a divorce from those (now referred to as ethologists) who studied the behavior of animals" (p. 343). They go on to note that learning is often adapted to specific contexts. For example, rats seem especially adapted to learning about odors. Phobias, such as fear of snakes, seem to be learned much more rapidly than other forms of associative learning. Birds that store quantities of food, such as the afore-mentioned Clarke's nutcracker, seem to have better spatial learning than non-storers, and indeed have larger hippocampi. Macphail and Bolhuis carefully review such examples, and conclude that there are no convincing examples of differences in the processes of learning and memory themselves, either within or between species. "The outcome," they write, "supports a general process as opposed to an ecological account of cognition" (p. 361).

As noted at the beginning of this article, though, Macphail and Bolthuis continued to hold language as a special case, raising the intelligence of humans above that of any other species. Developments over the past few decades have seen a blurring of this gap. The essence of language seems to lie in generative thought rather than in any power of communication. This not only narrows the gap between humans and other species, but also broadens the concept of intelligence, which can now be taken to include theory of mind, imaginative thinking, and creativity. Macphail and Bolthuis do concede that they have omitted discussion of whether apes possess theory of mind, which Byrne (1995) had earlier suggested to be critical to the evolution of intelligence. As suggested earlier, we may add mental time travel and even imagination as aspects of intelligence that may go far back in evolution. We saw too that humans might sacrifice some memory capacity as a trade for enhanced creativity and imagination. It now seems likely that these capacities do vary even among vertebrates other than humans, contrary to Macphail's suggestion that all non-human vertebrates are of equal intelligence. The picture seems complicated by suggestions of interactions between different capacities, as in the idea of a tradeoff between memory capacity and creativity. Given the advances in behavioral techniques and neurophysiological investigation, the challenge is to map out the mental capacities of different species, with perhaps less of an imperative to consider that humans are different.

This is not to undermine Macphail's work and influence. Revisiting his work now reminds us that there was much value in the behaviorist movement that dominated psychology for much of the 20th century, establishing an evolutionary continuity that was largely overlooked following the "cognitive revolution." He no doubt exaggerated the uniformity of intellect between species, at least if one overlooks language, but that kind of challenge spurs more critical research and better definitions of what is meant by intelligence. At the same time he did conform to the changing zeitgeist (with very little reference to Chomsky himself) emphasizing the special qualities possessed by humans, which may well have seeded the subsequent attempts to demonstrate cognition in other species, from birds to mammals (and perhaps not forgetting the worms). What may have seemed a stark contrast between humans and all other species has blended into a continuum, albeit one with added complexities and divergences. Macphail's writing set up the challenge, and attempting to answer it can only advance our knowledge of how animals think, and where humans fit into the overall scheme of things.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

Arbib, M. A. (2005). From monkey-like action recognition to human language: a n evolutionary framework for neurolinguistics. *Behav. Brain Sci.* 28, 105–168. doi: 10.1017/s0140525x05000038

Armstrong, D. F., Stokoe, W. C., and Wilcox, S. E. (1995). *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90022-8

Berwick, R. C., and Chomsky, N. (2016). *Why Only Us? Language and Evolution*. Cambridge, MA: The MIT Press.

Bickerton, D. (2014). *More Than Nature Needs: Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.

Boesch, C. (1993). "Aspects of transmission of tool-use in wild chimpanzees," in *Tools, Language and Cognition in Human Evolution*, eds K. R. Gibson and T. Ingold (Cambridge, UK: Cambridge University Press), 171–183.

Bohn, M., Kachel, G., and Tomasello, M. (2019). Young children spontaneously recreate core properties of language in a new modality. *Proc. Natl. Acad. Sci. U.S.A.* 116, 26072–26077. doi: 10.1073/pnas.1904871116

Bolhuis, J. J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11, 747–759. doi: 10.1038/nrn2931

Bugnyar, T., Reber, S. A., and Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nat. Commun.* 7:10506. doi: 10.1038/ncomms10506

Bulley, A., Redshaw, J., and Suddendorf, T. (2019). "The future directed functions of the imagination: From prediction to metaforesight," in *The Cambridge Handbook of the Imagination*, ed. A. Abraham (Cambridge: Cambridge University Press).

Butts, R. E. (1988). The grammar of reason: Hamann's challenge to Kant. *Synthese* 75, 251–283. doi: 10.1007/bf00872981

Byrne, R. W. (1995). *The Thinking Ape: Evolutionary Origins of Intelligence*. Oxford: Oxford University Press.

Byrne, R. W., Cartmill, E., Genty, E., Hobaiter, C., and Tanner, J. E. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Anim. Cogn.* 20, 755–769. doi: 10.1007/s10071-017-1096-4

Call, J., and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192. doi: 10.1016/j.tics.2008.02.010

Cavanaugh, J. C. (1982). Metamemory: A critical examination. *Child Dev.* 53, 11–28. doi: 10.2307/1129635

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1959). A review of Skinner's "Verbal Behavior.". *Language* 35, 26–58.

Chomsky, N. (1988). *Language and the Problem of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2010). "Some simple evo devo theses: How true might they be for language?," in *The Evolution of Human Language*, eds R. K. Larson, V. Déprez, and H. Yamakido (Cambridge: Cambridge University Press), 45–62. doi: 10.1017/cbo9780511817755.003

Chomsky, N. (2011). Biolinguistic explorations: design, development, evolution. *Int. J. Philos. Stud.* 15, 1–21. doi: 10.1080/09672550601143078

Chomsky, N. (2015a). Some core contested concepts. *J. Psychol. Res.* 44, 91–104. doi: 10.1007/s10936-014-9331-5

Chomsky, N. (2015b). *The Minimalist Program: 20th Anniversary Edition*. Cambridge, MA: MIT Press.

Clayton, N. S., Bussey, T. J., and Dickinson, A. (2003). Can animals recall the past and plan for the future? *Trends Cogn. Sci.* 4, 685–691. doi: 10.1038/nrn1180

Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.

Corballis, M. C. (2009). Mental time travel and the shaping of language. *Exp. Brain Res.* 192, 553–560. doi: 10.1007/s00221-008-1491-9

Corballis, M. C. (2013). Mental time travel: a case for evolutionary continuity. *Trends Cogn. Sci.* 17, 5–6. doi: 10.1016/j.tics.2012.10.009

Corballis, M. C. (2017). *The Truth About Language: What It Is, and Where It Came From*. Chicago, IL: University of Chicago Press.

Corballis, M. C. (2019). Language, memory, and mental time travel: an evolutionary perspective. *Front. Hum. Neurosci.* 13:217. doi: 10.3389/fnhum.2019.00217

Darwin, C. (1859). *On the Origin of Species*. London: John Murray.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. New York, NY: Appleton.

Deane, P. (1992). *Grammar in Mind: Explorations in Cognitive Syntax*. Berlin: Mouton de Gruyter.

Dediu, D., and Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Front. Psychol.* 4:397. doi: 10.3389/fpsyg.2013.00397

Descartes, R. (1984). *The Philosophical Writings of Descartes, Trans*, Vol. 2, eds J. Cottingham, R. Stoothoff, and D. Murdoch (Cambridge: Cambridge University Press).

Dobson, S., and Sherwood, C. C. (2011). Correlated evolution of brain regions involved in producing and processing facial expressions in anthropoid primates. *Biol. Lett.* 7, 86–88. doi: 10.1098/rsbl.2010.0427)

Dor, D. (2015). *The Instruction of Imagination: Language as a Social Communication Technology*. New York, NY: Oxford University Press.

Dunbar, R. (2004). *The Human Story*. London: Faber & Faber.

Evans, N. (2009). *Dying Words: Endangered Languages and What They Have to Tell Us*. Oxford: Wiley-Blackwell.

Evans, N., and Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behav. Brain Sci.* 32, 429–492. doi: 10.1017/S0140525X0999094X

Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahâ. *Curr. Anthropol.* 46, 621–646. doi: 10.1086/431525

Fano, G. (1992). *The Origins and Nature of Language*. Bloomington, IN: Indiana University Press.

Fernandez-Armesto, F. (2019). *Out of Our Minds: What We Think and How We Came to Think It*. London: Oneworld Publications.

Fischer, J., and Hammerschmidt, K. (2019). Towards a new taxonomy of primate vocal production learning. *Philos. Trans. R. Soc. B* 375:20190045. doi: 10.1098/rstb.2019.0045

Friederici, A. D. (2019). Hierarchy processing in human neurobiology: How specific is it? *Philos.Trans. R. Soc. B* 375:20180391. doi: 10.1098/rstb.2018.0391

Gärdenfors, P., and Osvath, M. (2010). "Prospection as a cognitive precursor to symbolic communication," in *The Evolution of Human Language*, eds R. K. Larson, V. Déprez, and H. Yamakido (Cambridge: Cambridge University Press), 103–114. doi: 10.1017/cbo9780511817755.007

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., and Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychol. Sci.* 12, 516–522. doi: 10.1111/1467-280.00395

Grice, H. P. (1989). *Studies in the Ways of Words*. Cambridge, MA: Cambridge University Press.

Grosse, W. C., Schreiber, J., Singer, T., Steinbeis, N., and Friederici, A. D. (2017). White matter maturation is associated with the emergence of Theory of Mind in early childhood. *Nat. Commun.* 8:14692. doi: 10.1038/ncomms14692

Gutnik, T., Weissenbacher, A., and Kuba, M. J. (2019). The underestimated giants: operant conditioning, visual discrimination and long-term memory in giant tortoises. *Anim. Cogn. Adv. Publ.* 23, 159–167. doi: 10.1007/s10071-019-01326-6

Hewes, G. W. (1973). Primate communication and the gestural origins of language. *Curr. Anthropol.* 14, 5–24.

Hobbes, T. (1651). *Leviathan: Part 1. Of man*. New York, NY: Sparknotes.

Hockett, C. F. (1960). The origins of speech. *Sci. Am.* 203, 88–96. doi: 10.1038/scientificamerican0960-88

Hoffmann, D. L., Standish, C. D., García-Diez, M., Pettitt, P. B., Milton, J. A., and Pike, A. W. G. (2018). U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science* 359, 912–915. doi: 10.1126/science.aap7778

Iverson, J. M., and Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature* 396:228. doi: 10.1038/24300

Janmaat, K. R. L., Polansky, L., Ban, S. D., and Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. *Proc. Natl. Acad. Sci.U.S.A.* 111, 16343–16348. doi: 10.1073/pnas.1407524111

Johansson, S. (2013). The talking Neandertals: What do fossils, genetics, and archeology say? *Biolinguistics* 7, 35–74.

Jowett, B. (1892). *The Dialogues of Plato*, Vol. 4. Oxford: Oxford University Press.

Jozet-Alves, C., Bertin, M., and Clayton, N. S. (2013). Evidence of episodic-like memory in cuttlefish. *Curr. Biol.* 23, R1033–R1036. doi: 10.1016/j.cub.2013.10.021

Kamil, A. C., and Balda, R. P. (1985). Cache recovery and spatial memory in Clark's Nutcrackers (*Nucifraga columbiana*). *J. Exp. Psychol.* 85, 95–111. doi: 10.1037/0097-7403.11.1.95

Kaminski, J., Call, J., and Fischer, J. (2004). Word learning in a domestic dog: Evidence for "fast mapping". *Science* 304, 1682–1683. doi: 10.1126/science.1097859

Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., and Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20904–20909. doi: 10.1073/pnas.1910095116

Knott, A. (2012). *Sensorimotor Cognition and Natural Language Syntax*. Cambridge, MA: MIT Press.

Koda, H., Kunieda, T., and Nishimura, T. (2018). From hand to mouth: monkeys require greater effort in motor preparation for voluntary control of vocalization than for manual actions. *R. Soc. Open Sci.* 5, 180879. doi: 10.1098/rsos.180879

Krupenye, C., Kano, F., Hirata, S., Call, J., and Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science* 354, 110–114. doi: 10.1126/science.aaf8110

Langergraber, K. E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., and Vigilant, L. (2012). Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. *Proc. Natl. Acad. Sci U.S.A.* 109, 15716–15721. doi: 10.1073/Pnas.1211740109

Lewis, A., Bernsten, D., and Call, J. (2019). Long term memory of past events in great apes. *Curr. Dir. Psychol. Sci.* 28, 117–123. doi: 10.1177/0963721418812781

Locke, J. (1690/2017). *An Essay Concerning Human Understanding*. London: Scolar.

Macphail, E. M. (1982). *Brain and Intelligence in Vertebrates*. Oxford: Clarendon Press.

Macphail, E. M., and Bolhuis, J. J. (2001). The evolution of intelligence: Adaptive specializations versus general process. *Biol. Rev.* 76, 341–364. doi: 10.1017/s146479310100570x

Martin, V. C., Schacter, D. L., Corballis, M. C., and Addis, D. R. (2011). A role for the hippocampus in encoding simulations of future events. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13858–13863. doi: 10.1073/pnas.1105816108

Martin-Ordas, G., Haun, D., Colmenares, F., and Call, J. (2010). Keeping track of time: evidence for episodic-like memory in great apes. *Anim. Cogn.* 13, 331–340. doi: 10.1007/s10071-009-0282-4

Menzel, C. R. (2005). "Progress in the study of chimpanzee recall and episodic memory," in *The Missing Link in Cognition: Origins of Self-Reflective*

Consciousness, eds H. S. Terrace and J. Metcalfe (Oxford: Oxford University Press), 188–224. doi: 10.1093/acprof:oso/9780195161564.003.0008

Mercado, E. III, Murray, S. O., Uyeyama, R. K., Pack, A. A., and Herman, L. M. (1998). Memory for recent actions in the bottlenosed dolphin (*Tursiops truncatus*): Repetition of arbitrary behaviors using an abstract rule. *Anim. Learn. Behav.* 26, 210–218. doi: 10.3758/bf03199213

Moser, M.-B., Rowland, D. C., and Moser, E. I. (2015). Place cells, grid cells, and memory. *Cold Spring Harb. Perspect. Biol.* 7:a21808. doi: 10.1101/cshperspect.a021808

Muir, L. J., and Richardson, I. E. G. (2005). Perception of sign language and its application to visual communication for deaf people. *J. Deaf Stud. Deaf Educ.* 10, 390–401. doi: 10.1093/deafed/eni037

Müller, F. M. (1861). *Lectures on the Science of Language*, Vol. 1. London: Longmans.

Niles, J. D. (2010). *Homo Narrans: The Poetics and Anthropology of Oral Literature*. Philadelphia: University of Pennsylvania Press.

Panoz-Brown, D., Corbin, H. E., Dalecki, S. J., Gentry, M., Brotheridge, S., and Crystal, J. D. (2016). Rats remember items in context using episodic memory. *Curr. Biol.* 26, 2821–2826. doi: 10.1016/j.cub.2016.08.023

Panoz-Brown, D., Iyer, V., Carey, L. M., Sluka, C. M., Rajic, G., and Crystal, J. D. (2018). Replay of episodic memories in the rat. *Curr. Biol.* 28, 1628–1634.

Pastalkova, E., Itskov, V. A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science* 321, 1322–1327. doi: 10.1126/science.1159775

Patterson, F. G. P., and Gordon, W. (2001). "Twenty-seven years of project Koko and Michael," in *All Apes Great and Small: Vol 1. African Apes*, eds B. Galdikas, N. E. Briggs, L. K. Sheeran, and J. Goodall (New York, NY: Kluver), 165–176. doi: 10.1007/0-306-47461-1_15

Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 108–178. doi: 10.1017/S0140525X08003543

Pilley, J. W., and Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behav. Process.* 86, 184–195. doi: 10.1016/j.beproc.2010.11.007

Pinker, S., and Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition* 95, 201–236. doi: 10.1016/j.cognition.2004.08.004

Plath, J. A., and Barron, A. B. (2017). The evolution of honey bee dance communication: a mechanistic perspective. *J. Exp. Biol.* 220, 4339–4346. doi: 10.1242/jeb.142778

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Raffaele, P. (2006). *Speaking Bonobo. Smithsonian Magazine*. Washington, D.C: Smithsonian Institution Offices.

Rauscher, F., Krauss, R. M., and Chen, Y. (1996). Gesture, speech and lexical access: the role of lexical movements in speech production. *Psychol. Sci.* 7, 226–231. doi: 10.1111/j.1467-9280.1996.tb00364.x

Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194. doi: 10.1016/s0166-2236(98)01260-0

Root-Gutteridge, H., Ratcliffe, V. F., Korzeniowska, A. T., and David Reby, D. (2019). Dogs perceive and spontaneously normalize formant-related speaker and vowel differences in human speech sounds. *Biol.Lett.* 15:20190555. doi: 10.1098/rsbl.2019.0555

Russon, A., and Andrews, K. (2001). Orangutan pantomime: elaborating the message. *Biol. Lett.* 7, 627–630. doi: 10.1098/rsbl.2010.0564

Savage-Rumbaugh, S., Shanker, S. G., and Taylor, T. J. (1998). *Apes, Language, and the Human Mind*. New York, NY: Oxford University Press.

Scott-Phillips, T. (2015). *Speaking Our Minds: Why Human Communication is Different, and How Language Evolved to Make It Special*. Basingstoke: Palgrave Macmillan.

Shepherd, S. V., and Freiwald, W. A. (2018). Functional networks for social communication in the macaque monkey. *Neuron* 99, 413–420. doi: 10.1016/j.neuron.2018.06.027

Simonyan, K., and Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist* 17, 197–208. doi: 10.1177/1073858410386727

Skinner, B. F. (1957). *Verbal Behavior*. New York, NY: Appleton-Century-Crofts.

Sperber, D., and Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind Lang.* 17, 3–23. doi: 10.1111/1468-0017.00186

Suddendorf, T. (2013). Mental time travel: continuities and discontinuities. *Trends Cogn. Sci.* 17, 151–152. doi: 10.1016/j.tics.2013.01.011

Suddendorf, T., and Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Soc. Gen. Psychol. Monogr.* 123, 133–167.

Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behav. Brain Sci.* 30, 299–351. doi: 10.1017/S0140525X07001975

Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS Biol.* 16:e2006532. doi: 10.1371/journal.pbio.2006532

Suzuki, T. N., and Zuberbühler, K. (2018). Animal syntax. *Curr. Biol.* 29, R663–R682.

Tello-Ramos, M. C., Branch, C. L., Kozlovsky, D. Y., Pitera, A. M., and Pravosudov, V. V. (2019). Spatial memory and cognitive flexibility trade-offs: to be or not to be flexible, that is the question. *Anim. Behav.* 147, 129–136. doi: 10.1016/j.anbehav.2018.02.019

Tomasello, M. (2008). *The Origins of Human Communication*. Cambridge, MA: MIT Press.

Tomasello, M. (2009). Universal grammar is dead. *Behav. Brain Sci.* 32, 470–471. doi: 10.1017/S0140525X09990744

Tooby, J., and DeVore, I. (1987). "The reconstruction of hominid evolution through strategic modeling," in *The Evolution of Human Behavior: Primate Models*, ed. W. G. Kinzey (Albany, NY: SUNY Press), 183–237.

Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1–12.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114

Varley, R. A., Klessinger, N. J. C., Romanowski, C. A. J., and Siegal, M. (2005). Agrammatic but numerate. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3519–3524. doi: 10.1073/pnas.0407470102

Villa, P., and Roebroeks, W. (2014). Neandertal demise: An archaeological analysis of the modern human superiority complex. *PLoS One* 9:e96424. doi: 10.1371/journal.pone.0096424

von Frisch, K. (1967). *The Dance Language and Orientation Of Honeybees*. Cambridge: Harvard University Press.

von Humboldt, W. (1836/1999). *On the Diversity of Human Language Construction and Its Influence on the Mental Development Of the Human Species*. Cambridge: Cambridge University Press.

Wagner, S., Nusbaum, H., and Goldin-Meadow, S. (2004). Probing the mental representation of gesture: is handwaving spatial? *J. Mem. Lang.* 50, 395–407. doi: 10.1016/j.jml.2004.01.002

Watson, J. B. (1928). *Psychological Care of Infant and Child*. New York, NY: W. W. Norton.

Wilson, A. G., Pizzo, M. J., and Crystal, J. D. (2013). Event-based prospective memory in the rat. *Curr. Biol.* 23, 1089–1093. doi: 10.1016/j.cub.2013.04.067

Zuberbühler, K. (2019). Evolutionary roads to syntax. *Anim. Behav.* 151, 259–265. doi: 10.1016/j.anbehav.2019.03.006

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 30 April 2020
doi: 10.3389/fpsyg.2020.00846

Reduced Frequency of Knowledge of Results Enhances Acquisition of Skills in Rats as in Humans

Alliston K. Reid and Paige G. Bolton Swafford*

Department of Psychology, Wofford College, Spartanburg, SC, United States

Macphail's (1985) null hypothesis challenged researchers to demonstrate any differences in intelligence between vertebrate species. Rather than focus on differences, we asked whether rats would show the same unexpected, counterintuitive features of skill learning observed in humans: Factors that degrade performance during acquisition often enhance performance in a subsequent retention/autonomy phase. Providing post-trial "knowledge of results" (KR) on 30–67% of trials instead of 100% degrades accuracy, yet increases retention in a subsequent phase without KR. We tested this feature by providing three groups of rats with KR on every trial (100% KR), 67% KR, or 0% KR. We also provided operant feedback in every trial for completing the left-right lever-press skill (food for correct sequences, timeout for all others). In the autonomy phase, we assessed their ability to complete the skill independently—in the absence of differential cues and KR feedback. In agreement with human performance in the autonomy phase, 67% KR yielded higher skill accuracy than providing 100% KR. Also, providing 67% KR improved skill accuracy above that observed with operant feedback alone (0% KR). Rather than degrading performance during acquisition, the 67% KR condition yielded unexpected higher accuracy than the other conditions. Accuracy increased systematically across our extended acquisition phase, which provided each rat with over 3600 trials compared to 20–30 trials for human studies. Providing limited KR promoted skill learning in rats as it does in humans, consistent with the conjecture that both species share common learning processes. Introducing difficulties to rats during training improved their autonomy.

Keywords: skill learning, knowledge of results, operant feedback, frequency of KR, autonomy, comparative psychology, Rats, Humans

Edited by:
Thomas Zentall,
University of Kentucky, United States

Reviewed by:
Federico Sanabria,
Arizona State University, United States
Carter W. Daniels,
Columbia University, United States

****Correspondence:***
Alliston K. Reid
reidak@wofford.edu

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *28 December 2019*
Accepted: *06 April 2020*
Published: *30 April 2020*

Citation:
Reid AK and Swafford PGB (2020) Reduced Frequency of Knowledge of Results Enhances Acquisition of Skills in Rats as in Humans. Front. Psychol. 11:846. doi: 10.3389/fpsyg.2020.00846

INTRODUCTION

Skill learning has been widely studied in humans, but far less in rats. In the last few decades, researchers have often focused on a counterintuitive feature of skill learning in humans related to the historical distinction between learning and performance dating back to Thorndike's (1927) law of effect. Empirical reviews by Schmidt and Bjork (1992), Soderstrom and Bjork (2015), and Johnson and Proctor (2017) have described this common feature across several motor- and verbal-learning paradigms: Factors that degrade performance during acquisition often enhance performance in a subsequent retention condition (see also Bjork and Bjork, 1992, 2014; Kantak and Winstein, 2012). Soderstrom and Bjork explained that introducing "desirable difficulties" (Bjork and Bjork, 2014) for the learner during acquisition can enhance later retention because the cognitive processes active during these "difficulties" link new information with knowledge that already exists in memory.

As implied above, skill-learning studies usually consist of two phases: an acquisition phase in which the skill is acquired in the presence of useful stimuli or feedback from performance, and a retention phase in which retention of the skill is measured when those cues and feedback are no longer provided. In studies with non-humans such as rats or pigeons, we usually call this retention phase an "autonomy" phase because we are interested in the degree to which the animal can complete the skill independently—in the absence of earlier guiding cues or informative feedback. Humans typically seek autonomy as we acquire behavioral skills related to sports, driving, military training, dance, our jobs, and games of skill. Applied behavior analysis often promotes skill learning and autonomy in children with developmental disabilities. Extensive practice of behavioral skills usually leads to autonomy in humans and trained animals, even rats and pigeons (e.g., Helton, 2007a,b; Reid et al., 2010, 2013a).

Behavioral skill learning (as distinguished from cognitive skills such as playing chess) could focus on the roles of (a) antecedent stimuli, or on (b) feedback from performance. Rat studies have focused on anticipatory cues, such as transfer of stimulus control from discriminative stimuli to new exteroceptive or endogenous cues that develop during practice. For example, Mackintosh (1965, 1974) described examples of transfer to proprioceptive control as rats run through mazes and suggested that "if sufficient training were given on a maze problem, control was gradually transferred from exteroceptive to proprioceptive stimuli" (1974, p. 554). Several rat and pigeon studies have observed the same counterintuitive feature of human skill learning: *Less effective* cues and *more difficult* behavioral skills degrade accuracy during acquisition phases, yet *enhance* accuracy in a subsequent autonomy phase (e.g., Rats: Reid et al., 2010, 2013a,b, 2014, 2017; Pigeons: Fox et al., 2014). We often described these observations with the metaphor: "Holding your child's hand too much may delay or prevent autonomy."

In contrast, skill-learning studies with humans have historically focused on feedback from performance. While many types of feedback may be provided, the most influential type has been "knowledge of results" (KR): post-trial feedback about the success of a behavior during skill acquisition (Adams, 1987; Wulf and Schmidt, 1989; Newell, 1991; Wulf and Shea, 2002; Wolpert et al., 2011). One might predict that providing KR on every trial (100% KR) would lead to faster acquisition and stronger retention. However, Winstein and Schmidt (1990) demonstrated a counterintuitive feature of providing different relative frequencies of KR as feedback to human participants: Practice with reduced 50% KR improved skill retention over 100% KR, even though performance during acquisition suffered. This observation has been widely (although not universally) replicated with different types of skill learning in humans. Similarly, Wulf and Schmidt (1989) evaluated the effect of KR on a more complex behavioral skill (generalized motor programs) by comparing 67% KR versus 100% KR. Practice with reduced 67% KR was more beneficial to the transfer of skill learning even with this more complex behavioral skill. These studies were consistent with the claim (Schmidt and Bjork, 1992; Bjork and Bjork, 2014) that introducing difficulties for the learner during acquisition can enhance later retention.

Unfortunately, most human studies of skill learning ignore the other type of feedback: the operant consequences of responding such as reinforcement and punishment effects. Studies with rats or pigeons consistently specify these operant consequences. To our knowledge, no published rat or pigeon studies have included KR feedback with operant feedback. Therefore, it is unknown whether rats and humans would respond similarly to acquisition procedures that provide KR feedback. If rats and humans show not only the same basic features of skill learning, but also the same counterintuitive features, this would seem like much stronger evidence that skill learning in rats and humans may involve similar processes. Therefore, the purpose of this study was to begin to answer that very question.

This experiment combined operant feedback with KR feedback, so it is helpful to clarify the difference between the procedures, as well as the ways in which each affects skill acquisition and retention. However, our understanding of both terms and their mechanisms of action have changed substantially over the last century. Thorndike (1927) considered both types of post-trial feedback to be consistent with his Law of Effect: food delivery and KR indicating performance successes would both strengthen the stimulus-response connection, whereas extinction and KR indicating performance errors would weaken the connection. This view changed substantially after Salmoni et al. (1984) reviewed many studies of skill learning that provided KR feedback to human participants and asked: How does KR work? They emphasized the surprising observation that reduced frequencies of KR typically led to greater retention of motor skills than when KR was presented on every trial. Their "guidance hypothesis" emphasized the informative, goal-directed properties of KR feedback, which was more compatible with information-processing views of human cognition than the reinforcing properties of feedback. Today, KR in human studies is generally believed to provide opportunities to allocate cognitive resources such as attention and encoding processes, or to effectively link new information with existing memory. Several articles have expressed doubts as to whether laboratory animals would have the information-processing capacities to benefit from KR feedback (e.g., Adams, 1987; Wulf and Schmidt, 1989; Wulf and Shea, 2002). This claim assumes that Macphail's's (1985) null hypothesis is wrong, but (until now) no studies with laboratory animals have tested the assumption that laboratory animals and humans would react differently to skill learning procedures that provide KR.

Our behavioral skill was a simple left-right (L-R) lever-press sequence guided by blue and green panel lights above the two levers. All other sequences were errors. The second press terminated the trial and provided both operant and KR feedback. Operant feedback was always immediate and preceded KR feedback. Following the KR procedures of Wulf and Schmidt (1989) in a between-groups design, we compared the effects of 67% KR with those of 100% KR. As a control condition, we also included a 0% KR group that received operant feedback every trial, but never KR feedback, in both acquisition and autonomy phases. Normally, pellet delivery and KR feedback would be perfectly correlated for the 100% KR group. To distinguish between these two factors, we degraded this correlation by providing reinforcement for correct trials on a random ratio 2 (RR-2) schedule of reinforcement. If KR affects rats in the same way as in humans, then sequence accuracy during the autonomy phase should be higher for the 67% KR group than the 100% KR group, but accuracy for the 67% KR group during the acquisition phase should be lower than the 100% KR group. If providing KR promotes skill learning beyond the effects of operant feedback alone, then accuracy for the 67% KR group should exceed that of the 0% KR group.

METHODS

Subjects

Twenty-six naïve 4-month-old female Long Evans rats (*Rattus norvegicus*) were housed individually in a facility that maintained constant temperature and humidity on a 12:12-h light:dark cycle. Body weight was maintained at 80–85% of free-feeding weight by providing supplemental food (Tekland Rodent Diet) after daily sessions in home cages with water freely available.

Apparatus

We utilized four standard Med Associates operant chambers for rats, measuring 30 cm × 24 cm × 22 cm. Each chamber was located inside an isolation chamber containing a ventilation fan and a 7-W nightlight. A sound generator produced constant 65-db white noise in each chamber. Each operant chamber contained two retractable levers on the front wall and two non-retractable levers on the rear wall. Each pair of levers was separated by 16.5 cm, center to center, and located 6 cm above the floor. The 5 cm × 5 cm magazine hopper was centered between the two response levers on the front wall, 3 cm above the floor. A 2.5-cm 28-V white panel lamp was located 2.5 cm above the two front levers, and a 28-V houselight was located at the center top of the rear wall. In some conditions, 28-V blue or green LEDs replaced the two white panel lamps on the rear wall. The wavelengths of these LEDs (blue: 465 nm, green: 515 nm) were selected to approximately match the peak photopigment sensitivity of the UV and M cones (358 and 510 nm, respectively) in Long-Evans rats (Jacobs et al., 2001). A Sonalert was available to produce 1000-Hz tones. The pellet dispenser provided 45-mg Research Diet pellets. All four chambers were controlled by a single Lenovo personal computer located in an adjacent room and programmed in MED-PC IV, which implemented all experimental conditions and recorded every event and time of occurrence with 10-ms resolution.

Procedure

We randomly assigned the 26 naïve rats to three experimental groups of 8,9 rats each. The groups differed in the frequency of qualitative KR feedback about the accuracy of responding, which was provided by the two white front panel lamps and the Sonalert tone for responding on the two rear levers in a discrete-trials procedure. The 0% KR group received no KR on any trial; the 100% KR group received KR on every trial; and the 67% KR group received KR on 67% of the trials.

We exposed all three groups to a sequence of training procedures that reinforced lever pressing on the front wall. Once lever-press training was completed, the experiment consisted of two experimental phases on the rear wall. The purpose of the 30-session Acquisition Phase was to allow each group of rats to learn the left-right (L–R) lever-press sequence (the skill) on the rear wall, guided by blue (left) and green (right) LED panel lights. In the final 10-session Autonomy Phase, the reinforcement contingencies were unchanged, but KR was eliminated, and the rear blue and green LED panel lights were replaced with white LEDs. Our primary measure was percentage L–R accuracy, which allowed us to assess (a) the speed in which the three groups learned the L–R sequence during the Acquisition Phase, and (b) the degree of L–R autonomy once we eliminated the differential stimuli and KR in the Autonomy Phase.

Training

Lever-press training

We exposed all rats to an autoshaping procedure for three sessions. The procedure inserted the right front retractable lever (adjacent to the food hopper) into the chamber 8 s before delivering a pellet, followed by a 52-s intertrial interval (ITI). Each press on that lever or any of the three other levers was reinforced, independent of the 8-s lever insertion. Sessions ended with the earlier of 45 min or 80 pellets.

We next exposed each rat to a shaping procedure in which presses to the front right retractable lever or either rear lever produced pellet delivery. The white panel lamp over the front right lever was illuminated, but the rear panel lamps remained off. Pellet deliveries were followed by 3-s ITI, signaled by extinguishing that panel light and the houselight. Sessions ended with the earlier of 30 min or 45 pellets. This training procedure terminated when the rat earned 45 pellets/session for three sessions. All subsequent conditions required subjects to press levers on the rear wall, so levers on the front wall were retracted for the rest of the experiment.

Rear wall

We exposed each subject to training sessions of fixed ratio-1 (FR-1) for pressing the right lever on the rear wall. The green panel light above that lever was illuminated, while the other panel lamps remained extinguished. Food delivery was followed by a 3-s ITI, signaled by extinguishing the panel light and houselight. Sessions ended with the earlier of 30 min or 45 reinforcers. This training procedure terminated when the rat earned 45 pellets/session for three sessions.

Following this training, we exposed each subject to training sessions for pressing the left lever on the rear wall on FR-1.

The blue panel light above that lever was illuminated, while the other panel lamps remained extinguished. Otherwise, conditions were the same as training on the right lever. The purpose of these rear-wall training sessions was to ensure all subjects received approximately equal exposure to the reinforcement conditions on the rear wall before the experiment began, given that subjects required differing amounts of lever-press training on the front wall.

Experiment: L–R Acquisition Phase

The Acquisition Phase lasted 30 sessions for all subjects. It required subjects to complete a L–R lever-press sequence guided by blue (left) and green (right) panel lights in a discrete-trials procedure. No feedback about response accuracy was provided until two lever presses occurred (correct or not), which always produced a 100-ms tone "beep" end-of-trial marker. At the beginning of each trial, the houselight and rear blue (left) and green (right) panel lights were illuminated. Pressing the left lever turned off the left panel light, leaving only the right (green) panel light on. Similarly, pressing the right lever first would turn off the right panel light, leaving only the left (blue) panel light on. This was intended to reduce perseveration on either lever. Completion of the correct L–R response sequence turned off the houselight and panel lights. In addition, a food pellet was delivered with probability 0.5, producing a random ratio 2 (RR-2) reinforcement schedule. The RR-2 schedule was intended to degrade the normal correlation in discrete-trials procedures containing both food delivery and KR feedback. Incorrect trials contained any other 2-response sequence (L–L, R–L, R–R) and resulted in a 3-s timeout (TO) in which the houselight and panel lights were extinguished. Responding during TO had no programmed consequences. The next trial began with the illumination of the houselight and blue and green panel lights. Sessions lasted for the earlier of 45 min or until 120 correct L–R trials occurred (producing approximately 60 pellets on the RR-2 schedule).

Following the operant consequences described above, the three experimental groups received KR feedback delayed by 200 ms with probabilities identified by the name of each group: 0% KR, 67% KR, or 100% KR. As **Figure 1** illustrates, the qualitative probabilistic KR feedback was different for correct trials and incorrect trials. Correct sequences produced KR in which both front white panel lights turned on for 2 s, while the Sonalert sounded four times in that 2-s interval (each 0.25-s on, 0.25-s off). Following incorrect sequences, the Sonalert sounded for a full second, and the front panel lights remained off.

Experiment: Autonomy Phase

The 10-session Autonomy Phase was identical to the Acquisition Phase except: (a) The blue and green LED panel lights were replaced by identical white LED lights, and (b) No KR feedback was provided in any trial to any group. The lamps functioned the same way within trials as before (e.g., during ITI and TO), but the white lights above both levers provided no discriminative cues to influence lever selection. The autonomy condition assessed how well each rat could complete the L–R sequence independently, without the differential cues provided by panel lights or potential influence of KR feedback.

RESULTS

Figure 2 depicts L-R sequence accuracy for the three KR groups across the Acquisition and Autonomy Phases, separated by the vertical dotted line. We adopted an alpha level of.05 for all statistical tests.

Acquisition Phase

We used a mixed-effects 2-way ANOVA across the 30 sessions of the Acquisition Phase to assess two potential main effects (group, sessions) and a potential group × sessions interaction. We observed a significant repeated-measures effect of sessions, $F(29, 667) = 20.105$, $p < 0.001$, $\eta_p^2 = 0.446$, confirming that accuracy generally increased across acquisition sessions. We also observed a statistically significant group × sessions interaction, indicating that some groups learned faster or slower than average, $F_{(58,\ 667)} = 1.531$, $p = 0.008$, $\eta_p^2 = 0.118$. This significant interaction should be expected to qualify any main effect of group. Even so, the ANOVA showed a nearly significant main effect of group, $F_{(2,\ 23)} = 3.369$, $p = 0.052$, $\eta_p^2 = 0.227$. An LSD *post hoc* test demonstrated that accuracy for the 67% KR group increased faster than for the 100% KR group, $p = 0.016$. There were no significant between-group differences during the Acquisition Phase between the 0% KR group and the 67% KR group, nor between the 0% KR group and the 100% KR group.

Autonomy Phase

Figure 2 illustrates continued increasing accuracy through autonomy sessions. A mixed-effects 2-way ANOVA demonstrated a main effect of sessions in the Autonomy Phase, Wilks Lambda = 0.262, $F(9, 15) = 4.698$, $p = 0.004$, $\eta_p^2 = 0.738$, observed power = 0.964. The sessions × group interaction was not statistically significant, indicating the groups learned similarly during the Autonomy Phase, Wilks Lambda = 0.442, $F_{(18,30)} = 0.839$, $p = 0.646$. The main effect of group was statistically significant, $F(2, 23) = 7.112$, $p = 0.004$, $\eta_p^2 = 0.382$, observed power = 0.895. The Autonomy Phase shows that L-R accuracy for the 67% KR group was significantly greater than the 100% KR group, as confirmed with the LSD *post hoc* test, $p = 0.001$. Similarly, the LSD *post hoc* test indicated that the 67% KR group achieved significantly higher accuracy during the Autonomy Phase than the 0% KR group, $p = 0.033$. The 0% KR group appears slightly higher than the 100% KR group, but the difference was not statistically significant, $p = 0.19$.

Analysis of Errors

Figure 3 identifies the frequencies of response-sequence errors (L-L, R-L, R-R) observed in the Acquisition and Autonomy Phases for each KR group. The patterns of these errors were remarkably consistent across the three groups.

DISCUSSION

The purpose of this study was to begin answering the question: Does KR affect rats the same way as with humans? Many studies have demonstrated the counterintuitive feature of human

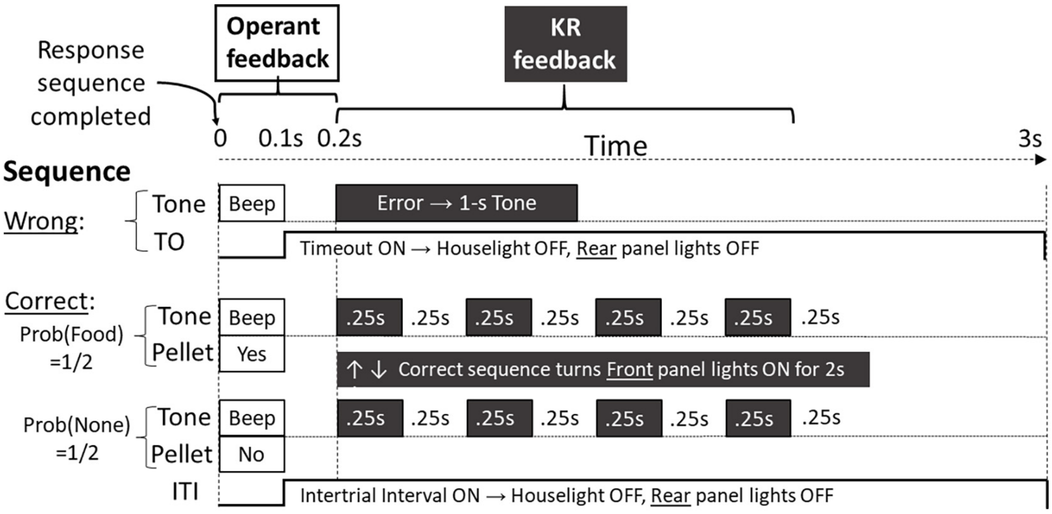


FIGURE 1 | Timing diagram for operant feedback and KR feedback, depicted in different fonts. Operant feedback is depicted in black text with white background, whereas KR feedback is depicted in white text with black background. See procedure for other details.

skill learning that practice with reduced relative frequency of KR improves skill retention over 100% KR, even though performance during acquisition often suffers (e.g., Wulf and Schmidt, 1989; Winstein and Schmidt, 1990; Schmidt and Bjork, 1992; Kantak and Winstein, 2012; Bjork and Bjork, 2014; Soderstrom and Bjork, 2015). Our rats demonstrated the same clear effect. Accuracy in the Autonomy Phase for the 67% KR group was greater than the 100% KR group: significantly greater accuracy with reduced KR feedback. The rats replicated the counterintuitive feature of skill learning that has come to define how motor skill learning occurs in humans. Yet performance during our Acquisition Phase did not suffer: Our 67% KR group yielded higher accuracy than the 100% KR group throughout both phases. This difference between rats and humans is surely due to the procedural differences in the number of trials experienced during acquisition. Human studies typically provide only about 20–30 acquisition trials (total), whereas our rats received an average of 3615 trials—more than 100 times as many (about 120 trials per session for 30 sessions). Performance in human KR studies improves for every group during acquisition phases (not only in the brief retention phase), consistent with the power law of practice. Therefore, one should expect improvement to continue when the duration of the Acquisition Phase is greatly extended, as in this experiment.

To our knowledge, this is the first experiment with rats that explicitly combined operant and KR feedback. By including a 0% KR group that never received any KR feedback, our procedure allowed us to measure whether the addition of KR feedback influences accuracy beyond that of operant feedback alone. **Figure 2** shows that the 67% KR group produced significantly higher accuracy than the 0% KR group during the Autonomy Phase. Thus, providing 67% KR feedback did improve autonomy above that of operant feedback alone.

KR experiments with humans have repeatedly demonstrated that acquisition is improved by providing some ("but not too much") KR feedback. However, they do not normally provide a 0% KR baseline condition for comparison. We observed that providing 67% KR was not "too much" compared to 0% KR because 67% KR was still beneficial. The differences between 100% KR and 0% KR were not statistically significant, but the differences were compatible with the claim that 100% KR was "too much" feedback to be beneficial for rats. An understanding of what entails "too much" feedback will require additional research. Nevertheless, our tentative observation is compatible with several hypotheses designed to explain how KR on every trial is less effective than less frequent KR, such as

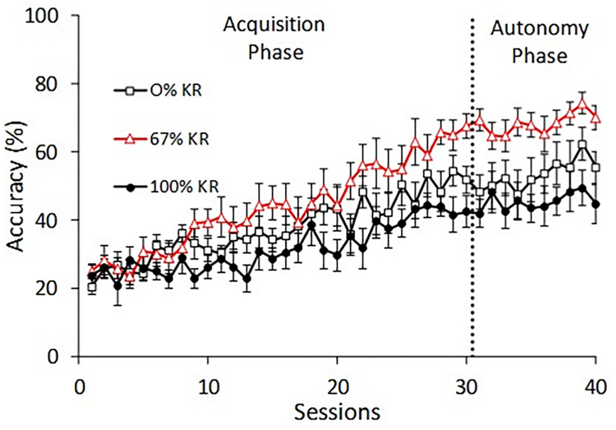


FIGURE 2 | Percentage accuracy of the L-R behavioral skill for Groups 0% KR, 67% KR, and 100% KR across the 30 sessions of the Acquisition Phase and 10 sessions of the Autonomy Phase. Error bars depict standard error of the mean.

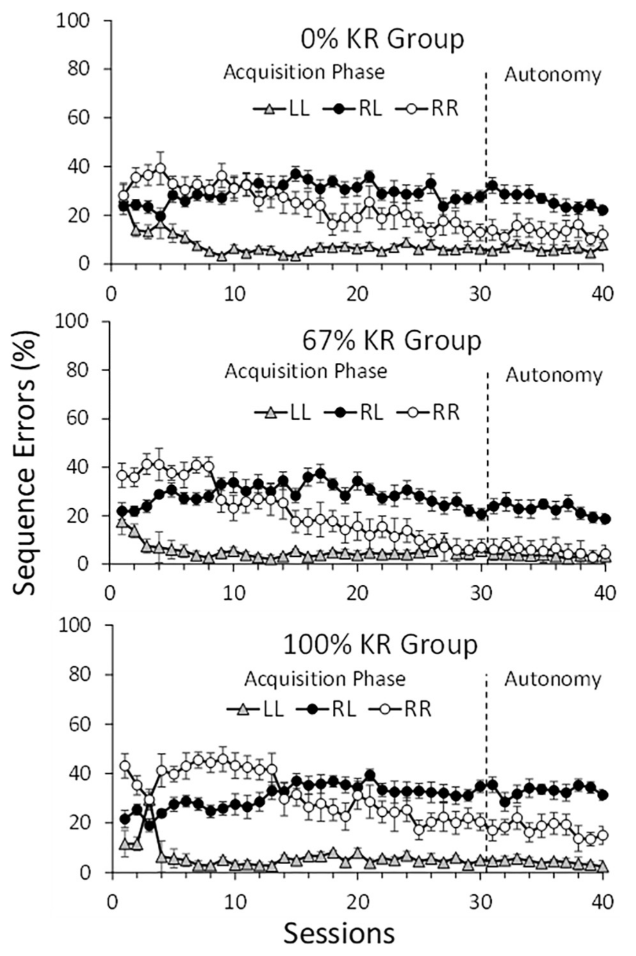


FIGURE 3 | Percentages of the three possible response-sequence errors (LL, RL, RR) depicted for the three KR groups across the Acquisition and Autonomy Phases. Note that the patterns of errors were similar for the three KR groups across both phases. Error bars depict standard error of the mean.

the guidance hypothesis (Salmoni et al., 1984) and explanations based on information processing. Of several theories attempting to explain this observation, Schmidt and Bjork (1992) suggested that frequent feedback in studies (with humans) could block important information processing activities from occurring during the acquisition phase that are required for learning. In contrast, when KR is given intermittently during the acquisition phase, human learners may be more able to evaluate performance in the absence of KR, and thus perform better in the retention phase (Winstein and Schmidt, 1990).

Schmidt and Bjork (1992) summarized procedures that share a common principle for learning motor and verbal skills in humans: "Introducing difficulties for the learner can enhance training" (p. 209). Providing reduced KR was one of these procedures, but others did not include KR. Bjork and Bjork (2014) recently described this principle as "Creating desirable difficulties to enhance learning" (p. 56). This principle also applies to acquisition of behavioral skills in rats and pigeons. As mentioned earlier, several rat and pigeon studies have observed the same counterintuitive feature of human skill learning: *Less effective* cues and *more difficult* behavioral skills degrade accuracy during acquisition phases, yet *enhance* accuracy in the subsequent autonomy phase (e.g., *Rats*: Reid et al., 2010, 2013a,b, 2017; *Pigeons*: Fox et al., 2014; Reid et al., 2014).

It is becoming increasingly apparent that skill learning in rats and humans share certain consistent features—even counterintuitive features. Some researchers may assume rats would not have the cognitive abilities to benefit from KR feedback (e.g., Wulf and Schmidt, 1989; Wulf and Shea, 2002) or that the mechanisms for learning in humans are not the same as those for animal conditioning with reinforcement (e.g., Winstein and Schmidt, 1990). Nevertheless, our rats replicated the major findings of KR procedures widely documented in human skill learning. The cognitive processes required for these procedures may be simpler than those proposed to explain results with humans. Identifying whether these similarities are due to common general principles of skill learning will require much more research, including explorations of the various forms of KR known to affect skill learning in humans (Winstein, 1991; Mazur, 2006). We believe this research will be worth the effort. We propose that experimental designs with animals (e.g., rats, pigeons, dogs) with clearly specified combinations of operant and KR feedback (e.g., **Figure 1**) have the potential to more clearly define the relationship between (and the importance of) these two classes of feedback that greatly influence both acquisition and retention in humans and laboratory animals. This discovery would be a substantial improvement in our understanding of skill acquisition across species.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AR designed the experiment, obtained funding, programmed all conditions, helped analyse the data, and wrote the final manuscript. PS supervised the research team to carry out the experiment, helped with data analysis, and contributed to early versions of the manuscript.

FUNDING

We are grateful to Wofford College for funding this research.

ACKNOWLEDGMENTS

We thank students enrolled in Learning & Adaptive Behavior at Wofford College for helping with the subjects in this experiment, and we thank Elizabeth Kyonka and James Mazur for fruitful discussions about this research. We are especially grateful to Wofford College for funding this research. Preliminary versions of this work were presented at the annual meeting of the XXX International Conference of the Spanish Society for Comparative Psychology, Ávila, Spain in 2018 and at the annual meeting of the Association for Behavior Analysis International in 2019.

REFERENCES

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychol. Bull.* 101, 41–74. doi: 10.1037/0033-2909.101.1.41

Bjork, E. L., and Bjork, R. A. (2014). "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, 2nd Edn, eds M. A. Gernsbacher, and J. Pomerantz (New York, NY: Worth), 59–68.

Bjork, R. A., and Bjork, E. L. (1992). "A new theory of disuse and an old theory of stimulus fluctuation," in *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes*, Vol. 2, eds A. Healy, S. Kosslyn, and R. Shiffrin (Hillsdale, NJ: Erlbaum), 35–67.

Fox, A. E., Reid, A. K., and Kyonka, E. G. E. (2014). Behavioral history and pigeons' "guiding cues" performance. *Psychol. Record.* 64, 403–413. doi: 10.1007/s40732-014-0060-9

Helton, W. S. (2007a). Deliberate practice in dogs: a canine model of expertise. *J. Gen. Psychol.* 134, 247–257. doi: 10.3200/genp.134.2.247-258

Helton, W. S. (2007b). Skill in expert dogs. *J. Exp. Psychol. Appl.* 13, 171–178. doi: 10.1037/1076-898x.13.3.171

Jacobs, G. H., Fenwick, J. A., and Williams, G. A. (2001). Cone-based vision of rats for ultraviolet and visible lights. *J. Exp. Biol.* 204, 2439–2446.

Johnson, A., and Proctor, R. W. (2017). *Skill Acquisition & Training: Achieving Expertise in Simple and Complex Tasks*. New York, NY: Routledge.

Kantak, S. S., and Winstein, C. J. (2012). Learning-performance distinction and memory processes for motor skills: a focused review and perspective. *Behav. Brain Res.* 228, 219–231. doi: 10.1016/j.bbr.2011.11.028

Mackintosh, N. J. (1965). Overtraining, transfer to proprioceptive control, and position reversal. *Q. J. Exp. Psychol.* 17, 26–36. doi: 10.1080/17470216508416405

Mackintosh, N. J. (1974). *The Psychology of Animal Learning*. London: Academic Press.

Macphail's, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Phil. Trans. R. Soc. Lond. B* 308, 37–51. doi: 10.1098/rstb.1985.0008

Mazur, J. E. (2006). *Learning and Behavior*, 6th Edn. Upper Saddle River, NJ: Pearson Prentice Hall.

Newell, K. M. (1991). Motor skill acquisition. *Annu. Rev. Psychol.* 42, 213–237.

Reid, A. K., DeMarco, G., Smith, K., Fort, T., and Cousins, E. (2013a). The influences of guiding cues on motor skill autonomy in rats. *Learn. Behav.* 41, 455–463. doi: 10.3758/s13420-013-0121-y

Reid, A. K., Rapport, H. F., and Le, T.-A. (2013b). Why don't guiding cues always guide in behavior chains? *Learn. Behav.* 41, 402–413. doi: 10.3758/s13420-013-0115-9

Reid, A. K., Folks, N., and Hardy, J. (2014). On the dynamics of stimulus control during guided skill learning in nonhumans. *Behav. Process.* 104, 72–83. doi: 10.1016/j.beproc.2014.01.017

Reid, A. K., Futch, S. E., Ball, K. M., Knight, A. G., and Tucker, M. (2017). Assessment of progressively delayed prompts on guided skill learning in rats. *Learn. Behav.* 45, 62–75. doi: 10.3758/s13420-016-0241-2

Reid, A. K., Nill, C. A., and Getz, B. R. (2010). Changes in stimulus control during guided skill learning in rats. *Behav. Process.* 84, 511–515. doi: 10.1016/j.beproc.2010.01.001

Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of results and motor learning: a review and critical reappraisal. *Psychol. Bull.* 95, 355–386. doi: 10.1037/0033-2909.95.3.355

Schmidt, R. A., and Bjork, R. A. (1992). New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychol. Sci.* 3, 207–217. doi: 10.1177/1745691617690642

Soderstrom, N. C., and Bjork, R. A. (2015). Learning versus performance: an integrative review. *Perspect. Psychol. Sci.* 10, 176–199. doi: 10.1177/1745691615569000

Thorndike, E. L. (1927). The law of effect. *Am. J. Psychol.* 39, 212–222.

Winstein, C. J. (1991). Knowledge of results and motor learning—Implications for physical therapy. *Phys. Ther.* 7, 140–148.

Winstein, C. J., and Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 677–691. doi: 10.1037/0278-7393.16.4.677

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nat. Rev. Neurosci.* 12, 739–751. doi: 10.1038/nrn3112

Wulf, G., and Schmidt, R. A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *J. Exp. Psychol. learn. Mem. Cogn.* 15, 748–757. doi: 10.1037/0278-7393.15.4.748

Wulf, G., and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon. Bull. Rev.* 9, 185–211. doi: 10.3758/bf03196276

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

REVIEW
published: 19 May 2020
doi: 10.3389/fpsyg.2020.00973

The Comparative Psychology of Intelligence: Some Thirty Years Later

*Irene M. Pepperberg**

Department of Psychology, Harvard University, Cambridge, MA, United States

After re-reading Macphail's (1987) essay "The Comparative Psychology of Intelligence" with all the associated commentaries, I was struck by how contemporary many of the arguments and counter-arguments still appear. Of course, we now know much more about the abilities of many more species (including their neurobiology) and fewer researchers currently favor explanations of behavior based solely on associative processes; however, the role of contextual variables in comparative psychology still remains cloudy. I discuss these issues briefly. Given my research interests involving the cognitive and communicative abilities of Grey parrots, the one aspect of the original article upon which I feel I can comment in depth involves Macphail's claims about the importance of language—and specifically syntax—in problem-solving and thus in placing humans above all other creatures. Granted, no other species has (or in my opinion is likely ever to acquire) everything that goes into what is considered "human language." Nevertheless, several other species have acquired symbolic representation, and considerable information now exists upon which to base an argument that such acquisition by itself enables more complex and "human-like" cognitive processes. Such processes may form the basis of the kind of intelligence that is measured—not surprisingly—with human-based tasks, including the use of such representations as a means to directly query non-human subjects in ways not unlike those used with young children.

Keywords: comparative psychology, avian cognition, animal intelligence, grey parrot, animal cognition

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Martin Giurfa,
UMR5 169 Centre de Recherches sur la Cognition Animale (CRCA), France
Bonnie M. Perdue,
Agnes Scott College, United States

Correspondence:
Irene M. Pepperberg
impepper@media.mit.edu;
impepper@wjh.harvard.edu

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *31 January 2020*
Accepted: *20 April 2020*
Published: *19 May 2020*

Citation:
Pepperberg IM (2020) The Comparative Psychology of Intelligence: Some Thirty Years Later. Front. Psychol. 11:973. doi: 10.3389/fpsyg.2020.00973

INTRODUCTION

Over 30 years have passed since Macphail's essay "The Comparative Psychology of Intelligence" was published in *Behavioral and Brain Sciences* along with numerous commentaries, critiques, and his rebuttals (Macphail, 1987). With some exceptions that I will not discuss below (e.g., the most notable being that recent decades have seen an unprecedented upsurge in both the reputation and number of publications in comparative psychology, in contrast to the adverse trends upon which Macphail commented in 1987—e.g., the emergence of a new journal, *Animal Cognition*; the independence of the *Journal of Comparative Psychology*; the founding of the *Comparative Cognition Society;* the publication of Call et al.'s 2017 *APA Handbook of Comparative Psychology*), many of the same arguments and counter-arguments might be found in contemporary literature. Granted, we now know much more about the neurobiology of many more species (particularly the fairly recent findings about the complexity of psittacine and corvid brains) and have reams of data about sophisticated abilities of previously unstudied or rarely studied creatures such as reptiles and even invertebrates. Too, a larger number of researchers are now less likely to reduce the characteristics of the tasks being studied to chains of associative processes, though some still disagree. The extent to which the effects of specific capacities that have been shaped by evolutionary pressures can be defined or explained by contextual variables remains cloudy; for example, some researchers propose that many living creatures begin life with certain equivalent core abilities, upon which more complex cognitive capacities may be built to varying degrees. I'll briefly discuss a few of these topics, then concentrate on what for me is a central issue—that of the effects of the acquisition of symbolic communication, albeit something much less than language, on the cognitive capacities of non-human subjects, and how such communication may expedite the study of such capacities. Almost all my comments will arise from the standpoint of a researcher on such avian capacities.

IS EVERYTHING REDUCIBLE TO ASSOCIATIONS?

This basic claim of Macphail is controversial, specifically because the answer to the question depends solely on one's theoretical framework. Some researchers still argue that all complex tasks can be reduced to a series of associations and others heatedly disagree—see, for example, Heyes (2016) associative-learning based arguments for the explanation of human imitation, other researchers' counter-claims for neonatal imitation (Meltzoff and Moore, 1999; Simpson et al., 2014), and the demands of still others for further research (Vincini et al., 2017). Despite an overall lack of consensus on exactly what does separate associative versus other forms of learning, many researchers seem to agree on mental representation, rule-based learning and symbolic processing as behaviors that differ in measurable ways from associative learning (e.g., McLaren et al., 2018; Church, 2019; Smith, 2019; Wills et al., 2019). Note that Macphail acknowledges only the uniqueness of language, of which symbolic processing is merely one aspect. The debate framed by Macphail thus clearly has not been resolved, but its parameters have widened considerably, with many cogent arguments for multiple levels of processing. I leave it to colleagues who specialize in areas outside of symbolic processing to address those aspects of the debate in full.

WHAT IS THE ROLE OF NEUROANATOMY?

As with the debate on associative learning, my knowledge of neuroanatomy is limited compared to that of others who will also likely be commenting on Macphail's paper. However, I wish to note, if only briefly, that the explosion of information on what is now known about non-human brains cannot be ignored, particularly with respect to avian cognition. Numerous papers have demonstrated that the architecture of neither a primate nor even a mammalian brain is required for complex cognitive processing (e.g., Iwaniuk et al., 2005; Jarvis et al., 2005; Güntürkün and Bugnyar, 2016; Gutiérrez-Ibáñez et al., 2018; Rinnert et al., 2019). Significantly, Olkowicz et al. (2016) found that parrots and corvids in particular have forebrain neuron counts equal to or greater than primates with much larger brains, and suggest that such avian neural densities are likely responsible for their high levels of intelligence. A recent study even argues for adult neurogenesis in the Grey parrot brain as being correlated with advanced cognitive processing (Mazengenya et al., 2018). Other studies demonstrate certain relationships (and thus suggest the possibility of shared forms of processing) among avian, mammalian, and reptilian brains based on common ancestry (Tosches et al., 2018). Additional research (see entries in the aforementioned *Handbook*) demonstrates many non-human abilities that compare favorably with those of humans. However, lest we are tempted to use this information to argue for the lack of differences between human-non-human abilities or among non-human species, we also know that even within related species—for example, among closely-related corvids and also among less-closely related parrots—there exist subtle and not-so-subtle brain variations, respectively, that likely are correlated with differences in the types and extent of processing abilities (see Basil et al., 1996; Gould et al., 2013; Chakraborty et al., 2015). Differences in the relative sizes of specific neural areas, the presence or absence of specific areas, and the overall internal organization will affect the complexity of a species' cognition and memory. Notably, even within the same species, individual differences exist with respect to competencies: we argue about human brilliance based on outliers (e.g., Einstein, Beethoven, Rembrandt, Shakespeare), but the 'average' human clearly does not exhibit such capacities, even when contextual variables are taken into account. Although we have not entirely determined the neural correlates of human intelligence (see Rhein et al., 2014; Wen et al., 2019), evidence for inter-individual differences obviously exists, both in brains and behavior. Surely interspecies differences can be at least as great as intraspecies ones?

SPECIES DIFFERENCES IN GENERAL?

At least one reptile fails a task in simple numerical cognition (although not in distinguishing larger from smaller objects; Petrazzini et al., 2017), and comparative work by researchers (e.g., Kamil and his students; see Olson, 1991; Olson et al., 1995) have shown that some bird species excel at certain spatial learning tasks and not on others such as match-to-sample. Some of these behavioral differences may be related to differences in brain structure (see previous section), but one might, like Macphail, argue that such differences are simply a matter of "contextual variables." Interestingly, some researchers now argue, consistent with Macphail's claims, that most species have very similar, basic "core" capacities, which are involved in representing certain aspects of objects, actions, number, space, and (possibly to a lesser extent), social interactions (Spelke and Kinzler, 2007)—that is, diverse species show remarkably similar levels of competence on a number of rather basic tasks. However, these so-called *core* capacities, which are present in most species at a very young age, are but the building blocks of complex cognitive processing. If all that are being studied are tasks that rely on these core capacities (e.g., more-less, object identity), then the few differences that emerge are, indeed, likely to depend on contextual variables. However, contra Macphail, additional research demonstrates that different species develop additional abilities, beyond those based on these core capacities, to different extents: if the tasks being studied require more than core capacities, differences exist in various abilities to process more and more complex information (Wright et al., 2018). Specifically, complex cognition is not only the ability to come to a decision by evaluating, or processing, current information on the basis of some representation of prior experiences (e.g., Kamil, 1984). Complex cognition must also include the capacity to choose, from among various possible sets of rules that have been acquired or have been taught, the set that appropriately governs the current processing of this data—that is, in order to solve a problem, the subject must first decide which rules are appropriate for the processing of data (i.e., determine which of many possible different types of problem is being posed) and then figure out what types of data are needed for the solution. According to this criterion, a subject that is limited to organizing information on the basis of a single set of rules—a subject that has little more than the core capacities that allow success at something like a matching procedure—will not have the occasion to demonstrate complex cognitive processing (Pepperberg, 1990).

How a subject develops more advanced capacities from core capacities, and exactly *why* different species acquire more advanced capacities than others in some domains and not in other domains—these questions still require complete answers. The answers likely lie in some confluence of evolutionary pressures in the form of environmental input and the wherewithal to process the information in this input, and thus in some differential aspect of neurobiology—but remain a subject for study at present.

THE ROLE OF LANGUAGE

Macphail argued that although no differences exist in the intelligence of the various non-human species, he made a special case for humans based on their acquisition of language, particularly syntax. The extent to which Macphail's arguments are still valid enter yet another murky realm. He argued that, in almost all cases, what passed for "language" in the various non-human programs he reviewed (e.g., Gardner and Gardner, 1969; Premack, 1971; Rumbaugh, 1977; Patterson, 1978; Terrace et al., 1979; Savage-Rumbaugh et al., 1983) were simple associations between objects and artificial symbols; he argued that these subjects' inability to create novel sentences meant that whatever success they had achieved was insufficient to raise their intellectual capacity. By ignoring some possible instances of novelty (e.g., Fouts and Rigby, 1977; Rumbaugh, 1977) and—most importantly—only briefly noting work on cetaceans (e.g., studies by Herman and Schusterman and their students), Macphail gave insufficient credit to these species' abilities to understand certain levels of rule-governed behavior. Herman's dolphins, for example, could respond with statistically significant accuracy to *novel* 5-element sentences such as "modifier + direct object + verb + modifier + indirect object" (e.g., fetch the right hoop—as opposed to the left one—and bring it to the top frisbee—as opposed to the bottom one; or could swim through both hoops in the given order; Herman et al., 1984; see Schusterman and Gisiner, 1988, for related work on sea lions).

To give Macphail credit, rule-governed behavior is only a simple form of syntax, and no non-human has demonstrated capacities fully comparable to all the possible intricacies of complex human communication. It must be noted, however, that not every human language includes all those intricacies—e.g., some lack complex embedding and constructs such as the passive (e.g., Everett, 2005; some controversy surrounds that claim, but Huttenlocher et al., 2010 has shown that when raised in impoverished settings, even American children's sentence structure lacks such complexities). Clearly, cetaceans in such training programs demonstrated far more complex behavior patterns than would be possible without such instruction. Notably, reducing their behavior to chains of association would require the same type of reduction for much of human language, a communication system that Macphail argued was unique.

Macphail also failed to fully appreciate what I had recently accomplished with a Grey parrot (Pepperberg, 1983), research that also contradicted claims of simple association rather than full referential abilities. The parrot, Alex, by showing that he could—even at that early stage in his training—vocally indicate different attributes of a single item (its color, shape, material, and overall label), demonstrated that, even for a novel item, he could interpret the various possible questions that could be posed (i.e., determine which attribute was being targeted), search his repertoire for the set of labels that were hierarchically organized under that attribute (e.g., if the question was "What color?," know to examine labels such as yellow, blue, green, etc. instead of paper, cork, wood, etc.), chose the one appropriate label, and then encode it vocally.

Of particular interest is that Premack (1983) had made a somewhat similar argument to that of Macphail concerning the effect of language on cognition but had come to a strikingly different conclusion. Premack (1983), sidestepping the controversy surrounding whether animals were capable of human-level language, claimed that non-humans who learned *symbolic representation*—in which a non-iconic symbol stands for an object, an attribute, an action, etc.—have an enhanced ability to perform tasks that require abstract thinking. He buttressed these claims with data demonstrating that those of his apes that had acquired such symbolic representation outperformed those that did not, particularly on tasks such as analogical reasoning. [Interestingly, although some evidence exists for such reasoning in subjects lacking symbolic representation (e.g., a relational MTS task, see data and discussion in Obozova et al., 2015), those claims have been critiqued by Vonk, 2015]. I discuss several experiments from my laboratory that give additional credence to Premack and suggest the limits of Macphail's argument. I do so with the following caveat: I have argued previously (summarized in Pepperberg, 1990) that training on symbolic representation (or its lack) is likely to affect only the ease with which animals learn and can be tested on certain concepts. I've also argued that although a system of two-way communication may enable a researcher to teach a concept that an animal subject may not easily acquire by other means, such acquisition is unlikely if the animal does not have the basic cognitive capacity for such acquisition. The first set of experiments I review relates to these points. However, I now am not entirely sure about a corollary of those two claims—that acquisition of symbolic representation does not affect how subjects manipulate information—I believe that, in some specific instances, such a change in processing ability may exist. Possibly, once a non-human understands that a symbol can be used to represent an object or an action, it can then understand how, for example, a three-dimensional entity can be represented by a two-dimensional one (e.g., optical illusions, Pepperberg and Nakayama, 2016) or that two symbols (e.g., one vocal and one visual) that separately represent the same object can then represent each other (a formal equivalence, Pepperberg, 2006b), or that symbols can be used as place-markers to assist in tasks requiring memory and evaluation of probability (Pepperberg and Pailian, 2017; Clements et al., 2018). I will discuss one such topic—various studies on numerical concepts—in depth. I will also suggest some limits to the functionality of symbols, based on the extent of their use and comprehension (Bowden et al., 2019).

When Symbolic Representation Affects Training/Testing and Therefore Results

I describe two studies in depth in which the use of symbolic processing plays an important role, not specifically because such representation allows for more complex processing, but because it facilitates training and/or eliminates confounds that can detract from the claims of success in other species. In those as well as the additional studies briefly referenced above, I have demonstrated that Grey parrots have succeeded on certain tasks that have proven challenging to other non-humans; I believe that the success of the birds may rely on their access to symbolic representation. For these and other tasks, exact comparisons between non-humans with and without symbolic representation cannot be made specifically because the parrot can choose one vocal response from its entire repertoire, much like children but unlike almost all other non-humans. These findings do not support Macphail's claims of a lack of difference in intelligence among species, but rather suggest that having even some language-like elements may be instrumental in assisting researchers to explore these differences.

A Study on Abstract Relations: Bigger/Smaller

The ability to predicate a response on a relational rather than an abstract basis is frequently used as a metric for comparing cross-species abilities, because understanding relations (darker than, bigger than, etc.) is supposedly a more complex task than learning to respond to an absolute concept (e.g., redness; see discussions in Schusterman and Krieger, 1986; Pepperberg and Brezinsky, 1991). Responding on a relative basis requires a subject to compare stimulus choices and then derive and use an underlying, more abstract (and thus general) concept; it is the comparison that is crucial, because in a task such as "lighter than" the right answer in one trial ("gray" in a task pitting black against gray) may be the incorrect in the next trial (pitting white against gray). In contrast, learning an absolute stimulus value requires only that a subject form a single association (e.g., choose gray; Thomas, 1980). Because tasks that involve relative concepts often allow organisms to learn something about both absolute and relative concepts concurrently (Premack, 1978), researchers who use such tasks for cross-species evaluations of cognitive capacity must determine the extent to which their subjects rely on relative information in problem-solving. I'll present a few examples without going into a detailed review of the literature.

It is not that subjects unable to use symbols completely fail in demonstrating relative concepts, but rather that these subjects tend to focus primarily on absolute concepts and that demonstrating their understanding of relational concepts can be challenging. In one set of such studies, starlings that are taught to discriminate a set of rising tones from a set of descending tones and are then asked to transpose to a novel set in a totally different key, can transfer solely under very specific conditions. Such data show that they respond on a relative basis as a secondary strategy, only after acquiring information on an absolute basis (Hulse et al., 1984, 1990; Cynx et al., 1986; Page et al., 1989; MacDougal-Shackleton and Hulse, 1995). This so-called "frequency range constraint" may derive from ethological priorities, where changing the overall pitch changes the meaning and importance of the signal (see discussion in Pepperberg, 1999).

Other studies have examined relative luminance or size. Here, pigeons also tend to discriminate on the basis of absolute, rather than relative, brightness, and rhesus monkeys do the same for both brightness and size, although changes in experimental design (e.g., how stimuli are presented) may result in their showing some understanding of the relative concept in transposition trials (see Pasnak and Kurtz, 1987; Wills and Mackintosh, 1999). Interestingly, horses seem capable of size transpositions, although—like almost all the other subjects tested—only for the direction in which they were trained—that is, to stimuli that are either relatively larger or relatively smaller, but not to both within the same experiment (Hanggi, 2003). And, of course, responding to "larger" is only meaningful if a subject can also respond to "smaller," so that the task is not simply the ability to respond to "more" (possibly a preference related to foraging).

Two studies with symbol-using subjects, on relative size, however, demonstrate the worth of symbolic representation. Here, the subject is taught a label for both of the concepts that are being tested, rather than having to derive the concepts over large numbers of trials. In one study on a sea lion, Schusterman and Krieger (1986) demonstrated that their subject understood the concepts of both bigger and smaller and transposed to objects of novel sizes, such that the previously correct choice would now be incorrect; their subject, however, was not tested on items completely different in shape or material from those used in training. My Grey parrot, Alex, after learning to respond to "What color bigger/smaller?" for three sets of items, was able to transfer, without additional training, to a large number of sets involving sizes outside the training paradigm and to totally novel objects with respect to shape, color, and material; he also spontaneously transferred to the questions "What matter bigger/smaller?" and, when the two objects were of the same size, spontaneously responded "none," transferring his understanding of that label from a study on a lack of same/difference (Pepperberg, 1987a, 1988; Pepperberg and Brezinsky, 1991). That these subjects had the ability not only to respond to the largest *or* the smallest item that was present, but also could recognize that on *any* trial, *either* bigger or smaller could be queried, demonstrated a far greater understanding of the relative concept than had been shown by any other non-humans. Their training in symbolic representation was not likely responsible for this understanding but enabled them to display such capacities at a higher level.

The Müller-Lyer Illusion

Few studies that examine how non-humans perceive optical illusions are directly comparable to those with humans. Grey parrots that have some referential use of English speech, however, allow for such comparative studies, as these birds can be tested just as are humans, by asking them to describe exactly what they have seen. Specifically, no studies had previously been performed on an avian subject that, *without any training on the actual task,* could—as would be possible for Alex—simply state vocally whether or not an optical illusion had been observed. My colleagues and I began a series of such studies by examining the Müller-Lyer illusion (Pepperberg et al., 2008) because it is well-represented in the scientific literature; in the classic form, humans underestimate or overestimate the length of a line that has arrows attached, respectively, either inwardly < > or outwardly > <. Many explanations exist as to why humans are subject to the illusion (see review in Pepperberg et al., 2008), but our main interest was in determining how it would be processed by the avian visual system, which is notably anatomically and neurobiologically distinct from that of humans (see review in Shimizu et al., 2010 for both similarities and differences). Would a parrot, separated from humans by 300 million years of evolution (Hedges et al., 1996) also be duped into thinking that the two horizontal lines in the illusion differed in length because of the placement of the arrows?

Some evidence existed for the illusion in ring doves (Warden and Baar, 1929), pigeons (Nakamura et al., 2006), and chickens (Winslow, 1933), but the data were not conclusive. Intensive training procedures were generally necessary to enable these birds to discriminate the initial stimulus and subjects were then tested on their recognition of similar patterns. Results often depended on, for example, statistical averaging over 100s of trials of pecking/touching behavior to a very limited set of choices and thus were often highly variable and dependent upon details of the experimental design (reviewed Pepperberg et al., 2008). Rosa Salva et al. (2014) clearly discuss how intensive training and the requirement that subjects perform such physically manipulative responses result in failures that may be avoided when subjects engage in incidental learning and can respond in a more naturalistic manner. Thus, having a subject that, like humans, would have an extensive vocal repertoire from which it could choose any utterance (from over 100 possibilities), and could simply be asked to describe what it sees, on only a few trials per type of stimulus, without any prior training on any materials related in any way to those stimuli, would avoid these issues. Also, because such vocal responses and lack of extensive training did enable us to query Alex only a relatively few times compared to other non-humans, we could prevent, at least to some extent, a decrement in perceiving the illusion that can occur over time (Mountjoy, 1958; Predebon, 1998, 2006).

Alex was shown the Brenano version of the task (>–<–>), considered to be equivalent to the presentation of two separate figures (e.g., Sadza and de Weert, l984), to ensure that he focused on both illusions simultaneously; the horizontal lines were of different colors and he was queried as to "What color bigger/smaller?," a concept he already understood (Pepperberg and Brezinsky, 1991). Controls were lines with the arrows replaced by perpendicular lines; if he saw the illusion as do humans, he would give a color response on the standard queries and say "none" (his response to the absence of a size differential) in the controls. Requiring responses with respect to both bigger and smaller forced Alex to attend to and interpret each question individually, unlike most other non-human subjects. To test the extent to which he saw the illusion, we varied the pitch of the arrow from the standard 45° and the thickness of the horizontal lines. Again, if Alex responded as did humans, thick horizontal lines would decrease the extent to which he saw the illusion, as would angles that approached 90°.

Alex's data were scored as 'illusion reported' if he named the shaft color that human observers would report, as 'no illusion' if he reported "none," and 'opposite the illusion' if he reported the color opposite to the illusion response. After accounting for mistrials due to inattention, he reported the illusion in about 88% of the trials in which human observers would have reported the classic illusion (such a rate is consistent with his overall accuracy in color labeling, 80–85%; Pepperberg, 1999) and he showed a lessened or absent illusion in control trials where humans would not have reported the illusion, as the line thickness increased and the arrow angles altered (Pepperberg et al., 2008). Interestingly, even with the relatively limited number of trials we administered, we needed to account for some habituation and decrement of response due to inattention; such findings suggest how the effects of extended training and testing may have affected the results of previous studies.

His data suggest that even if avian systems for visual input differ from those of mammals, similar processing may occur within various neural structures. The importance of the data, however, lie not only in finding out how a Grey parrot perceives the world in which we co-exist, but also in being able to compare his responses *directly* to humans who view the same stimuli. This study emphasizes ways in which symbolic representation affects *how* non-humans can be tested on certain concepts. Thus, although one may argue that this study involves perception more than intelligence, my point is that the results show that symbolic representation enables us to directly compare how perceived information is processed, which is part and parcel of intelligence.

Other Studies

Alex was tested on many other types of tasks, including the concept of same-different (Pepperberg, 1987a). A review of that entire topic is the basis for a separate paper (Pepperberg, in review), but the central issue is as follows. Same-different is more than identity versus non-identity or the difference in entropy between stimuli (e.g., Young and Wasserman, 2001). Rather, it is a task that, according to Premack's (1983) stringent criteria, requires a feature analysis of the objects being compared, recognition that objects can simultaneously exhibit attributes that involve *both* similarity and difference, and the ability to understand which attributes are being targeted based on questions of either similarity *or* difference. Because an appropriate response requires that a subject (a) attend to multiple aspects of two different objects; (b) determine, from a verbal question, whether the response is to be based on sameness *or* difference; (c) determine, from the exemplars, *exactly* what is same or different (i.e., what are their colors/shapes/materials?); and then (d) produce, verbally, the label for the hierarchical category of the appropriate attribute, the task is another instance in which symbolic representation is likely critical for success (Premack, 1983).

Furthermore, Alex was not the only bird that my students and I have studied; numerous experiments on another Grey parrot, Griffin—one with a less extensive repertoire but with otherwise similar experiences to those of Alex—provide further evidence for the importance of symbolic representation in the study of non-human intelligence. Detailed descriptions of such studies with respect to the importance of symbolic representation are reviewed elsewhere (Pepperberg, in press) but, as noted above, his demonstration of capacities comparable to that of a 7-year-old child on topics such as Piagetian probability (Clements et al., 2018) likely involved his understanding that symbols can be used as place-markers to assist in tasks requiring memory and evaluation of chance. His ability to label occluded objects correctly and recognize Kanizsa figures also likely depended upon his symbolic understanding, in those instances for transferring his knowledge that a three-dimensional entity can be represented by vocal label into knowledge that the same entity can also be represented by a two-dimensional depiction (Pepperberg and Nakayama, 2016).

When Symbolic Representation May Enable Advanced Information Processing

For the topics discussed and referenced above, non-humans without symbolic representation were often able to demonstrate certain levels of competence, but non-humans with such representation were able to demonstrate either somewhat higher levels of such competence or were able to demonstrate their competence simply more efficiently. For the topic discussed below, the data suggest that the capacity for symbolic representation may actually have affected whether the non-human subject actually *could* demonstrate the given capacity; non-humans lacking the levels of Alex's representation have not, at least at present, shown such levels of intelligence.

Exact Numbers, Including a 'Zero-Like' Concept

Numerical competence can be defined so as to include a wide range of abilities, ranging from a simplistic understanding of more-versus-less to full comprehension of various forms of set theory. What makes numerical competence interesting as an overall topic is that number is not an inherent attribute of an object, as is color, shape, or material, but rather a descriptor that is applicable to any discrete collection of entities. What makes numerical competence relevant to the theme of this paper, however, is its relationship to symbolic representation, and the argument that a full understanding of number begins with the ability to use symbols to designate exact quantities (see Wiese, 2003).

It is, of course, true that some basic understanding of number is a widespread phenomenon; use of a primitive, approximate number system (ANS) has been observed in almost every species examined, from fish (Petrazzini et al., 2015) to bears (Vonk and Beran, 2012), from preverbal children (Wynn, 1990) to preliterate hunter–gatherers (Frank et al., 2008) (but note reptile exception above, Petrazzini et al., 2017). Although the ANS allows for some level of numerical discrimination, ANS tasks are not symbol-based and precision under the ANS decreases sharply in all mathematical operations as the number of items involved increases; for example, accurate comparison of two numerical sets is possible only when they differ by a sufficient ratio (Halberda et al., 2008). Consequently, species possessing an ANS can generally choose the greater of sets consisting of one, two and three items exactly, but their accuracy decreases when larger numbers and smaller ratios are involved and the ANS is not useful when discrimination among sets of even moderately larger quantities (e.g., eight versus nine) is required to solve a problem or to achieve success on a task.

In contrast, symbolic representation of number—the understanding that individual symbols represent exact, specific quantities—enables advanced capacities such as counting principles, precise addition, subtraction, etc. Acquisition of symbolic representation of number is a slow, multi-year process, even for human children (Fuson, 1988; Carey, 2009), and was once thought limited only to humans (reviewed in Pepperberg and Carey, 2012). Notably, Hurford (1987) and Dehaene (1992) have also suggested a close correlation between labeling and number skills, in the sense that numerical cognition is "a layered modular architecture, the preverbal representation of approximate numerical magnitudes supporting the progressive emergence of language-dependent abilities such as verbal counting" (Dehaene, 1992, p. 35).

Indeed, only a very few non-humans have acquired exact symbolic number representation: two apes, Matsuzawa's Ai (Matsuzawa, 1985) and Boysen's Sheba (Boysen and Berntson, 1989), and my subject, the Grey parrot Alex (Pepperberg, 1987b, 1994). Sheba's instruction on symbolic representation primarily involved use of physical Arabic numerals; in contrast, at the time that their numerical training began, Ai and Alex had already been trained to identify objects and colors and Alex had also begun to recognize shapes based on their numbers of corners (Asano et al., 1982; Pepperberg, 1983). Thus, these two latter subjects had to reorganize how they categorized objects in their world. They had to learn that a new set of labels, either physical symbols or the vocal labels "one," "two," "three," and so forth represented a novel classification strategy; that is, one based on both physical similarity within a group (e.g., that the objects were, for example, all keys) and a group's quantity (the exact number of a set), rather than solely by physical characteristics of group members (being metal). They also had to generalize this new class of number labels to sets of novel items and items in random arrays; Alex, unlike other subjects, also had to extend his understanding to heterogeneous collections. All three subjects eventually expanded their competency to more advanced numerical processes (Pepperberg, 1999, 2006a). And, as we shall see, they all understood, at least to some extent, that numbers are flexible tools that can be used to assess both cardinal and ordinal relations (Wiese, 2003). Such behavior is not easily reducible to simple associative learning; however, these competencies were acquired by subjects who had not demonstrated communication skills comparable to human language.

All three of these subjects acquired the ability to identify sets of objects exactly (i.e., their accuracy did not decrease as the size of the set increased as in the case of the ANS). Initial studies showed that Sheba and Alex could quantify sets up to six (Pepperberg, 1987b; Boysen, 1993), and that Ai could distinguish sets up to nine, although for the largest quantity she seemed to use a fairly accurate form of estimation rather than counting (Biro and Matsuzawa, 2001a). All three subjects were equally accurate when asked to examine novel sets and sets placed in random arrays. Such behavior is not possible without the use of symbolic representation. Interestingly, when asked to distinguish between numerical sets, data from monkeys without such training obeyed ANS rules (Brannon and Merritt, 2011), whereas those with symbolic representation were considerably more accurate, particularly when the two sets were large and differed by one unit (Livingstone et al., 2010).

The two apes and the parrot also acquired a zero-like concept. Ai and Sheba were specifically trained on the concept (Boysen and Berntson, 1989; Biro and Matsuzawa, 2001b). Alex, in contrast, spontaneously transferred his use of "none," which he had originally learned to produce so as to designate the absence of a common attribute in a same/different task (Pepperberg, 1988), to now designate the absence of a set of objects ("none" present, Pepperberg and Gordon, 2005). Apropos of the topic of this review, such a transfer was possible *only* because Alex could access his entire verbal repertoire during sessions—that is, was not limited to choosing among a small number of possible response keys—and, in this case, was also able to use his vocal abilities to manipulate the experimenter into asking him the question that led to his demonstrating this transfer (Pepperberg and Gordon, 2005). Again, he had received no training on use of any symbol, vocal or physical, to represent absence of quantity.

Alex, without training, was also able to quantify subsets in a heterogeneous array: given four groups of items that varied in two colors and two object categories (e.g., blue and red keys and trucks), he was able to label the number of items uniquely defined by the conjunction of one color and one object category (e.g., "How many blue key?") with an accuracy >80% (Pepperberg, 1994). Notably, the study replicated work with adult humans (Trick and Pylyshyn, 1989), who use an exact, rather than an ANS, in this task. Young children who, like Alex, had been taught to label homogeneous sets exclusively, may fail this task; they may be at the stage where they can label exact quantities, but often give the total number of items instead of that of the targeted subset (see Siegel, 1982; Greeno et al., 1984). Interestingly, *unlike the other subjects*, Alex was never trained on number comprehension; nevertheless, when tested, his comprehension accuracy was somewhat superior to that of production (Pepperberg and Gordon, 2005). Again, such abilities are based on symbolic representation.

Ordinality

Whereas the apes had been trained from the start of their studies to use Arabic numerals, and learned their quantifications in numerical order, much like children (see Carey, 2009), Alex had been trained only on vocal numerical labels, and, in contrast, first learned "three" and "four" (simultaneously), then "two" and "five" (again, simultaneously), and lastly "six" (pronounced "sih") and "one" (again, simultaneously). Note that, unlike the other subjects, he couldn't simply point to an answer. Instead, for each label, he had to learn to configure his vocal tract appropriately (see Patterson and Pepperberg, 1994, 1998), a somewhat difficult process (Pepperberg, 1999). After demonstrating his accuracy—as noted above, *without any training*—on comprehension (Pepperberg and Gordon, 2005), he was then taught to identify Arabic numerals (production and comprehension), using the same labels as he had used for numerical sets of objects, *but in the absence of any of these sets.* Thus, given a tray containing all the plastic or wooden Arabic numerals of different colors from 1 to 6, he learned to correctly respond to queries of, for example, "What color is 'four'?" or "What number is 'blue'?" Of particular interest is that he then *spontaneously* inferred the ordinality of his labels, as tested by his stating the color of the larger or smaller Arabic digit in a paired set or "none" if they represented the same quantity (Pepperberg, 2006c). These data also demonstrated that he was capable of a formal stimulus equivalence (Sidman et al., 1989), a behavior that is again dependent upon symbolic representation. Notably, ordinality did *not* arise spontaneously in the apes, even though their numerals had been learned in order (Boysen et al., 1993; Tomonaga et al., 1993); they all required significant amounts of training.

Given that both apes and Alex understood symbolic representation, why was Alex the only subject to spontaneously demonstrate ordinality? As noted several times already, his use of symbols would not have qualified as language under any definition. The issue is one that I will discuss after presenting additional evidence of Alex's numerical capacities.

Addition and More About Zero

Like Sheba (Boysen and Berntson, 1989), Alex also spontaneously demonstrated the ability to sum sets of objects and label that sum (Pepperberg, 2006b). Initially, Alex was presented with two cups placed on a tray, under which various quantities of objects were hidden from view. He was briefly (either 2–3 or 10–15 s, depending upon the experiment) shown each quantity, after which the cup covering that quantity was replaced. After both cups were replaced, he was asked about the total number of objects, which could vary from 0 to 6. For all but a few control trials, the objects varied in mass and contour, so that Alex had to respond on the basis of number. Under the shorter time constraints, his accuracy for all sets, with the exception of 5 + 0, was just below 90%. Interestingly, under the short time constraint, he consistently labeled 5 + 0 as "six," but was 100% accurate when given the longer time to examine the sets, suggesting that he needed the additional time actually to count the sets; when prevented from counting, he used the largest label available. Including those trials when he was given more time, his accuracy was overall 90%.

Unlike Sheba (or any other subjects), he was also asked to sum 0 + 0. Again, unlike Sheba and Ai, who had had extensive training to associate a null set with a label representing zero (Boysen and Berntson, 1989; Biro and Matsuzawa, 2001b), Alex had *spontaneously* associated "none" with a null set—he had had no formal training. Thus, asking him to label the total absence of something was a test of the extent of his *untrained* abilities. Interestingly, he mostly refused to answer, as though he realized that his standard number labels would not be correct. When forced to respond, on three of eight trials he eventually responded "one," using the smallest quantity label he possessed. Note that Ai also sometimes confounded 0 and 1 (Biro and Matsuzawa, 2001b). Alex's responses demonstrated that his overall understanding of the use of "none" for zero (i.e., with respect to both this and the earlier study) was comparable to that of a child just learning the concept, or of humans in cultures that do not see zero as a quantity to be labeled (Bialystok and Codd, 2000). Moreover, his occasional use of "one" clearly demonstrated that he was not simply saying "none" when he didn't know what else to say. Unfortunately, for various reasons, we did not pursue training of his use of "none" to represent zero.

We did, however, pursue his understanding of addition. In a subsequent study, we extended questions to sets totaling to eight, to adding three rather than two hidden sets (which required additional memory), and to asking him to add hidden Arabic numerals rather than sets of objects (Pepperberg, 2012). Although his death precluded testing on all possible arrays, his accuracy was statistically significant, not subject to the vagaries of an ANS, and suggested that his capacities with respect to addition were, like those of Sheba (Boysen and Berntson, 1989), spontaneously transferrable from object sets to symbolic representations of the sets. Alex's data, however, involved sums slightly beyond those of chimpanzees.

Inference of Cardinality From Ordinality

One aspect that seemed to be unique to all the non-humans during their acquisition of symbolic number labels was the lack of the so-called "bootstrapping" process that is present in young children (Carey, 2009): although, as noted earlier, the process by which children learn their first few numbers (1–4) is extremely slow (i.e., proceeds over the course of several years), they also simultaneously learn a number line—they learn to state their numerals in a specific order—even though initially the line may make little sense and the order in which they recite their numerals can be variable (Siegel, 1982; Fuson, 1988). Eventually, they learn the successor function—the ordering of their numerals stabilizes and they realize that the value of each digit in their number line is exactly one more than the previous digit—and then the bootstrapping process engages: without any further instruction they infer the meaning of numbers above 4 from this number line. In contrast, no non-human showed savings in learning as the successive numerals 5, 6, 7, etc. were added to their repertoire.

As we noted above, however, Alex's labels were initially all trained vocally, and acquisition of each label required that he learn to produce the various sounds involved. Thus, his slow acquisition of larger numbers might have been a reflection merely of this difficulty in vocal acquisition. My colleague and I set out to test this possibility (Pepperberg and Carey, 2012).

We taught Alex to identify, vocally, Arabic numerals 7 and 8 in the absence of their respective quantities (an almost year-long process), trained him that $6 < 7 < 8$ (a rapid, ~2 month process, even when interspersed with other tasks), then tested how 7 and 8 related to his other Arabic labels (Pepperberg and Carey, 2012). If he inferred the new complete number line, he could be tested on whether he, like children (≥4 years old), *spontaneously* understood that "seven" represented exactly one more than "six," that "eight" represented two more than "six" and one more than "seven," by labeling appropriate physical sets on *first* trials. Data already showed he knew that the label "six" represented six items exactly, not approximately (Pepperberg and Carey, 2012); if he succeeded, we could claim that he induced cardinal meanings of "seven" and "eight" from their ordinal positions on an implicit count list, something *no* ape (although evolutionarily closer to humans) had yet achieved.

Alex learned the novel symbolic Arabic numeral labels, placed them appropriately in his inferred number line *without* training, and quantified, on first trials, novel sets of seven and eight physical items—he did *not* have to be taught the relationship between the labels and the novel sets (Pepperberg and Carey, 2012). Thus, he responded as would children, and in a way that has *not* yet been demonstrated in any other non-human.

Why Did Alex Differ From Other Non-humans?

According to Premack (1983), symbolic representation of number should have been adequate to enable Ai and Sheba to infer ordinality from cardinality and vice versa. According to Macphail (1987), only full human language would have allowed any of the non-humans to succeed, explaining the failure of the apes but not the success of the parrot. How do we resolve this conundrum? Could the *extent* of Alex's symbolic representation be the issue? Might it be that Alex's symbolic representation abilities, although far from encompassing the range of abilities that define full human language, was more 'language-like' than that of the other non-humans that were trained with numbers?

Sheba's symbolic representation was limited solely to that of numerical sets. However, her representation abilities in this realm were robust—her ability to demonstrate spontaneous addition of novel combinations of Arabic numerals showed more than a simple association between a particular symbol and a particular set (Boysen, 1993). It is nevertheless possible that her limited range of symbolic understanding was not sufficient to enable the emergence of ordinality-cardinality comprehension. Furthermore, unlike children, she had never explicitly been taught a number line, and thus had no reason to expect any relationship between numbers and ordered lists.

Ai, in contrast, had had considerable training on the labeling of objects and colors (Matsuzawa, 1985), and in fact she was often required to label all three attributes of a set. However, she, like Sheba, also had not been taught an ordered list before, or simultaneously along with, training on the cardinal values of each numeral.

Alex had been taught labels for objects, materials, colors, and shapes (with respect to numbers of corners) before being taught numbers, but also was not trained with a number line. Moreover, Alex's number labels were not even trained in order, as noted above. Alex had, however, much more overall training on symbolic representation. He was not always asked to label everything about a set (e.g., if given three blue keys, to state "three blue key" like Ai; Matsuzawa, 1985), but had to parse his sets with respect to specific categorical labels: he could be asked "How many?" and had to respond "three key" *or* "What color?" and respond "blue key" *or* "What shape?" and respond "4-corner key" ("four-corner" being Alex's label for square items) for the same set (Pepperberg, 1983, 1999). He had also been trained on abstract concepts of same/different, such that he had to look at a pair of items and respond not as to whether they were identical or not, but, in accordance with the specific question ("What's same?" versus "What's different?"), provide the label of the one appropriate attribute (e.g., "color," "shape," "matter"; Pepperberg, 1987a). Alex, therefore, had to have acquired a more complex understanding of how his labels—vocal symbols—represented the world compared to the other non-human subjects. He knew not only that "green" was associated with, for example, both a specific key and a bean, but also that it was, along with a specific subset of other labels, hierarchically grouped under another label, "color," and likewise for his various shape, object, and material labels. His use of the order attribute + noun when multiple labels were required for identification arose through observing such use by his trainers. Alex had also already begun to parse individual labels with respect to beginnings and endings; for example, using "banerry" (banana/cherry) for an apple, and producing a label such as "carrot"—after hearing it only briefly—from his existing labels "key" and "parrot," as well as other spontaneous rearrangements in which he carefully parsed and appropriately edited beginning and endings (e.g., "grape" to "grate" to "grain" to "chain" to "cane," etc., Pepperberg, 1999). Later research provided additional examples (e.g., "spool" from the "s" sound and "wool"; Pepperberg, 2007). Thus, he had acquired some limited understanding of order that he might have abstracted for use with numbers. Notably, other apes that had learned something about label order (e.g., "put x in y," Kanzi; Savage-Rumbaugh and Lewin, 1994), were not trained or tested on number concepts, nor were Herman or Schusterman's cetaceans to any extent (see, however, Mitchell et al., 1985), and therefore comparative data are lacking.

Of course, the question still arises as to whether we can ever determine the extent to which symbolic representation—whether in the absence of ordering or with only limited understanding of ordering (rudimentary syntax)—can affect changes in cognitive processing in non-humans. A recent study suggests that to have such an effect, even symbolic representation by itself must be rich and varied, and not be a simple case of associative learning. This project involved extremely limited symbolic learning and a match-to-sample task (Bowden et al., 2019). Here researchers compared children's (ages 3–5) and monkeys' abilities to learn to use different icons to represent different types of matching strategies—for example, a circle meant match with respect to color whereas a cross meant match with respect to shape. Both sets of subjects learned the basic rule, but only the children could generalize to novel colors and shapes. Although children in this age range still lack full expressive language and are just beginning to learn about symbols (see discussion in Deloache, 2004), their levels of symbolic representation far outstrip those of monkeys that had been trained only on two specific symbols. Thus, not only symbolic representation, but also the extent and richness of this representation, seem critical in enabling non-humans to succeed on various tasks designed to examine their intelligence.

A ROLE FOR FORMAL SYNTAX?

Obviously, some of what I have described so far could occur only in a subject that always had its entire repertoire available for use, that was a vocal learner, and that had human interaction for extended periods outside of training and testing sessions such that spontaneous expressions could both be noted and evoke responses from caretakers. The complicated tasks about which I have written, however, do not require formal syntax. Thus, despite all my comments on aspects of communication that could be considered to involve some type of ordering, I have not yet commented to any extent on Macphail's basic claim with respect to the effects of the formal syntax of human language on cognitive processing. The reason for my hesitancy is the paucity of existing data and the dearth of recent (and the poor prospect for future) studies on interspecies communication, making the probability of acquiring additional knowledge unlikely (note Pepperberg, 2017).

What we currently know is of limited value: in the laboratory, some species that have learned symbolic representation have also learned something about rule-governed behavior, something clearly much simpler than formal syntax. Have researchers instilled this behavior or does it build on something already existent in nature? Some level of rule-governed behavior may exist for some species in the wild but not for others. Cetaceans learned to respond to particular orderings of symbols, but evidence for order-related meaning in their natural communication system is lacking (review in Suzuki and Zuberbühler, 2019). For some bird species that learn their songs, note and syllable order is crucial for meaning; for other species it is not (see review in Weisman and Ratcliffe, 1987). Such may also be the case for certain bird calls—in particular instances, when the order of the elements is altered, birds fail to respond in playback tests (Suzuki et al., 2019), suggesting that some sort of rules for the production and comprehension of vocalizations exist in species separated from humans by over 300 million years of evolution (Hedges et al., 1996). Evidence for what could, in some sense, be considered combinatorial order in non-human primate vocalizations in the wild has been demonstrated only in a few—primarily monkey—species (reviewed in Zuberbüler and Neumann, 2017; Suzuki and Zuberbühler, 2019), but so far seems to be used mostly in terms of modifying the level of communicative intent. Kanzi, a bonobo, demonstrated rule-governed behavior in symbolic comprehension in the laboratory (Savage-Rumbaugh and Lewin, 1994), but no evidence for such behavior has yet been discovered in the wild. Whether meanings of parrot vocalizations in the wild depend on order has not been examined, but Alex exhibited clear—if limited—sensitivity to order in various non-numerical aspects of his vocal symbolic behavior (e.g., his ability to segment, Pepperberg, 2007, his use of sentence frames "I want X" and "Wanna go Y," where he knew that an object "X" must follow "want" but a location "Y" must follow "wanna go," Pepperberg, 1999). Such comprehension could have been transferred to his understanding of, for example, numerical label order. Another parrot, Griffin, has also demonstrated some rule-governed behavior, as shown during a study of his acquisition of label order (Pepperberg and Shive, 2001). However, evidence for formal syntax, as opposed to simple rule-governed behavior (what can be seen as a proto-syntax) is still lacking.

These findings lead to a number of questions, particularly with respect to syntax. Does the ability to learn rule-governed behavior patterns in addition to symbolic representation simply indicate higher processing power, such that non-humans that are capable of acquiring such competence would therefore be expected to succeed on more complex tasks, and that humans with their demonstration of fully syntactic language are consequently at the apex of such behavior? Or does acquisition of rule-governed behavior in addition to symbolic representation even affect how the individual can process information? Or have brains and corresponding behavioral complexity evolved in lock-step, each synergistically supporting the next evolutionary stage (Pepperberg, 2007, 2010)? We have yet to fully determine the answers to these questions. It is likely that the effect of syntax on intelligent behavior is not easily specified, nor is the type of non-linguistic task for which formal syntax would be necessary.

CONCLUSION

My overall conclusion is that differences *do* exist among various species' abilities, that these differences are not due entirely to contextual variables, but that when individuals of these species are given appropriate training, the differences are not as great as we once may have thought. The training cannot, however, merely be with respect to simple associations between a limited number of labels and their corresponding items, but must be rich enough to encompass concepts and enable the subjects to transfer the learned concepts to novel situations. Addition of some level of rule-governed behavior would also seem important.

However, one additional issue must also be addressed when looking at the differences between humans and all other species as well as across non-human species, and when looking at the effects of language or at least symbolic representation on performance: most of the tasks that Macphail describes for comparing intelligence across species involve detecting contingencies, and little else; however, ". . .tasks likely to be relevant to comparative psychologists interested in intelligence" (Macphail, 1987)—tasks that actually involve more complex forms of information processing—are all designed by humans, generally are based on tasks that are appropriate for, and presented to, adult humans and/or human children, and thus are inherently biased in favor of humans who have language. Several original commentators argue the point about how the tasks proposed by Macphail fail to examine many of the more interesting qualities of information processing, but none of the commentaries fully examine the extent to which symbolic representation—or the lack thereof—may affect how well any tasks can be solved by non-humans when these are tasks that humans design from their point of view as language learners, or even how well the tasks designed for non-humans by humans will actually test what they are designed to test. From the studies I have described, it is obvious that I am not against using human-based tasks to test non-humans. My data do, however, suggest the existence of inherent biases in such tasks, given how training non-humans on symbolic representation turns out to be so important for their ability to succeed on these tasks. As researchers, we need to be aware of such issues when making claims about non-human intelligence.

In sum, despite the still-contemporary feel of "The Comparative Psychology of Intelligence," article and commentaries, much has indeed changed in the intervening 30 years. In 1987, an extremely high number of papers in psychological journals (see Burghardt, 2006) still generally involved only rodents or pigeons performing some kind of experiment using operant conditioning. Now, we can access studies about everything from ants to lizards to horses to elephants that examine everything from visual illusions to social cooperation to spatial orientation to delayed gratification. New techniques have given us access to levels of neurophysiological and neuroanatomical information that enable incredibly detailed cross-species analyses. We have new statistical tools and modeling algorithms—and the computational power to use them—that few could have foreseen. Many of us studying non-humans are asked to collaborate on projects spanning AI to SETI. Macphail (1987) may not have foreseen the actual future of comparative psychology, but we must give him credit for instigating a variety of controversies, stimulating the wide-ranging discussions, and generating the types of challenges that have led to many new avenues of research.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This manuscript was written with the support of donors to *The Alex Foundation*. The research described was supported by donors to *The Alex Foundation* and fellowships and grants cited in the referenced journal articles.

ACKNOWLEDGMENTS

I thank Francesca Cornero and Gabriella Smith for their helpful comments.

REFERENCES

Asano, T., Kojima, T., Matsuzawa, T., Kubota, K., and Murofushi, T. (1982). Object and color naming in chimpanzees (*Pan troglodytes*). *Proc. Jpn. Acad.* 58B, 118–122. doi: 10.2183/pjab.58.118

Basil, J. A., Kamil, A. C., Balda, R. P., and Fite, K. V. (1996). Differences in hippocampal volume among food-storing corvids. *Brain Behav. Evol.* 47, 156–164. doi: 10.1159/000113235

Bialystok, E., and Codd, J. (2000). Representing quantity beyond whole numbers: some, none and part. *Can. J. Exp. Psychol.* 54, 117–128. doi: 10.1037/h0087334

Biro, D., and Matsuzawa, T. (2001a). "Chimpanzee numerical competence," in *Primate Origins of Human Cognition and Behavior*, ed. T. Matsuzawa (Hong Kong: Springer), 199–225. doi: 10.1007/978-4-431-09423-4_10

Biro, D., and Matsuzawa, T. (2001b). Use of numerical symbols by the chimpanzee (*Pan troglodytes*): cardinals, ordinals, and the introduction of zero. *Anim. Cogn.* 4, 193–199. doi: 10.1007/s100710100086

Bowden, M., Jackson, B., Church, B., and Smith, J. D. (2019). Instructed matching strategies: learning symbols to guide decision making. *Poster Presented at the 2019 Meeting of the Psychonomic Society*, Montreal.

Boysen, S. T. (1993). "Counting in chimpanzees: nonhuman principles and emergent properties of number," in *The Development of Numerical Competence: Animal and Human Models*, eds S. T. Boysen and E. J. Capaldi (Hillsdale, NJ: Erlbaum), 39–59.

Boysen, S. T., and Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *J. Comp. Psychol.* 103, 23–31. doi: 10.1037/0735-7036.103.1.23

Boysen, S. T., Berntson, G. G., Shreyer, T. A., and Quigley, K. S. (1993). Processing of ordinality and transitivity by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 107, 208–215. doi: 10.1037/0735-7036.107.2.208

Brannon, E. M., and Merritt, D. J. (2011). "Evolutionary foundations of the approximate number system," in *Space, Time, and Number: Searching for the Foundations of Mathematical Thought*, eds S. Dehaene and E. M. Brannon (London: Academic Press), 207–224. doi: 10.1016/b978-0-12-385948-8.00014-1

Burghardt, G. M. (2006). Editorial. *J. Comp. Psychol.* 120, 77–78.

Call, J., Burghardt, G., Pepperberg, I. M., Snowdon, C., and Zentall, T. R. (eds) (2017). *The APA Handbook of Comparative Psychology*. Washington, DC: APA.

Carey, S. (2009). *The Origin of Concepts*. New York, NY: Oxford.

Chakraborty, M., Walløe, S., Nedergaard, S., Fridel, E. E., Dabelsteen, T., Pakkenberg, B., et al. (2015). Core and shell song systems unique to the parrot brain. *PLoS ONE* 10:e0118496. doi: 10.1371/journal.pone.0118496

Church, B. A. (2019). The evolutionary emergence of explicit categorization. *Paper Presented at the 2019 Meeting of the Psychonomic Society*, Montreal, Canada.

Clements, K., Gray, S. L., Gross, B., and Pepperberg, I. M. (2018). Initial evidence for probabilistic learning by a Grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 132, 166–177. doi: 10.1037/com0000106

Cynx, J., Hulse, S. H., and Polyzois, S. (1986). A psychophysical measure of pitch discrimination loss resulting from a frequency range constraint in European starlings (*Sturnus vulgaris*). *J. Exp. Psychol. Anim. Behav. Proc.* 12, 394–402. doi: 10.1037/0097-7403.12.4.394

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition* 44, 1–42. doi: 10.1016/0010-0277(92)90049-n

Deloache, J. S. (2004). Becoming symbol minded. *Trends Cognit. Sci.* 8, 66–70. doi: 10.1016/j.tics.2003.12.004

Everett, D. L. (2005). Cultural constraints on grammar and cognition in the Pirahã: another look at the design features of human language. *Curr. Anthropol.* 46, 621–646. doi: 10.1086/431525

Fouts, R., and Rigby, R. (1977). "Man-chimpanzee communication," in *How Animals Communicate*, ed. T. Sebeok (Bloomington, IN: Indiana University Press), 1034–1054.

Frank, M. C., Everett, D. L., Fedorenko, E., and Gibson, E. (2008). Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* 108, 819–824. doi: 10.1016/j.cognition.2008.04.007

Fuson, K. (1988). *Children's Counting and Concepts of Number*. New York, NY: Springer-Verlag.

Gardner, R. A., and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 165, 664–672. doi: 10.1126/science.165.3894.664

Gould, K. L., Gilbertson, K. E., Seyfer, A. L., Brantner, R. M., Hrvol, A. J., Kamil, A. C., et al. (2013). Differences in relative hippocampus volume and number of hippocampus neurons among five corvid species. *Brain Behav. Evol.* 81, 56–70. doi: 10.1159/000345560

Greeno, J. G., Riley, M. S., and Gelman, R. (1984). Conceptual competence and children's counting. *Cognit. Psychol.* 16, 94–143. doi: 10.1016/0010-0285(84)90005-7

Güntürkün, O., and Bugnyar, T. (2016). Cognition without cortex. *TiCS* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Gutiérrez-Ibáñez, C., Iwaniuk, A. N., and Wylie, D. R. (2018). Parrots have evolved a primate-like telencephalic-midbrain-cerebellar circuit. *Sci. Rep.* 8:9960. doi: 10.1038/ s41598-018-28301-4

Halberda, J., Mazzocco, M., and Feigenson, L. (2008). Individual differences in nonverbal number acuity predicts maths achievement. *Nature* 455, 665–669.

Hanggi, E. B. (2003). Discrimination learning based on relative size concepts in horses (*Equus caballus*). *Appl. Anim. Behav. Sci.* 83, 201–213. doi: 10.1016/s0168-1591(03)00136-9

Hedges, S. B., Parker, P. H., Sibley, C. G., and Kumar, S. (1996). Continental breakup and the ordinal diversification of birds and mammals. *Nature* 381, 226–229. doi: 10.1038/381226a0

Herman, L., Richards, D. G., and Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition* 16, 129–219. doi: 10.1016/0010-0277(84)90003-9

Heyes, C. (2016). Homo imitans? Seven reasons why imitation couldn't possibly be associative. *Philos. Trans. R. Soc. B* 371:20150069. doi: 10.1098/rstb.2015.0069

Hulse, S. H., Cynx, J., and Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *J. Exp. Psychol. Gen.* 113, 28–54.

Hulse, S. H., Page, S. C., and Braaten, R. F. (1990). Frequency range size and the frequency range constraint in auditory perception by European starlings (*Sturnus vulgaris*). *Anim. Learn. Behav.* 18, 238–245. doi: 10.3758/bf03205281

Hurford, J. R. (1987). *Language and Number*. Oxford: Blackwell.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., and Hedges, L. V. (2010). Source of variability in children's language growth. *Cogn. Psychol.* 61, 343–365.

Iwaniuk, A. N., Dean, K. M., and Nelson, J. E. (2005). Interspecific allometry of the brain and brain regions in parrots (*Psittaciformes*): comparisons with other birds and primates. *Brain Behav. Evol.* 65, 40–59. doi: 10.1159/000081110

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159.

Kamil, A. C. (1984). "Adaptation and cognition: knowing what comes naturally," in *Animal Cognition*, eds H. L. Roitblat, T. G. Bever, and H. S. Terrace (Hillsdale, NJ: Erlbaum), 533–544.

Livingstone, M. S., Srihasam, K., and Morocz, I. A. (2010). The benefit of symbols: monkeys show linear, human-like, accuracy when using symbols to represent scalar value. *Anim. Cogn.* 13, 711–719. doi: 10.1007/s10071-010-0321-1

MacDougal-Shackleton, S. A., and Hulse, S. H. (1995). Concurrent absolute and relative pitch processing by European starlings (*Sturnus vulgaris*). *J. Comp. Psychol.* 110, 139–146. doi: 10.1037/0735-7036.110.2.139

Macphail, E. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature* 315, 57–59. doi: 10.1038/315057a0

Mazengenya, P., Bhagwandin, A., Manger, P. R., and Ihunwo, A. O. (2018). Putative adult neurogenesis in Old World parrots: the congo African Grey parrot (*Psittacus erithacus*) and the timneh Grey parrot (*Psittacus timneh*). *Front. Neuroanat.* 12:7. doi: 10.3389/fnana.2018.00007

McLaren, I. P. L., McAndrew, A., Angerer, K., McLaren, R., Forrest, C., Bowditch, W., et al. (2018). Mackintosh lecture: association and cognition: two processes one system. *Q. J. Exp. Psychol.* 72, 98–117.

Meltzoff, A. N., and Moore, M. K. (1999). "Resolving the debate about early imitation," in *The Blackwell Reader in Developmental Psychology*, eds A. Slater and D. Muir (Oxford: Blackwell), 151–155.

Mitchell, R. W., Yao, P., Sherman, P. T., and O'Regan, M. (1985). Discriminative responding of a dolphin (*Tursiops truncates*) to differentially rewarded stimuli. *J. Comp. Psychol.* 99, 218–225. doi: 10.1037/0735-7036.99.2.218

Mountjoy, P. T. (1958). Effects of exposure time and intertrial interval upon decrement to the Müller-Lyer illusion. *J. Exp. Psychol.* 56, 97–102. doi: 10.1037/h0041307

Nakamura, N. N., Fujita, K. F., Ushitani, T. U., and Miyata, H. M. (2006). Perception of the standard and the reversed Müller-Lyer figures in pigeons (*Columba livia*) and humans (*Homo sapiens*). *J. Comp. Psychol.* 120, 252–261. doi: 10.1037/0735-7036.120.3.252

Obozova, T., Smirnova, A., Zorina, Z., and Wasserman, E. A. (2015). Analogical reasoning in Amazons. *Anim. Cogn.* 18, 1363–1371. doi: 10.1007/s10071-015-0882-0

Olkowicz, S., Kocourek, M., Lucan, R. K., Porteš, M., Tecumseh Fitch, W., Herculano-Houzel, S., et al. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7255–7260. doi: 10.1073/pnas.1517131113

Olson, D. J. (1991). Species differences in spatial memory among Clark's nutcrackers, scrub jays, and pigeons. *J. Exp. Psychol. Anim. Behav. Proc.* 17, 363–376. doi: 10.1037/0097-7403.17.4.363

Olson, D. J., Kamil, A. C., Balda, R. P., and Nims, P. J. (1995). Performance of four seed caching corvid species in operant tests of nonspatial and spatial memory. *J. Comp. Psychol.* 109, 173–181. doi: 10.1037/0735-7036.109.2.173

Page, S. C., Hulse, S. H., and Cynx, J. (1989). Relative pitch perception in the European starling (*Sturnus vulgaris*): further evidence for an elusive phenomenon. *J. Exp. Psychol. Anim. Behav. Proc.* 15, 137–146. doi: 10.1037/0097-7403.15.2.137

Pasnak, R., and Kurtz, S. L. (1987). Brightness and size transposition by rhesus monkeys. *Bull. Psychonom. Soc.* 25, 109–112. doi: 10.3758/bf03330299

Patterson, D. K., and Pepperberg, I. M. (1994). A comparative study of human and parrot phonation: I. Acoustic and articulatory correlates of vowels. *JASA* 96, 634–648.

Patterson, D. K., and Pepperberg, I. M. (1998). A comparative study of human and Grey parrot phonation: II. Acoustic and articulatory correlates of stop consonants. *JASA* 103, 2197–2213.

Patterson, F. G. (1978). The gestures of a gorilla: language acquisition in another pongid. *Brain Lang.* 5, 72–97. doi: 10.1016/0093-934x(78)90008-1

Pepperberg, I. M. (in press). "Symbolic communication in the Grey parrot," in *Cambridge Handbook of Animal Cognition*, eds A. Kaufman and J. Call (Cambridge: Cambridge University Press).

Pepperberg, I. M. (1983). Cognition in the African Grey parrot: preliminary evidence for auditory/vocal comprehension of the class concept. *Anim. Learn. Behav.* 11, 179–185.

Pepperberg, I. M. (1987a). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. *Anim. Learn. Behav.* 15, 423–432.

Pepperberg, I. M. (1987b). Evidence for conceptual quantitative abilities in the African Grey parrot: labeling of cardinal sets. *Ethology* 75, 37–61.

Pepperberg, I. M. (1988). Acquisition of the concept of absence by an African Grey parrot: learning with respect to questions of same/different. *J. Exp. Anal. Behav.* 50, 553–564.

Pepperberg, I. M. (1990). Cognition in an African Grey parrot (*Psittacus erithacus*): further evidence for comprehension of categories and labels. *J. Comp. Psychol.* 104, 41–52.

Pepperberg, I. M. (1994). Numerical competence in an African Grey parrot. *J. Comp. Psychol.* 108, 36–44.

Pepperberg, I. M. (1999). *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Cambridge, MA: Harvard University Press.

Pepperberg, I. M. (2006a). Grey parrot numerical competence: a review. *Anim. Cogn.* 9, 377–391. doi: 10.1007/s10071-006-0034-7

Pepperberg, I. M. (2006b). Grey parrot (*Psittacus erithacus*) numerical abilities: addition and further experiments on a zero-like concept. *J. Comp. Psychol.* 120, 1–11. doi: 10.1037/0735-7036.120.1.1

Pepperberg, I. M. (2006c). Ordinality and inferential abilities of a Grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 120, 205–216. doi: 10.1037/0735-7036.120.3.205

Pepperberg, I. M. (2007). Grey parrots do not always 'parrot': phonological awareness and the creation of new labels from existing vocalizations. *Lang. Sci.* 29, 1–13. doi: 10.1016/j.langsci.2005.12.002

Pepperberg, I. M. (2010). Vocal learning in Grey parrots: a brief review of perception, production, and cross-species comparisons. *Brain Lang*. 115, 81–91. doi: 10.1016/j.bandl.2009.11.002

Pepperberg, I. M. (2012). Further evidence for addition and numerical competence by a Grey parrot (*Psittacus erithacus*). *Anim. Cogn.* 15, 711–717. doi: 10.1007/s10071-012-0470-5

Pepperberg, I. M. (2017). Animal language studies: what happened? *Psychol. Bull. Rev*. 24, 181–185. doi: 10.3758/s13423-016-1101-y

Pepperberg, I. M., and Brezinsky, M. V. (1991). Acquisition of a relative class concept by an African Grey parrot (*Psittacus erithacus*): discriminations based on relative size. *J. Comp. Psychol.* 105, 286–294. doi: 10.1037/0735-7036.105.3.286

Pepperberg, I. M., and Carey, S. (2012). Grey parrot number acquisition: the inference of cardinal value from ordinal position on the numeral list. *Cognition* 125, 219–232. doi: 10.1016/j.cognition.2012.07.003

Pepperberg, I. M., and Gordon, J. D. (2005). Number comprehension by a Grey parrot (*Psittacus erithacus*), including a zero-like concept. *J. Comp. Psychol.* 119, 197–209. doi: 10.1037/0735-7036.119.2.197

Pepperberg, I. M., and Nakayama, K. (2016). Robust representation of shape by a Grey parrot (*Psittacus erithacus*). *Cognition* 153, 146–160. doi: 10.1016/j.cognition.2016.04.014

Pepperberg, I. M., and Pailian, H. (2017). Evolution of mechanisms underlying visual working memory manipulation: when 'bird-brain' is a compliment. *Paper Presented at the Vision Science Society Annual Conference*, St. Petersburg, FL.

Pepperberg, I. M., and Shive, H. A. (2001). Simultaneous development of vocal and physical object combinations by a Grey parrot (*Psittacus erithacus*): bottle caps, lids, and labels. *J. Comp. Psychol.* 115, 376–384. doi: 10.1037/0735-7036.115.4.376

Pepperberg, I. M., Vicinay, J., and Cavanagh, P. (2008). Processing of the Müller-Lyer illusion is processed by a Grey parrot (*Psittacus erithacus*). *Perception* 37, 765–781. doi: 10.1068/p5898

Petrazzini, M. E. M., Agrillo, C., Izard, V., and Bisazza, A. (2015). Relative versus absolute numerical representation: can guppies represent "fourness?" *Anim. Cogn.* 18, 1007–1017. doi: 10.1007/s10071-015-0868-y

Petrazzini, M. E. M., Fraccaroli, I., Gariboldi, F., Gariboldi, F., Agrillo, C., Bisazza, A., et al. (2017). Quantitative abilities in a reptile (*Podarcis sicula*). *Biol. Lett.* 13:20160899. doi: 10.1098/rsbl.2016.0899

Predebon, J. (1998). Decrement of the Brentano Müller-Lyer illusion as a function of inspection time. *Perception* 27, 183–192. doi: 10.1068/p270183

Predebon, J. (2006). Decrement of the Müller-Lyer and Poggendorff illusions: the effects of inspection and practices. *Psychol. Res.* 70, 384–394. doi: 10.1007/s00426-005-0229-6

Premack, D. (1971). Language in chimpanzees? *Science* 172, 808–822.

Premack, D. (1978). "On the abstractness of human concepts: why it would be difficult to talk to a pigeon," in *Cognitive Processes in Animal Behavior*, eds S. H. Hulse, H. Fowler, and W. K. Honig (Hillsdale, NJ: Erlbaum), 425–451.

Premack, D. (1983). The codes of man and beast. *Behav. Brain Sci.* 6, 125–167.

Rhein, C., Mühle, C., Richter-Schmidinger, T., Alexopoulos, P., Doerfler, A., and Kornhuber, J. (2014). Neuroanatomical correlates of intelligence in healthy young adults: the role of basal ganglia volume. *PLoS ONE* 9:e93623. doi: 10.1371/journal.pone.0093623

Rinnert, P., Kirschhock, M. E., and Nieder, A. (2019). Neuronal correlates of spatial working memory in the endbrain of crows. *Curr. Biol.* 29, 2616–2624.

Rosa Salva, O., Sovrano, V. A., and Vallortigara, G. (2014). What can fish brains tell us about visual perception? *Front. Neural Circuit* 8:119. doi: 10.3389/fncir.2014.00119

Sadza, K. J., and de Weert, C. M. M. (1984) Influence of color and luminance on the Müller-Lyer illusion. *Percept. Psychophys.* 35, 214–220. doi: 10.3758/bf03205934

Savage-Rumbaugh, E. S., Pate, J. L., Lawson, J., Smith, S. T., and Rosenbaum, S. (1983). Can a chimpanzee make a statement? *J. Exp. Psychol.: Gen.* 112, 457–492. doi: 10.1037/0096-3445.112.4.457

Savage-Rumbaugh, S., and Lewin, R. (1994). *Kanzi*: *The Ape at the Brink of the Human Mind*. New York, NY: John Wiley and Sons.

Schusterman, R. J., and Gisiner, R. (1988). Artificial language comprehension in dolphins and sea lions: the essential skills. *Psychol. Rec.* 38, 311–348. doi: 10.1007/bf03395027

Schusterman, R. J., and Krieger, K. (1986). Artificial language comprehension and size transposition by a California sea lion (*Zalophus californianus*). *J. Comp. Psychol.* 100, 348–355. doi: 10.1037/0735-7036.100.4.348

Shimizu, T., Patton, T. B., and Husband, S. A. (2010). Avian visual behavior: organization of the telencephalon. *Brain Behav. Evol.* 75, 204–217. doi: 10.1159/000314283

Sidman, M., Wynne, C. K., Maguire, R. W., and Barnes, T. (1989). Functional classes and equivalence relations. *J. Exp. Anal. Behav.* 52, 261–274. doi: 10.1901/jeab.1989.52-261

Siegel, L. S. (1982). "The development of quantity concepts: perceptual and linguistic factors," in *Children's Logical and Mathematical Cognition*, ed. C. J. Brainerd (Berlin: Springer-Verlag), 123–155. doi: 10.1007/978-1-4613-9466-2_4

Simpson, E. A., Murray, L., Paukner, A., and Ferrari, P. F. (2014). The mirror neuron system as revealed through neonatal imitation: presence from birth, predictive power and evidence of plasticity. *Philos. Trans. R. Soc. B* 369:20130289. doi: 10.1098/rstb.2013.0289

Smith, J. D. (2019). Animal metacognition: the dolphin's tale. *Paper Presented at the 2019 Meeting of the Psychonomic Society*, Montreal.

Spelke, E. S., and Kinzler, K. D. (2007). Core knowledge. *Dev. Sci.* 10, 89–96.

Suzuki, T. N., Griesser, M., and Wheatcroft, D. (2019). Syntactic rules in avian vocal sequences as a window into the evolution of compositionality. *Anim. Behav.* 151, 267–274. doi: 10.1016/j.anbehav.2019.01.009

Suzuki, T. N., and Zuberbühler, K. (2019). Animal syntax. *Curr. Biol.* 29, R669–R671. doi: 10.1016/j.cub.2019.05.045

Terrace, H. S., Petitto, L. A., Sanders, R. J., and Bever, T. J. (1979). Can an ape create a sentence? *Science* 206, 891–902. doi: 10.1126/science.504995

Thomas, R. K. (1980). Evolution of intelligence: an approach to its assessment. *Brain Behav. Evol.* 17, 454–472. doi: 10.1159/000121814

Tomonaga, M., Matsuzawa, T., and Itakura, S. (1993). Teaching ordinals to a cardinal trained chimpanzee. *Primate Res.* 9, 67–77. doi: 10.2354/psj.9.2_67

Tosches, M. A., Yamawaki, T. M., Naumann, R. K., Jacobi, A. A., Tushev, G., and Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles. *Science* 360, 881–888. doi: 10.1126/science.aar4237

Trick, L., and Pylyshyn, Z. (1989). *Subitizing and the FNST Spatial Index Model* (Tech. *Rep. No. COGMEM* 44). London: University of Ontario.

Vincini, S., Jhang, Y., Buder, E. H., and Gallagher, S. (2017). Neonatal imitation: theory, experimental design, and significance for the field of social cognition. *Front. Psychol.* 8:1323. doi: 10.3389/fpsyg.2017.01323

Vonk, J. (2015). Corvid cognition: something to crow about? *Curr. Biol.* 25, R.69–R.71.

Vonk, J., and Beran, M. J. (2012). Bears "count" too: quantity estimation and comparison in black bears (*Ursus americanus*). *Anim. Behav.* 84, 231–238. doi: 10.1016/j.anbehav.2012.05.001

Warden, D. J., and Baar, J. (1929). The Müller-lyer illusion in the ring dove. *Turtur risorius*. *J. Comp. Psychol.* 9, 275–292. doi: 10.1037/h0071052

Weisman, R., and Ratcliffe, L. (1987). How birds identify species information in song: a pattern recognition approach. *Learn. Motiv.* 18, 80–98. doi: 10.1016/0023-9690(87)90024-5

Wen, X., Cant, J., Xiang, Y., Huang, R., and Mo, L. (2019). The neural correlates of intelligence comparison. *Cereb. Cortex* 29, 253–264.

Wiese, H. (2003). Iconic and non-iconic stages in number development: the role of language. *Trends Cogn. Sci.* 7, 385–390. doi: 10.1016/s1364-6613(03)00192-x

Wills, A. J., Edmunds, C. E. R., LePelly, M. E., et al. (2019). Dissociable learning processes, associative theory, and testimonial reviews: a comment on Smith and Church (2018). *Psychon. Bull. Rev.* 26, 1988–1993. doi: 10.3758/s13423-019-01644-3

Wills, S., and Mackintosh, N. J. (1999). Relational learning in pigeons? *Q. J. Exp. Psychol.* 52B, 31–52.

Winslow, C. N. (1933). Visual illusions in the chick. *Arch. Psychol.* 153, 1–83.

Wright, A. A., Kelly, D., and Katz, J. (2018). Comparing cognition by integrating concept learning, proactive interference, and list memory. *Learn. Behav.* 46, 107–123. doi: 10.3758/s13420-018-0316-3

Wynn, K. (1990). Children's understanding of counting. *Cognition* 36, 155–193. doi: 10.1016/0010-0277(90)90003-3

Young, M. E., and Wasserman, E. A. (2001). Entropy and variability discrimination. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 278–293. doi: 10.1037/0278-7393.27.1.278

Zuberbüler, K., and Neumann, C. (2017). "Referential communication in nonhuman animals," in *The Handbook of Comparative Psychology*, Vol. 1, eds J. Call, G. Burghardt I, M. Pepperberg, C. T. Snowdon, and T. R. Zentall (Washington, DC: APA), 645–661. doi: 10.1037/0000011-031

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

REVIEW
published: 08 July 2020
doi: 10.3389/fpsyg.2020.01692

Macphail's Null Hypothesis of Vertebrate Intelligence: Insights From Avian Cognition

Amalia P. M. Bastos† and Alex H. Taylor†*

School of Psychology, The University of Auckland, Auckland, New Zealand

Macphail famously criticized two foundational assumptions that underlie the evolutionary approach to comparative psychology: that there are differences in intelligence across species, and that intelligent behavior in animals is based on more than associative learning. Here, we provide evidence from recent work in avian cognition that supports both these assumptions: intelligence across species varies, and animals can perform intelligent behaviors that are not guided solely by associative learning mechanisms. Finally, we reflect on the limitations of comparative psychology that led to Macphail's claims and suggest strategies researchers can use to make more advances in the field.

Keywords: null hypothesis, avian cognition, comparative psychology, differences in intelligence, avian intelligence

Edited by:
Thomas Zentall,
University of Kentucky, United States

Reviewed by:
Michael Colombo,
University of Otago, New Zealand
Lucia Regolin,
University of Padua, Italy

****Correspondence:***
Amalia P. M. Bastos
a.bastos@auckland.ac.nz

†ORCID:
Amalia P. M. Bastos
0000-0003-3037-687X
Alex H. Taylor
0000-0003-3492-7667

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *30 April 2020*
Accepted: *22 June 2020*
Published: *08 July 2020*

Citation:
Bastos APM and Taylor AH (2020)
Macphail's Null Hypothesis
of Vertebrate Intelligence: Insights
From Avian Cognition.
Front. Psychol. 11:1692.
doi: 10.3389/fpsyg.2020.01692

INTRODUCTION

Euan Macphail sparked great controversy in the 1980s, following his synthesis of the current state of comparative psychology. Macphail argued that, given the body of evidence available at the time, there appeared to be no quantitative or qualitative differences in intelligence across species (Macphail, 1982, 1985; Macphail and Bolhuis, 2001), and that seemingly intelligent behavior is underpinned by associative learning (Macphail, 1982, 1985; Macphail and Bolhuis, 2001). In Macphail's general process view, species differences in performance within tasks could be ascribed to contextual variables, rather than to any real differences in their underlying cognition. He directly pitted this view against the widely-regarded ecological view (Macphail and Bolhuis, 2001), which takes an evolutionary perspective of cognition, suggesting that species evolve cognitive adaptations to their environment, just as they do physiological adaptations. Here, we will consider two of these lines of Macphail's criticism in light of recent developments in avian cognition: namely, that there are differences in intelligence across species, and that intelligent behavior cannot be explained by associative learning alone.

DIFFERENCES IN INTELLIGENCE ACROSS SPECIES

Macphail's null hypothesis of vertebrate intelligence posits that all animals use the same general mechanisms, to the same level of ability, to solve cognitive tasks. Whilst this may be true in considering some basic processes such as operant conditioning, which appear to be universal across species, this hypothesis fails to consider differences in intelligence at finer, and more ecologically relevant, scales (Shettleworth, 1987).

A strong line of evidence suggesting not only that intelligence is quantitatively different across species, but that these differences emerge as a direct consequence of their ecology, explores the relationship between species' social complexity and cognitive task performance. The social intelligence hypothesis predicts that complex social environments require better memory and overall cognitive capacity, and so social complexity drives the evolution of intelligence (Dunbar, 1998, 2008). Comparative studies across both parrot and corvid species provide support for this hypothesis. For example, parrots living in complex groups involving fission-fusion dynamics outperform those that form smaller and more stable family groups in string-pulling tasks testing means-end comprehension (Krasheninnikova et al., 2013). Similarly, pinyon jays, which live in large flocks of up to five hundred individuals, outperform western scrub-jays, which form small family groups, in tasks of transitive inference (Bond et al., 2003). Pinyon jays also outperform two less social species, western scrub-jays and Clark's nutcrackers, in both color and spatial reversal tasks (Bond et al., 2007). Correlations between social complexity and cognitive capacity may be particularly strong in the corvids and parrots, due to the long life expectancy in these species, which might facilitate exposure to a greater number of social partners over time (Boucherie et al., 2019).

There is also good evidence that quantitative cognitive differences between species are driven by their ecological differences in comparative work between caching and non-caching corvids. While caching and non-caching corvids perform, similarly, in a color-based task, caching species outperform non-caching species in a spatial task (Olson et al., 1995). Findings from studies such as these suggest that ecology plays an important role in shaping the cognitive abilities of species. Given that comparative psychology has so far been restricted to a minority of species, it seems likely that as a greater number of species are tested, more differences in intelligence are likely to emerge (Elepfandt, 1987; Kamil, 1987; Shultz and Dunbar, 2010; van Horik and Emery, 2011), generating more clear and testable links between differences in ecology and cognitive ability.

Evidence for convergent evolution in the cognitive abilities of great apes, corvids, and parrots also suggests that quantitative differences in intelligence do exist across species, and that these differences relate to their ecology (Emery, 2004; Clayton, 2012; Emery et al., 2012; Güntürkün and Bugnyar, 2016; Auersperg and von Bayern, 2019). One clear prediction that the contextual variable argument makes is that differences in methodology should become more difficult to control for and, therefore, have a greater impact on task performance, the further apart two species are phylogenetically. This is because species that are more similar are more likely to share more of the same perceptual abilities and biases than those that are more distantly related. Thus, if species do not differ in intelligence, as Macphail claims, we should expect problem solving performances to differ more as phylogenetic distance increases, due to contextual variables becoming harder to control.

However, the great apes, parrots, and corvids, despite being evolutionarily distant taxa, converge in several of their cognitive abilities (Emery, 2004, 2006; Emery and Clayton, 2004; Seed et al., 2009; Clayton, 2012; Emery et al., 2012; Güntürkün and Bugnyar, 2016; Auersperg and von Bayern, 2019). The Piagetian framework for object permanence describes different stages of development for this ability, which requires an individual to understand that an object continues to exist when hidden within a container (Piaget, 1954). Its final stage requires an understanding of invisible displacement: that is, tracking a container which presumably contains the hidden object as it moves behind a series of screens or occluders, and guessing where it may have been left once the container is shown to be empty. The great apes (de Blois et al., 1998; Call, 2001; Collier-Baker and Suddendorf, 2006; Collier-Baker et al., 2006; Mallavarapu, 2009), corvids (Pollok et al., 2000; Zucca et al., 2007; Hoffmann et al., 2011; Ujfalussy et al., 2013), and parrots (Pepperberg and Funk, 1990; Pollok et al., 2000) succeed at the final stage of object permanence, even though other species of both birds and mammals do not. Four species of lemurs (Deppe et al., 2009; Mallavarapu, 2009) succeed only at visible displacement tasks, where the reward can be seen as it moves between two or more occluders. Several other mammals also fail to understand invisible displacement tasks (for a review, see Jaakkola, 2014), even though they understand visible displacements, suggesting that contextual variables are not to blame. Similarly, ring doves can successfully retrieve a hidden reward, but fail to track its displacement within a container (Dumas and Wilkie, 1995). Given that parrots, corvids, and the great apes show similar performance whilst more closely related mammalian and avian species fail, it seems likely that stage 6 object permanence – the ability to understand invisible displacements–emerged convergently in the great apes and these two avian taxa, and represents a real quantitative difference in cognitive ability across species.

A similar convergence in capacity appears in the object transposition task. In this task, a reward is hidden under one of two cups, and their positions are changed. In children, the ability to solve the transposition task emerges later than the ability to solve invisible displacement tasks (Sophian and Sage, 1983; Sophian, 1984; Barth and Call, 2006), suggesting that this is a more challenging type of problem. A large number of mammals either fail to solve transposition tasks or may use associative strategies to guide their choices, including cats (Doré et al., 1996), dogs (Doré et al., 1996; Rooijakkers et al., 2009; Fiset and Plourde, 2013), wolves (Fiset and Plourde, 2013), wild boars (Albiach-Serrano et al., 2012), pigs (Albiach-Serrano et al., 2012), goats (Nawroth et al., 2015), dolphins (Jaakkola et al., 2010), and bears (Hartmann et al., 2017). Despite this selection of species including herbivores, omnivores, and carnivores, as well both domesticated and wild animals, only parrots (Pepperberg et al., 1997; Auersperg et al., 2014) and primates (Beran and Minahan, 2000; Call, 2001, 2003; Beran et al., 2005; Barth and Call, 2006; Rooijakkers et al., 2009) have been conclusively shown to succeed at object transposition tasks. Rather than relying on associative learning strategies, these two taxa appear able to represent the change to the objects' spatial locations.

Another example is the ability to reason through inference by exclusion (Güntürkün and Bugnyar, 2016). In tests of inference by exclusion, subjects must infer that one of two stimuli contains or is associated with a reward, after a demonstration that the other stimulus is not. Where two cups are used, for example, they must reason that if the reward is not hidden in the cup shown to be empty, then it must be in the other one. Several species of corvids

(Schloegl et al., 2009; Mikolasch et al., 2012; Shaw et al., 2013; Jelbert et al., 2015), parrots (Schloegl et al., 2009; Mikolasch et al., 2011; Pepperberg et al., 2013; O'Hara et al., 2015, 2016; Bastos and Taylor, 2019; Subias et al., 2019), and apes (Call, 2004, 2006; Hill et al., 2011) readily reason in this way. The ability to reason by exclusion is present in some New World monkeys (Sabbatini and Visalberghi, 2008; Marsh et al., 2015; Takahashi et al., 2015). Some capuchin monkeys are capable of this form of inference, whilst squirrel monkeys fail at both auditory and visual versions of the task (Marsh et al., 2015). This ability seems to be absent from other mammalian species including rats, golden hamsters, and tree shrews (Takahashi et al., 2015). This pattern suggests that the ability to reason through inference by exclusion varies quantitatively across species and has emerged convergently in the primate and avian lineages.

The ability to reason about use probabilistic information to make predictions about uncertain events also appears to have evolved convergently in the great apes and parrots (Rakoczy et al., 2014; Bastos and Taylor, 2020). When choosing between two hidden samples taken from two mixed populations of rewarding and unrewarding objects, capuchin monkeys appear to use a heuristic strategy of simply avoiding the sample from the population with the greatest absolute number of unrewarding objects (Tecwyn et al., 2017). On the other hand, both the great apes and the New Zealand mountain parrot, the kea, make their choices by relying on probabilistic information, by comparing the ratios of objects within and between the two populations (Rakoczy et al., 2014; Bastos and Taylor, 2020). This ability, known as true statistical inference, has so far not been conclusively shown outside of these two taxa, as other studies on primates and birds have not been able to exclude the absolute number heuristic as a potential strategy (Clements et al., 2018; Roberts et al., 2018; De Petrillo and Rosati, 2019; Placì et al., 2019).

Macphail went further than suggesting there are no quantitative differences in intelligence between species. He also suggested there are no qualitative differences in intelligence across species (Macphail, 1982, 1985; Macphail and Bolhuis, 2001). Research in this area has rather focused on whether there are types of thought that are unique to humans (Penn et al., 2008), rather than whether different animal species might think in qualitatively different ways. At present, therefore, it is not clear whether this hypothesis has been tested sufficiently to make conclusions either way. One route to testing this hypothesis further is focusing more on testing whether there are differences in the information processing biases, errors and limits of species showing similar levels of performance at different behavioral tasks (Taylor, 2014).

INTELLIGENT BEHAVIOR BEYOND PURE ASSOCIATION

Another of Macphail's claims is that all intelligent behavior can be explained by associative learning alone (Macphail, 1982, 1985; Macphail and Bolhuis, 2001). However, critics of Macphail have highlighted that a purely associative view of cognition is insufficient to explain the intelligent behaviors observed in vertebrates (Shettleworth, 1987), including birds. There are certainly areas of the literature on avian cognition where there is great debate as to whether the problem solving performances of birds can be explained by associative learning alone. For example, there is currently debate surrounding the role of associative learning and more complex cognition in research on planning in ravens (Redshaw et al., 2017; Lind, 2018; Dickerson et al., 2018; Hampton and Hampton, 2019), stone-dropping in corvids (Taylor and Gray, 2009; Cheke et al., 2011; Taylor et al., 2011; Jelbert et al., 2014; Logan et al., 2014; Ghirlanda and Lind, 2017; Hennefield et al., 2018, 2019), and string-pulling in a wide variety of birds (Taylor et al., 2010b, 2012; for a review of the species tested on string pulling, see Jacobs and Osvath, 2015). However, several lines of evidence indicate the presence of specific cognitive mechanisms other than associative learning in birds.

First, work on the innate cognitive capacities of birds has shown that prior experience is not required for complex problem solving to emerge. Without any prior experience, chicks can solve several problems in the physical realm, including detouring around a barrier by moving away from a desired object (Regolin et al., 1995), mentally representing the possible location of a hidden object when choosing between two different screens (Vallortigara et al., 1998; Chiandetti and Vallortigara, 2011), and recognizing partially-hidden objects by representing their complete outline (Regolin and Vallortigara, 1995; Regolin et al., 2004). Research on imprinted ducklings has also revealed an innate ability to distinguish between the abstract concepts of "same" and "different": when imprinted on two identical objects, ducklings preferred to approach pairs of identical objects rather than pairs of different objects, even though the objects in either case were different from those they were originally imprinted on (Martinho and Kacelnik, 2016). Given that these studies used inexperienced chicks and ducklings, this line of work strongly suggests that intelligence operates on more cognitive processes than associative learning alone. Work in chicks also offers further support for an innate approximate number system. Inexperienced chicks can distinguish between both small quantities from one to four (Rugani et al., 2013a) and larger quantities between five and ten (Rugani et al., 2013b). This capacity develops in birds into a numerical ability of surprising complexity. A seminal study in pigeons trained subjects to select images including one, two, or three shapes in ascending order, after which pigeons were asked to order sets with numerosities between one and nine (Scarf et al., 2011). Pigeons succeeded in this task despite never having received training on stimuli including between four and nine shapes, suggesting that they represent one through nine on an ordinal scale.

Work on the social cognition of birds has found clear evidence of birds performing beyond the predictions of associative learning. In a recent prosociality experiment, African gray parrots readily transferred tokens through a window to a conspecific who could exchange them for a food reward, when they could not exchange them themselves (Brucks and von Bayern, 2020). The study's control conditions suggest that this response was not driven by associative learning alone, as token transfers occurred significantly less often when their partner was unable to exchange tokens, or when the partner was absent. Similarly, an associative account would suggest that their tendency to transfer tokens would increase over time, but most subjects acted prosocially in their first trial. Caching studies provide evidence that birds can flexibly use information learnt in an egocentric manner to make allocentric predictions about the behavior of conspecifics in their environment. For example, Western scrub-jays, which pilfer other individual's caches, strategically relocate their caches (Emery and Clayton, 2001; Dally et al., 2005, 2006) in response to novel cues of a conspecific's presence, so as to reduce the likelihood of their caches being stolen in the future. In order to do this, individuals must have pilfered others' caches before, but need not have observed a pilfering event by another individual (Emery and Clayton, 2001), suggesting that they can project their own experience onto others. An associative learning explanation fails to acknowledge how they might shift between these egocentric and allocentric perspectives. A more recent study on ravens shows these birds will re-cache food when they believe they are being watched, and not as a learned response to a conspecific's gaze (Bugnyar et al., 2016). Ravens were similarly fast to cache when they heard sound recordings of a conspecific in a nearby compartment with a peephole, which could have granted the conspecific visual access to the cache, and when a conspecific was fully visible in the nearby compartment. In contrast, ravens cached slower and made more improvements to their caches in a control non-observed condition where they could hear a conspecific in a nearby compartment, but this conspecific was neither visible nor had a peephole available to look through. Ravens, therefore, appeared to flexibly use their egocentric experiences, in this case looking through a peephole at the caching chamber, to predict that another individual at the peephole would be able to see them caching.

Work on tool use in birds have produced a number of intriguing findings, suggesting that birds are capable of sophisticated technical intelligence (Weir et al., 2002; Taylor et al., 2007, 2010a; Tebbich et al., 2007; Bird and Emery, 2009; von Bayern et al., 2009, 2018; Wimpenny et al., 2009; Auersperg et al., 2010, 2011b, 2012b; Teschke and Tebbich, 2011; St Clair and Rutz, 2013; Laumer et al., 2016; Jelbert et al., 2018, 2019; Fayet et al., 2020). While some of these studies suggested that birds might be capable of mental trial and error during tool use, conclusive evidence that birds can mentally represent tool problems only emerged recently, from a study on New Caledonian crows (Gruber et al., 2019). This showed that these birds can pre-plan a sequence of behaviors up to three steps ahead, taking an available tool to the correct apparatus (the sub-goal) in order to retrieve another tool, which only then could be used to obtain a food reward (the overall goal). New Caledonian crows correctly planned and executed this sequence of behaviors even though all components of the sequence were out-of-sight of each other. This, therefore, required the crows to mentally represent the location and identity of the correct out-of-sight sub-goal and then use this representation to from a plan to solve the problem without error. Clear evidence of future-directed thought also comes from work on caching corvids. Western scrub-jays can anticipate their future needs, storing food that is unlikely to be available the following morning in a particular location (Alexis et al., 2007; Cheke and Clayton, 2011), regardless of their current satiation state (Correia et al., 2007). Evidence for the use of mental representations during tool manufacture has also emerged recently. After learning to insert a tool of a particular size into a vending machine, New Caledonian crows, when given a sheet of paper, were able manufacture tools of the correct size to insert into the machine (Jelbert et al., 2018). This was despite no tool template being available at the time of manufacture for use as a reference. Instead, the crows had to rely solely on their mental representation of the tool's size. Additionally, crows were not rewarded at test for making tools of the correct size. Instead, half of all tools made were rewarded irrespective of size, meaning there was no differential reinforcement for making the correct size tool at test.

Recently, work has begun to show that birds can solve problems that require domain-general intelligence, rather than problems involving domain-specific, ecologically relevant behaviors such as tool use and caching. Initial evidence that birds might have more domain-general cognitive processes comes from studies examining the ability of non-tool users to solve tool problems (Bird et al., 2009; Auersperg et al., 2010, 2011a,b, 2012b, 2016; Laumer et al., 2016). More recently, a study in kea showed that they can not only make accurate probabilistic comparisons between the two sampling events, as described above, but also integrate information across different domains (Bastos and Taylor, 2020). In one experiment, the two jars contained a physical barrier, and the otherwise identical populations of tokens were unevenly distributed above and below these barriers. Kea considered the physical constraint imposed by the barriers, adjusting their predictions of the likely sampling outcomes from the two jars. Another experiment in this study provided the kea with social information on sampling biases: one human demonstrators showed they had a preference for rewarding tokens by taking them from a jar rewarding tokens were in the minority, while the other demonstrated they were an unbiased blind sampler by taking rewarding tokens from a jar where such tokens were in the majority. When both these demonstrators sampled from jars with an even split of rewarding and unrewarding tokens, kea preferentially selected the samples from the biased demonstrator. These results showed that kea integrated either social or physical information into their probabilistic predictions, performing comparably to human infants (Téglás et al., 2007; Xu and Denison, 2009; Denison and Xu, 2010, 2014; Denison et al., 2012) and chimpanzees (Eckert et al., 2018a,b; Rakoczy et al., 2014), and outperforming monkeys (Tecwyn et al., 2017).

ECHOES OF MACPHAIL'S CRITICISMS IN THE 21ST CENTURY

Despite recent research not finding support for several of Macphail's claims, it is important to consider why Macphail may have raised these points in the first place, and why they are relevant today. The reasoning behind Macphail's null hypothesis for differences in intelligence across vertebrates highlights a flaw that has pervaded comparative psychology for many years: it is often impossible to tell why animals fail at a task. Differences in apparatus, methodology, motivation, and other contextual factors may affect species' performance in cognitive tasks. As highlighted by Macphail, failure at a task might be a true reflection of the species' ability, or the result may be caused by some contextual variable in that task. Researchers may attempt to resolve this in two ways: either by presenting an identical task across species, or by modifying some contextual variables in the task so it is better suited to a particular species. However, these two solutions are equally problematic.

When contextual variables are changed to suit a particular species, this makes it even more difficult to establish the reason for a species' failure at the task (Caldwell and Whiten, 2002; Schloegl et al., 2009; Liedtke et al., 2011; Auersperg et al., 2012a; Krasheninnikova et al., 2019; Farrar et al., 2020). Small changes in contextual variables may affect how a species interprets a task and therefore affect their performance, and this is likely to make it difficult to compare performances in a task across multiple species. One clear example of this comes from work on the trap-tube, a problem where an animal must extract food from a tube with a tool while avoiding a trap set into the lower surface of the tube. Apes' performance at this task changed dramatically once subjects were allowed to pull food with a tool toward them, rather than push food away (Mulcahy and Call, 2006), with learning speed increasing greatly and subjects passing the key "inverted-tube" control, where the tube was turned upside down, rendering the trap irrelevant. This example highlights how small changes to a task can affect animal performance greatly and offers a cautionary reminder of how hard it can be to interpret failure by a species at a cognitive task.

Even presenting an identical task to two very different species may lead to false positives, or false negatives, when the two species interpret the same task differently. This has been highlighted in a number of studies where animals failed at tasks involving a human demonstrator, but could have performed better had that contextual variable been changed (Erdőhegyi et al., 2007; Mikolasch et al., 2012; Shaw et al., 2013; Nawroth et al., 2014; Jelbert et al., 2016). Given that failure at a task does not necessarily represent a species' true cognitive abilities, negative results often become ambiguous and difficult to interpret, contributing to a "file-drawer effect" and publication bias (Fanelli and Fanelli, 2012; Farrar and Ostojić, 2019).

These issues in comparative psychology are highlighted in a landmark study by Maclean and colleagues (MacLean et al., 2014), which presented two identical tasks across thirty-six species to measure self-control: an A-not-B task, where a reward was visibly moved between two locations after being previously rewarded in only one of them, and a cylinder task, where an opaque cylinder containing food was presented and then substituted for a translucent cylinder. According to the authors, greater self-control should enable species to successfully switch search locations in the A-not-B task and avoid reaching directly for the food in the cylinder task. However, the study failed to consider how different species may perceive these tasks (Jelbert et al., 2016; Kabadayi et al., 2017): for example, that birds may perform poorly in the A-not-B task due to a poor innate understanding of human hands (Jelbert et al., 2016), rather than an inability to exert self-control. In support of this critique, New Caledonian crow performed poorly at this task without experience tracking hands, but after hand-tracking training actually performed comparably to the great apes in the same task (Jelbert et al., 2016).

Macphail's view suggests that errors such as these could be ruled out by exhaustively varying perceptual task features and other contextual variables to ensure that they are not responsible for subjects' failures, but in real terms this is often impossible (Kamil, 1987). One potential solution to this problem is to present pre-test baselines to different species (Jelbert et al., 2016). These baselines would comprise simple tasks that the animal would be expected to succeed at, given that the testing methods – or contextual variables – were appropriate. Success at such baselines could act as a checkpoint prior to test, ensuring that all species in the experiment understood the basic requirements of the test. Provided that a species succeeded at these baseline tasks, it would be possible to more confidently attribute failure at test to a lack of understanding, rather than other aspects of the task. For example, in the New Caledonian crow A-not-B study mentioned above, subjects first experienced hand-tracking training, watching the experimenter's hand bait a container among multiple other hand movements involving several cups and their lids (Jelbert et al., 2016). Had the subjects not first experienced this baseline training, it would not have been possible to determine if failure at the subsequent A-not-B task was due to a lack of experience with tracking human hands, or reflected a failure to inhibit a response to investigate a previously-rewarded container. Similarly, as highlighted earlier, the performance of various mammal species that pass visible displacement tasks, but fail invisible displacement tasks, provides stronger evidence for this failure being due to cognition rather than contextual variables, because the visible displacement task acts as a baseline test for the more complex invisible displacement task.

Another criticism of MacLean et al. (2014) is a lack of clarity on exactly which cognitive mechanisms were being tested (Beran, 2015; Brucks et al., 2017). It is unclear whether the self-control measures discussed in the study might reflect a single cognitive process, or a combination of several mechanisms (Beran, 2015). Self-control has been used as a term to describe either the ability to incur a cost in order to obtain a more valuable reward instead of a less-costly, lower-value reward (Beran, 2015), or the ability to inhibit a response (MacLean et al., 2014). These two abilities are not necessarily underpinned by the same cognitive process. Work on dogs shows that even inhibition alone is not consistent across different tasks, suggesting that different tests of the same ability are not actually tapping into the same cognitive mechanisms (Brucks et al., 2017). Similarly, a recent study in pheasant chicks shows that comparisons across multiple tasks might not accurately reflect cognitive ability (van Horik and Madden, 2016). In the pheasant study, two hundred chicks experienced three foraging tasks, meant to assess whether individual variation in performance was robust and driven by real differences in cognitive ability. The study failed to find any consistent differences in problem solving ability between individuals across the three tasks, suggesting that motivation, and not cognitive capacity, was the main driver for these differences.

One way to help move past these issues would be to focus more on exploring how animals succeed at some tasks, and how they fail at others, rather than whether they simply fail or succeed at certain problems. The signature-testing approach, and research focused on cognitive processes rather than task performance, are a viable strategy for this (Kacelnik, 2009; Taylor and Gray, 2009; Seed et al., 2012; Carruthers and Fletcher, 2013; Taylor, 2014). A process-driven approach allows researchers to generate specific hypotheses about which errors, biases, limits, and specific patterns of performance identify particular cognitive mechanisms, and design experimental tasks that tease these potential processes apart. This is analogous to the strong inference approach (Platt, 1964), which aims to successively rule out alternative hypotheses through the design of experiments that specifically test these hypotheses with clear predicted outcomes for each alternative explanation. Researchers can triangulate several of these signatures within or between tasks (Heyes, 1993; Taylor, 2014), to pinpoint exactly which of several qualitative forms of intelligence different species are capable of. This approach could provide a more powerful system through which we can better address Macphail's null hypothesis, particularly in terms of qualitative differences in intelligence. Several of the studies discussed earlier provide clear examples of behavioral signatures that constrain the possible cognitive mechanisms an animal can be using to solve a problem. For example, the presence of a distance effect bias in pigeons' numerical discriminations, where pigeons are more accurate and quicker to discriminate number pairs when the numerical distance between them is greater, provides a clear behavioral signature that numbers are represented on an ordinal scale (Scarf et al., 2011). Similarly, the ability of crows to solve problems without mistake when downstream aspects of the problem are out-of-sight, shows they are not limited by having problems out-of-sight and so provides a clear signature for pre-planning, as decisions have to be made using mental representations of the problem (Gruber et al., 2019).

Finally, a Bayesian framework may provide useful tools in interpreting animal performances from a statistical viewpoint. Given that research questions and methods are appropriately framed, the Bayesian framework can distinguish between a lack of power in the data, and direct support for the null hypothesis (Wagenmakers, 2007; Stevens, 2017; Wagenmakers et al., 2018). In the frequentist framework, these two forms of non-significance are often confounded. This leaves researchers with inconclusive data which often ends up unpublished (the file-drawer effect; Fanelli and Fanelli, 2012; Farrar et al., 2020). In contrast, Bayesian analyses can be much more informative than their frequentist counterparts when animals fail at an experimental task. The Bayesian framework enables researchers to provide claims both for and against the existence of particular cognitive capacities in their target species, rather than it being unclear whether negatives are due to low sample size or a true failure at a task. While clearly this framework does not resolve all of the issues surrounding the interpretation of 'evidence of absence' in comparative psychology, it does offer a route toward bringing more quantitative and qualitative differences in intelligence to light in the literature (Stevens, 2017).

DISCUSSION

Macphail's support of a null hypothesis for no quantitative differences in intelligence across species, and his claim that all intelligent behavior is association-based, fall short in the light of recent research in avian cognition. Avian cognition provides clear evidence for robust differences in intelligence among avian species, as well as between birds and other taxa, and for problem solving that extends beyond simple associative learning.

However, Macphail's criticisms of comparative psychology are relevant to this day and can inspire researchers to make more advances in this field. Thirty-five years ago, Macphail highlighted the difficulty in establishing whether animals fail at a task because they cannot understand it, or because their performance was affected by variations in methodology (Macphail, 1982, 1985; Macphail and Bolhuis, 2001). Today, much of the field still grapples with this distinction. Researchers often cannot tell why subjects fail at some tasks but not others, and comparative psychology suffers from widespread publication bias (Farrar et al., 2020).

Macphail also highlighted that some of the preconceptions of the field at the time but had not been appropriately tested. One of these was the belief that intelligence varies predictably across species, with humans showing the greatest intelligence, followed by their closest relatives (Jensen, 1980). According to this view, one might expect an inverse correlation in intelligence with evolutionary distance from humans and other primates. Macphail argued that such a *scala naturae* assumption might be erroneous (Macphail, 1985), so helping move the field past this early anthropocentric attitude and toward the present day, where researchers focus on testing intelligence across a phylogenetically broad array of animal species, albeit often still with tests that have been designed for human intelligence.

In sum, while some of Macphail's claims do not hold up to the current body of evidence, a number of his criticisms of the field of comparative psychology still hold in the present day. We suggest three strategies researchers can use to combat these issues: (i) using baseline tasks to ensure that contextual variables cannot explain subjects' failure at test (Jelbert et al., 2016); (ii) focusing on a signature-testing, process-driven approach, that specifically seeks to pinpoint the cognitive mechanisms that animals rely on to solve problems (Kacelnik, 2009; Taylor and Gray, 2009; Seed et al., 2012; Carruthers and Fletcher, 2013; Taylor, 2014); and (iii) taking advantage of the Bayesian framework to distinguish between support for the null hypothesis and a lack of statistical power (Wagenmakers, 2007; Stevens, 2017; Wagenmakers et al., 2018). Put together, these three strategies can help researchers identify both quantitative and qualitative differences in intelligence between species, learn from animals' successes as well as their failures, and triangulate evidence for complex cognition that is not rooted exclusively in associative learning.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was made possible through a Rutherford Discovery Fellowship (AT).

REFERENCES

Albiach-Serrano, A., Bräuer, J., Cacchione, T., Zickert, N., and Amici, F. (2012). The effect of domestication and ontogeny in swine cognition (*Sus scrofa scrofa* and S. s. domestica). *Appl. Anim. Behav. Sci.* 141, 25–35. doi: 10.1016/j.applanim.2012.07.005

Alexis, D. M., Raby, C. R., Clayton, N. S., and Dickinson, A. (2007). Planning for the future by western scrub-jays. *Nature* 445, 919–921. doi: 10.1038/nature05575

Auersperg, A. M. I., Borasinski, S., Laumer, I., and Kacelnik, A. (2016). Goffin's cockatoos make the same tool type from different materials. *Biol. Lett.* 12:20160689. doi: 10.1098/rsbl.2016.0689

Auersperg, A. M. I., Gajdon, G. K., and Huber, L. (2010). Kea, *Nestor notabilis*, produce dynamic relationships between objects in a second-order tool use task. *Anim. Behav.* 80, 783–789. doi: 10.1016/j.anbehav.2010.08.007

Auersperg, A. M. I., Gajdon, G. K., and von Bayern, A. M. P. (2012a). A new approach to comparing problem solving, flexibility and innovation. *Commun. Integ. Biol.* 5, 140–145. doi: 10.4161/cib.18787

Auersperg, A. M. I., Szabo, B., von Bayern, A. M. P., and Kacelnik, A. (2012b). Spontaneous innovation in tool manufacture and use in a Goffin's cockatoo. *Curr. Biol.* 22, R903–R904. doi: 10.1016/j.cub.2012.09.002

Auersperg, A. M. I., Huber, L., and Gajdon, G. K. (2011a). Navigating a tool end in a specific direction: stick-tool use in kea (*Nestor notabilis*). *Biol. Lett.* 7, 825–828. doi: 10.1098/rsbl.2011.0388

Auersperg, A. M. I., von Bayern Auguste, M. P., Gajdon, G. K., Huber, L., and Kacelnik, A. (2011b). Flexibility in problem solving and tool use of kea and new caledonian crows in a multi access box paradigm. *PLoS One* 6:e20231. doi: 10.1371/journal.pone.0020231

Auersperg, A. M. I., Szabo, B., von Bayern Auguste, M. P., and Bugnyar, T. (2014). Object permanence in the Goffin cockatoo (*Cacatua goffini*). *J. Comp. Psychol.* 128, 88–98. doi: 10.1037/a0033272

Auersperg, A. M. I., and von Bayern, A. M. P. (2019). Who's a clever bird - now? A brief history of parrot cognition. *Behaviour* 156, 391–407. doi: 10.1163/1568539X-00003550

Barth, J., and Call, J. (2006). Tracking the displacement of objects: a series of tasks with great apes (*Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, and *Pongo pygmaeus*) and young children (*Homo sapiens*). *J. Exp. Psychol. Anim. Behav. Process.* 32, 239–252. doi: 10.1037/0097-7403.32.3.239

Bastos, A. P. M., and Taylor, A. H. (2019). Kea (*Nestor notabilis*) represent object trajectory and identity. *Sci. Rep.* 9:19759. doi: 10.1038/s41598-019-56380-4

Bastos, A. P. M., and Taylor, A. H. (2020). Kea show three signatures of domain-general statistical inference. *Nat. Commun.* 11:828.

Beran, M., Beran, M., and Menzel, C. (2005). Chimpanzees (*Pan troglodytes*) use markers to monitor the movement of a hidden item. *Primates* 46, 255–259. doi: 10.1007/s10329-005-0150-z

Beran, M. J. (2015). The comparative science of "self-control": what are we talking about? *Front. Psychol.* 6:51. doi: 10.3389/fpsyg.2015.00051

Beran, M. J., and Minahan, M. F. (2000). Monitoring spatial transpositions by Bonobos (*Pan paniscus*) and Chimpanzees (*P. troglodytes*). *Int. J. Comp. Psychol.* 13, 1–15.

Bird, C. D., and Emery, N. J. (2009). Rooks use stones to raise the water level to reach a floating worm. *Curr. Biol.* 19, 1410–1414. doi: 10.1016/j.cub.2009.07.033

Bird, C. D., Emery, N. J., and Gallistel, C. R. (2009). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10370–10375. doi: 10.1073/pnas.0901008106

Bond, A. B., Kamil, A. C., and Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Anim. Behav.* 65, 479–487. doi: 10.1006/anbe.2003.2101

Bond, A. B., Kamil, A. C., and Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*). *J. Comp. Psychol.* 121, 372–379. doi: 10.1037/0735-7036.121.4.372

Boucherie, P. H., Loretto, M. C., Massen, J. J. M., and Bugnyar, T. (2019). What constitutes "social complexity" and "social intelligence" in birds?: lessons from ravens. *Behav. Ecol. Sociobiol.* 73, 1–14. doi: 10.1007/s00265-018-2607-2

Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., and Range, F. (2017). Measures of dogs' inhibitory control abilities do not correlate across tasks. *Front. Psychol.* 8:849. doi: 10.3389/fpsyg.2017.00849

Brucks, D., and von Bayern, A. M. P. (2020). Parrots voluntarily help each other to obtain food rewards. *Curr. Biol.* 30, 292.e5–297.e5. doi: 10.1016/j.cub.2019.11.030

Bugnyar, T., Reber, S. A., and Buckner, C. (2016). Ravens attribute visual access to unseen competitors. *Nat. Commun.* 7:10506. doi: 10.1038/ncomms10506

Caldwell, C., and Whiten, A. (2002). Evolutionary perspectives on imitation: is a comparative psychology of social learning possible? *Anim. Cogn.* 5, 193–208. doi: 10.1007/s10071-002-0151-x

Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *J. Comp. Psychol.* 115, 159–171. doi: 10.1037/0735-7036.115.2.159

Call, J. (2003). Spatial rotations and transpositions in orangutans (*Pongo pygmaeus*) and chimpanzees (*Pan troglodytes*). *Primates* 44, 347–357. doi: 10.1007/s10329-003-0048-6

Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, and *Pongo pygmaeus*). *J. Comp. Psychol.* 118, 232–241. doi: 10.1037/0735-7036.118.2.232

Call, J. (2006). Inferences by exclusion in the great apes: the effect of age and species. *Anim. Cogn.* 9, 393–403. doi: 10.1007/s10071-006-0037-4

Carruthers, P., and Fletcher, L. (2013). "Behavior-reading versus mentalizing in animals," in *Anonymous Agency and Joint Attention* eds J. Metcafe, H. Terrace (New York, NY: Oxford University Press).

Cheke, L., Bird, C., and Clayton, N. (2011). Tool-use and instrumental learning in the Eurasian jay (*Garrulus glandarius*). *Anim. Cogn.* 14, 441–455. doi: 10.1007/s10071-011-0379-4

Cheke, L. G., and Clayton, N. S. (2011). Eurasian jays (*Garrulus glandarius*) overcome their current desires to anticipate two distinct future needs and plan for them appropriately. *Biol. Lett.* 8, 171–175. doi: 10.1098/rsbl.2011.0909

Chiandetti, C., and Vallortigara, G. (2011). Intuitive physical reasoning about occluded objects by inexperienced chicks. *Proc. Biol. Sci.* 278, 2621–2627. doi: 10.1098/rspb.2010.2381

Clayton, N. (2012). Corvid cognition: feathered apes. *Nature* 484, 453–454. doi: 10.1038/484453a

Clements, K. A., Gray, S. L., Gross, B., and Pepperberg, I. M. (2018). Initial evidence for probabilistic reasoning in a grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 132, 166–177. doi: 10.1037/com0000106

Collier-Baker, E., Davis, J., Nielsen, M., and Suddendorf, T. (2006). Do chimpanzees (*Pan troglodytes*) understand single invisible displacement? *Anim. Cogn.* 9, 55–61. doi: 10.1007/s10071-005-0004-5

Collier-Baker, E., and Suddendorf, T. (2006). Do chimpanzees (*Pan troglodytes*) and 2-year-old children (*Homo sapiens*) understand double invisible displacement? *J. Comp. Psychol.* 120, 89–97. doi: 10.1037/0735-7036.120.2.89

Correia, S. P. C., Dickinson, A., and Clayton, N. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Curr. Biol.* 17, 856–861. doi: 10.1016/j.cub.2007.03.063

Dally, J. M., Emery, N. J., and Clayton, N. S. (2005). Cache protection strategies by western scrub-jays, *Aphelocoma californica*: implications for social cognition. *Anim. Behav.* 70, 1251–1263. doi: 10.1016/j.anbehav.2005.02.009

Dally, J. M., Emery, N. J., and Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science* 312, 1662–1665. doi: 10.1126/science.1126539

de Blois, S. T., Novak, M. A., and Bond, M. (1998). Object permanence in orangutans (*Pongo pygmaeus*) and squirrel monkeys (*Saimiri sciureus*). *J. Comp. Psychol.* 112, 137–152. doi: 10.1037/0735-7036.112.2.137

De Petrillo, F., and Rosati, A. G. (2019). Rhesus macaques use probabilities to predict future events. *Evol. Hum. Behav.* 40, 436–446. doi: 10.1016/j.evolhumbehav.2019.05.006

Denison, S., Reed, C., and Xu, F. (2012). The emergence of probabilistic reasoning in very young infants: evidence from 4.5- and 6-month-olds. *Dev. Psychol.* 49, 243–249. doi: 10.1037/a0028278

Denison, S., and Xu, F. (2010). Integrating physical constraints in statistical inference by 11-month-old infants. *Cogn. Sci.* 34, 885–908. doi: 10.1111/j.1551-6709.2010.01111.x

Denison, S., and Xu, F. (2014). The origins of probabilistic inference in human infants. *Cognition* 130, 335–347. doi: 10.1016/j.cognition.2013.12.001

Deppe, A., Deppe, A., Wright, P., Wright, P., Szelistowski, W., and Szelistowski, W. (2009). Object permanence in lemurs. *Anim. Cogn.* 12, 381–388. doi: 10.1007/s10071-008-0197-5

Dickerson, K. L., Ainge, J. A., and Seed, A. M. (2018). The role of association in pre-schoolers' solutions to "spoon tests" of future planning. *Curr. Biol.* 28, 2309.e2–2313.e2. doi: 10.1016/j.cub.2018.05.052

Doré, F. Y., Fiset, S., Goulet, S., Dumas, M., and Gagnon, S. (1996). Search behavior in cats and dogs: interspecific differences in working memory and spatial cognition. *Anim. Learn. Behav.* 24, 142–149. doi: 10.3758/BF03198962

Dumas, C., and Wilkie, D. M. (1995). Object permanence in ring doves (*Streptopelia risoria*). *J. Comp. Psychol.* 109, 142–150. doi: 10.1037/0735-7036.109.2.142

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evol. Anthropol. Issues News Rev.* 6, 178–190. doi: 10.1002/(SICI)1520-650519986:53.0.CO

Dunbar, R. I. M. (2008). Cognitive constraints on the structure and dynamics of social networks. *Group Dyn. Theory Res. Pract.* 12, 7–16. doi: 10.1037/1089-2699.12.1.7

Eckert, J., Call, J., Hermes, J., Herrmann, E., and Rakoczy, H. (2018a). Intuitive statistical inferences in chimpanzees and humans follow Weber's law. *Cognition* 180, 99–107. doi: 10.1016/j.cognition.2018.07.004

Eckert, J., Rakoczy, H., Call, J., Herrmann, E., and Hanus, D. (2018b). Chimpanzees consider humans' psychological states when drawing statistical inferences. *Curr. Biol.* 28, 1959.e3–1963.e3. doi: 10.1016/j.cub.2018.04.077

Elepfandt, A. (1987). Comparative cognition: inadequate approach, precipitate conclusions. *Behav. Brain Sci.* 10, 661–662. doi: 10.1017/S0140525X00055059

Emery, N. J. (2004). "Are corvids 'feathered apes'? Cognitive evolution in crows, jays, rooks, and jackdaws," in *Comparative Analysis of Minds*, ed. S. Watanabe (Tokyo: Keio University Press), 181–213.

Emery, N. J. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philos. Trans. R. Soc. B Biol. Sci.* 361, 23–43. doi: 10.1098/rstb.2005.1736

Emery, N. J., and Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414, 443–446. doi: 10.1038/35106560

Emery, N. J., and Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306, 1903–1907. doi: 10.1126/science.1098410

Emery, N. J., Van Horik, J. O., Topál, J., Maestripieri, D., Hopper, L. M., Raby, C. R., et al. (2012). "Convergent evolution of cognition in corvids, apes and other animals," in *Oxford Library of Psychology. The Oxford Handbook of Comparative Evolutionary Psychology* J. Vonk & T. K. Shackelford (Oxford: Oxford University Press). doi: 10.1126/science.1098410

Erdőhegyi, Á, Topál, J., Virányi, Z., and Miklósi, Á (2007). Dog-logic: inferential reasoning in a two-way choice task and its restricted use. *Anim. Behav.* 74, 725–737. doi: 10.1016/j.anbehav.2007.03.004

Fanelli, D., and Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics* 90, 891–904. doi: 10.1007/s11192-011-0494-7

Farrar, B., Boeckle, M., and Clayton, N. (2020). Replications in comparative cognition: what should we expect and how can we improve? *Anim. Behav. Cogn.* 7, 1–22. doi: 10.26451/abc.07.01.02.2020

Farrar, B. G., and Ostojić, L. (2019). The illusion of science in comparative cognition. *PsyArXiv* doi: 10.31234/osf.io/hduyx

Fayet, A. L., Hansen, E. S., and Biro, D. (2020). Evidence of tool use in a seabird. *Proc. Natl. Acad. Sci. U.S.A.* 117, 1277–1279. doi: 10.1073/pnas.1918060117

Fiset, S., and Plourde, V. (2013). Object permanence in domestic dogs (*Canis lupus familiaris*) and gray wolves (*Canis lupus*). *J. Comp. Psychol.* 127, 115–127. doi: 10.1037/a0030595

Ghirlanda, S., and Lind, J. (2017). 'Aesop's fable' experiments demonstrate trial-and-error learning in birds, but no causal understanding. *Anim. Behav.* 123, 239–247. doi: 10.1016/j.anbehav.2016.10.029

Gruber, R., Schiestl, M., Boeckle, M., Frohnwieser, A., Miller, R., Gray, R. D., et al. (2019). New caledonian crows use mental representations to solve metatool problems. *Curr. Biol.* 29, 686.e3–692.e3. doi: 10.1016/j.cub.2019.01.008

Güntürkün, O., and Bugnyar, T. (2016). Cognition without cortex. *Trends Cogn. Sci.* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Hampton, R., and Hampton, R. (2019). Parallel overinterpretation of behavior of apes and corvids. *Learn. Behav.* 47, 105–106. doi: 10.3758/s13420-018-0330-5

Hartmann, D., Davila-Ross, M., Wong, S. T., Call, J., and Scheumann, M. (2017). Spatial transposition tasks in Indian sloth bears (*Melursus ursinus*) and Bornean sun bears (*Helarctos malayanus euryspilus*). *J. Comp. Psychol.* 131, 290–303. doi: 10.1037/com0000077

Hennefield, L., Hwang, H. G., and Povinelli, D. J. (2019). Going meta: retelling the scientific retelling of aesop's the crow and the pitcher. *J. Folklore Res.* 56, 45–70. doi: 10.2979/jfolkrese.56.2_3.04

Hennefield, L., Hwang, H. G., Weston, S. J., and Povinelli, D. J. (2018). Meta-analytic techniques reveal that corvid causal reasoning in the Aesop's Fable paradigm is driven by trial-and-error learning. *Anim. Cogn.* 21, 735–748. doi: 10.1007/s10071-018-1206-y

Heyes, C. M. (1993). Anecdotes, training, trapping and triangulating: do animals attribute mental states? *Anim. Behav.* 46, 177–188. doi: 10.1006/anbe.1993.1173

Hill, A., Collier-Baker, E., and Suddendorf, T. (2011). Inferential reasoning by exclusion in great apes, lesser apes, and spider monkeys. *J. Comp. Psychol.* 125, 91–103. doi: 10.1037/a0020867

Hoffmann, A., Rüttler, V., and Nieder, A. (2011). Ontogeny of object permanence and object tracking in the carrion crow, *Corvus corone*. *Anim. Behav.* 82, 359–367. doi: 10.1016/j.anbehav.2011.05.012

Jaakkola, K. (2014). Do animals understand invisible displacement? A critical review. *J. Comp. Psychol.* 128, 225–239. doi: 10.1037/a0035675

Jaakkola, K., Guarino, E., Rodriguez, M., Erb, L., and Trone, M. (2010). What do dolphins (*Tursiops truncatus*) understand about hidden objects? *Anim. Cogn.* 13, 103–120. doi: 10.1007/s10071-009-0250-z

Jacobs, I. F., and Osvath, M. (2015). The string-pulling paradigm in comparative psychology. *J. Comp. Psychol.* 129, 89–120. doi: 10.1037/a0038746

Jelbert, S. A., Hosking, R. J., Taylor, A. H., and Gray, R. D. (2018). Mental template matching is a potential cultural transmission mechanism for New Caledonian crow tool manufacturing traditions. *Sci. Rep.* 8, 8956–8958. doi: 10.1038/s41598-018-27405-1

Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., et al. (2019). New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proc. Biol. Sci.* 286, 20182332. doi: 10.1098/rspb.2018.2332

Jelbert, S. A., Taylor, A. H., Cheke, L. G., Clayton, N. S., and Gray, R. D. (2014). Using the Aesop's fable paradigm to investigate causal understanding of water displacement by new caledonian crows. *PLoS One* 9:e92895. doi: 10.1371/journal.pone.0092895

Jelbert, S. A., Taylor, A. H., and Gray, R. D. (2015). Reasoning by exclusion in new caledonian crows (*Corvus moneduloides*) cannot be explained by avoidance of empty containers. *J. Comp. Psychol.* 129, 283–290. doi: 10.1037/a0039313

Jelbert, S. A., Taylor, A. H., and Gray, R. D. (2016). Does absolute brain size really predict self-control? Hand-tracking training improves performance on the A-not-B task. *Biol. Lett.* 12:20150871. doi: 10.1098/rsbl.2015.0871

Jensen, A. R. (1980). *Bias in Mental Testing*. London: Metheun.

Kabadayi, C., Krasheninnikova, A., O'Neill, L., van de Weijer, J., Osvath, M., and von Bayern, A. (2017). Are parrots poor at motor self-regulation or is the cylinder task poor at measuring it? *Anim. Cogn.* 20, 1137–1146. doi: 10.1007/s10071-017-1131-5

Kacelnik, A. (2009). Tools for thought or thoughts for tools? *Proc. Natl. Acad. Sci. U.S.A.* 106, 10071–10072. doi: 10.1073/pnas.0904735106

Kamil, A. C. (1987). A synthetic approach to the study of animal intelligence. *Nebraska Symp. Motiv.* 35, 257–308.

Krasheninnikova, A., Berardi, R., Lind, M., O'Neill, L., and von Bayern, A. M. P. (2019). Primate cognition test battery in parrots. *Behaviour* 2019, 721–761. doi: 10.1163/1568539X-0003549

Krasheninnikova, A., Bräger, S., and Wanker, R. (2013). Means–end comprehension in four parrot species: explained by social complexity. *Anim. Cogn.* 16, 755–764. doi: 10.1007/s10071-013-0609-z

Laumer, I. B., Bugnyar, T., and Auersperg, A. M. I. (2016). Flexible decision-making relative to reward quality and tool functionality in Goffin cockatoos (*Cacatua goffiniana*). *Sci. Rep.* 6:28380. doi: 10.1038/srep28380

Liedtke, J., Liedtke, J., Werdenich, D., Werdenich, D., Gajdon, G., Gajdon, G., et al. (2011). Big brains are not enough: performance of three parrot species in the trap-tube paradigm. *Anim. Cogn.* 14, 143–149. doi: 10.1007/s10071-010-0347-4

Lind, J. (2018). What can associative learning do for planning? *R. Soc. Open Sci.* 5:180778. doi: 10.1098/rsos.180778

Logan, C. J., Jelbert, S. A., Breen, A. J., Gray, R. D., and Taylor, A. H. (2014). Modifications to the aesop's fable paradigm change new caledonian crow performances. *PLoS One* 9:e103049. doi: 10.1371/journal.pone.0103049

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., et al. (2014). The evolution of self-control. *Proc. Natl. Acad. Sci.* 111, E2140–E2148. doi: 10.1073/pnas.1323533111

Macphail, E. M. (1982). *Brain and Intelligence in Vertebrates*. Oxford: Clarendon Press.

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 37–51. doi: 10.1098/rstb.1985.0008

Macphail, E. M., and Bolhuis, J. J. (2001). The evolution of intelligence: adaptive specializations versus general process. *Biol. Rev.* 76, 341–364. doi: 10.1017/S146479310100570X

Mallavarapu, S. (2009). *Object Permanence in Orangutans, Gorillas, and Black-and-White Ruffed Lemurs*. Ph.D. thesis, retrieved from ProQuest Dissertations and Theses database (3376320).

Marsh, H. L., Vining, A. Q., Levendoski, E. K., and Judge, P. G. (2015). Inference by exclusion in lion-tailed macaques (*Macaca silenus*), a hamadryas baboon (*Papio hamadryas*), capuchins (*Sapajus apella*), and squirrel monkeys (*Saimiri sciureus*). *J. Comp. Psychol.* 129, 256–267. doi: 10.1037/a0039316

Martinho, A. III, and Kacelnik, A. (2016). Ducklings imprint on the relational concept of "same or different". *Science* 353, 286–288. doi: 10.1126/science.aaf4247

Mikolasch, S., Kotrschal, K., and Schloegl, C. (2011). en food. *Biol. Lett.* 7, 875–877. doi: 10.1098/rsbl.2011.0500

Mikolasch, S., Mikolasch, S., Kotrschal, K., Kotrschal, K., Schloegl, C., and Schloegl, C. (2012). Is caching the key to exclusion in corvids? The case of carrion crows (*Corvus corone corone*). *Anim. Cogn.* 15, 73–82. doi: 10.1007/s10071-011-0434-1

Mulcahy, N. J., and Call, J. (2006). How great apes perform on a modified trap-tube task. *Anim. Cogn.* 9, 193–199. doi: 10.1007/s10071-006-0019-6

Nawroth, C., von Borell, E., and Langbein, J. (2014). Exclusion performance in dwarf goats (*Capra aegagrus hircus*) and sheep (*Ovis orientalis aries*). *PLoS One* 9:e93534. doi: 10.1371/journal.pone.0093534

Nawroth, C., von Borell, E., and Langbein, J. (2015). Object permanence in the dwarf goat (*Capra aegagrus hircus*): perseveration errors and the tracking of complex movements of hidden objects. *Appl. Anim. Behav. Sci.* 167, 20–26. doi: 10.1016/j.applanim.2015.03.010

O'Hara, M., Auersperg, A. M. I., Bugnyar, T., and Huber, L. (2015). Inference by exclusion in goffin cockatoos (*Cacatua goffini*). *PLoS One* 10:e0134894. doi: 10.1371/journal.pone.0134894

O'Hara, M., Schwing, R., Federspiel, I., Gajdon, G. K., and Huber, L. (2016). Reasoning by exclusion in the kea (*Nestor notabilis*). *Anim. Cogn.* 19, 965–975. doi: 10.1007/s10071-016-0998-x

Olson, D. J., Kamil, A. C., Balda, R. P., and Nims, P. J. (1995). Performance of four seed-caching corvid species in operant tests of nonspatial and spatial memory. *J. Comp. Psychol.* 109, 173–181. doi: 10.1037/0735-7036.109.2.173

Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 109–130. doi: 10.1017/S0140525X08003543

Pepperberg, I. M., and Funk, M. S. (1990). Object permanence in four species of psittacine birds: an African Grey parrot (*Psittcaus erithacus*), an Illiger mini macaw (*Ara maracana*), a parakeet (*Melopsittacus undulatus*), and a cockatiel (*Nymphicus hollandicus*). *Learn. Behav.* 18, 97–108. doi: 10.3758/bf03205244

Pepperberg, I. M., Koepke, A., Livingston, P., Girard, M., and Hartsfield, L. A. (2013). Reasoning by inference: further studies on exclusion in grey parrots (*Psittacus erithacus*). *J. Comp. Psychol.* 127, 272–281. doi: 10.1037/a0031641

Pepperberg, I. M., Willner, M. R., and Gravitz, L. B. (1997). Development of piagetian object permanence in a grey parrot (*Psittacus erithacus*). *J. Comp. Psychol.* 111, 63–75. doi: 10.1037/0735-7036.111.1.63

Piaget, J. (1954). *The Construction of Reality in the Child*. New York, NY: Basic Books.

Placì, S., Padberg, M., Rakoczy, H., and Fischer, J. (2019). Long-tailed macaques extract statistical information from repeated types of events to make rational decisions under uncertainty. *Sci. Rep.* 9, 12107–12112. doi: 10.1038/s41598-019-48543-0

Platt, J. R. (1964). Strong inference. *Science* 146, 347–353. doi: 10.1126/science.146.3642.347

Pollok, B., Prior, H., and Güntürkün, O. (2000). Development of object permanence in food-storing magpies (*Pica pica*). *J. Comp. Psychol.* 114, 148–157. doi: 10.1037/0735-7036.114.2.148

Rakoczy, H., Clüver, A., Saucke, L., Stoffregen, N., Gräbener, A., Migura, J., et al. (2014). Apes are intuitive statisticians. *Cognition* 131, 60–68. doi: 10.1016/j.cognition.2013.12.011

Redshaw, J., Taylor, A. H., and Suddendorf, T. (2017). Flexible planning in ravens? *Trends Cogn. Sci.* 21, 821–822. doi: 10.1016/j.tics.2017.09.001

Regolin, L., Marconato, F., and Vallortigara, G. (2004). Hemispheric differences in the recognition of partly occluded objects by newly hatched domestic chicks (*Gallus gallus*). *Anim. Cogn.* 7, 162–170. doi: 10.1007/s10071-004-0208-0

Regolin, L., and Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Percept. Psychophys.* 57, 971–976. doi: 10.3758/BF03205456

Regolin, L., Vallortigara, G., and Zanforlin, M. (1995). Object and spatial representations in detour problems by chicks. *Anim. Behav.* 49, 195–199. doi: 10.1016/0003-3472(95)80167-7

Roberts, W., MacDonald, H., and Lo, K. (2018). Pigeons play the percentages: computation of probability in a bird. *Anim. Cogn.* 21, 575–581. doi: 10.1007/s10071-018-1192-0

Rooijakkers, E. F., Kaminski, J., and Call, J. (2009). Comparing dogs and great apes in their ability to visually track object transpositions. *Anim. Cogn.* 12, 789–796. doi: 10.1007/s10071-009-0238-8

Rugani, R., Cavazzana, A., Vallortigara, G., and Regolin, L. (2013a). One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. *Anim. Cogn.* 16, 557–564. doi: 10.1007/s10071-012-0593-8

Rugani, R., Vallortigara, G., and Regolin, L. (2013b). Numerical abstraction in young domestic chicks (*Gallus gallus*). *PLoS One* 8:e65262. doi: 10.1371/journal.pone.0065262

Sabbatini, G., and Visalberghi, E. (2008). Inferences About the location of food in capuchin monkeys (*Cebus apella*) in two sensory modalities. *J. Comp. Psychol.* 122, 156–166. doi: 10.1037/0735-7036.122.2.156

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334:1664. doi: 10.1126/science.1213357

Schloegl, C., Dierks, A., Gajdon, G. K., Huber, L., Kotrschal, K., and Bugnyar, T. (2009). What you see is what you get? Exclusion performances in ravens and keas. *PLoS One* 4:e6368. doi: 10.1371/journal.pone.0006368

Seed, A., Emery, N., and Clayton, N. (2009). Intelligence in corvids and apes: a case of convergent evolution? *Ethology* 115, 401–420. doi: 10.1111/j.1439-0310.2009.01644.x

Seed, A., Seddon, E., Greene, B., and Call, J. (2012). Chimpanzee 'folk physics': bringing failures into focus. *Philos. Trans. Biol. Sci.* 367, 2743–2752. doi: 10.1098/rstb.2012.0222

Shaw, R. C., Plotnik, J. M., and Clayton, N. S. (2013). Exclusion in corvids: the performance of food-caching Eurasian jays (*Garrulus glandarius*). *J. Comp. Psychol.* 127, 428–435. doi: 10.1037/a0032010

Shettleworth, S. J. (1987). Intelligence: more than a matter of associations. *Behav. Brain Sci.* 10:679. doi: 10.1017/S0140525X00055266

Shultz, S., and Dunbar, R. I. M. (2010). Species differences in executive function correlate with hippocampus volume and neocortex ratio across nonhuman primates. *J. Comp. Psychol.* 124, 252–260. doi: 10.1037/a0018894

Sophian, C. (1984). Spatial transpositions and the early development of search. *Dev. Psychol.* 20, 21–28. doi: 10.1037/0012-1649.20.1.21

Sophian, C., and Sage, S. (1983). Developments in infants' search for displaced objects. *J. Exp. Child Psychol.* 35, 143–160. doi: 10.1016/0022-0965(83)90075-9

St Clair, J. J. H., and Rutz, C. (2013). New Caledonian crows attend to multiple functional properties of complex tools. *Philos. Trans. R. Soc. B Biol. Sci.* 368, 20120415. doi: 10.1098/rstb.2012.0415

Stevens, J. R. (2017). Replicability and reproducibility in comparative psychology. *Front. Psychol.* 8:862. doi: 10.3389/fpsyg.2017.00862

Subias, L., Griffin, A. S., and Guez, D. (2019). Inference by exclusion in the red-tailed black cockatoo (*Calyptorhynchus banksii*). *Integ. Zool.* 14, 193–203. doi: 10.1111/1749-4877.12299

Takahashi, M., Ueno, Y., and Fujita, K. (2015). Inference in a social context: a comparative study of capuchin monkeys (*Cebus apella*), tree shrews (*Tupaia belangeri*), hamsters (*Mesocricetus auratus*), and rats (*Rattus norvegicus*). *J. Comp. Psychol.* 129, 402–411. doi: 10.1037/a0039732

Taylor, A. H. (2014). Corvid cognition. *Wiley Interdiscipl. Rev. Cogn. Sci.* 5, 361–372. doi: 10.1002/wcs.1286

Taylor, A. H., Elliffe, D., Hunt, G. R., and Gray, R. D. (2010a). Complex cognition and behavioural innovation in new caledonian crows. *Proc. Biol. Sci.* 277, 2637–2643. doi: 10.1098/rspb.2010.0285

Taylor, A. H., Medina, F. S., Holzhaider, J. C., Hearne, L. J., Hunt, G. R., and Gray, R. D. (2010b). An investigation into the cognition behind spontaneous string pulling in new caledonian crows. *PLoS One* 5:e9345. doi: 10.1371/journal.pone.0009345

Taylor, A. H., Elliffe, D. M., Hunt, G. R., Emery, N. J., Clayton, N. S., and Gray, R. D. (2011). New caledonian crows learn the functional properties of novel tool types. *PLoS One* 6:e26887. doi: 10.1371/journal.pone.0026887

Taylor, A. H., and Gray, R. D. (2009). Animal cognition: aesop's fable flies from fiction to fact. *Curr. Biol.* 19, R731–R732. doi: 10.1016/j.cub.2009.07.055

Taylor, A. H., Hunt, G. R., Holzhaider, J. C., and Gray, R. D. (2007). Spontaneous metatool use by new caledonian crows. *Curr. Biol.* 17, 1504–1507. doi: 10.1016/j.cub.2007.07.057

Taylor, A. H., Knaebe, B., and Gray, R. D. (2012). An end to insight? New Caledonian crows can spontaneously solve problems without planning their actions. *Proc. Biol. Sci.* 279, 4977–4981. doi: 10.1098/rspb.2012.1998

Tebbich, S., Seed, A. M., Emery, N. J., and Clayton, N. S. (2007). Non-tool-using rooks, *Corvus frugilegus*, solve the trap-tube problem. *Anim. Cogn.* 10, 225–231. doi: 10.1007/s10071-006-0061-4

Tecwyn, E., Denison, S., Messer, E., and Buchsbaum, D. (2017). Intuitive probabilistic inference in capuchin monkeys. *Anim. Cogn.* 20, 243–256. doi: 10.1007/s10071-016-1043-9

Téglás, E., Girotto, V., Gonzalez, M., and Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proc. Natl. Acad. Sci. U. S.A.* 104, 19156–19159. doi: 10.1073/pnas.0700271104

Teschke, I., and Tebbich, S. (2011). Physical cognition and tool-use: performance of Darwin's finches in the two-trap tube task. *Anim. Cogn.* 14, 555–563. doi: 10.1007/s10071-011-0390-9

Ujfalussy, D., Ujfalussy, D., Miklósi, Á, Miklósi, Á, Bugnyar, T., and Bugnyar, T. (2013). Ontogeny of object permanence in a non-storing corvid species, the jackdaw (*Corvus monedula*). *Anim. Cogn.* 16, 405–416. doi: 10.1007/s10071-012-0581-z

Vallortigara, G., Regolin, L., Rigoni, M., and Zanforlin, M. (1998). Delayed search for a concealed imprinted object in the domestic chick. *Anim. Cogn.* 1, 17–24. doi: 10.1007/s100710050003

van Horik, J., and Emery, N. J. (2011). Evolution of cognition. *Wiley Interdiscipl. Rev. Cogn. Sci.* 2, 621–633. doi: 10.1002/wcs.144

van Horik, J. O., and Madden, J. R. (2016). A problem with problem solving: motivational traits, but not cognition, predict success on novel operant foraging tasks. *Anim. Behav.* 114, 189–198. doi: 10.1016/j.anbehav.2016.02.006

von Bayern, A. M. P., Danel, S., Auersperg, A. M. I., Mioduszewska, B., and Kacelnik, A. (2018). Compound tool construction by New Caledonian crows. *Sci. Rep.* 8, 15676–15678. doi: 10.1038/s41598-018-33458-z

von Bayern, A. M. P., Heathcote, R. J. P., Rutz, C., and Kacelnik, A. (2009). The role of experience in problem solving and innovative tool use in crows. *Curr. Biol.* 19, 1965–1968. doi: 10.1016/j.cub.2009.10.037

Wagenmakers, E., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., et al. (2018). Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* 25, 35–57. doi: 10.3758/s13423-017-1343-3

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804. doi: 10.3758/bf03194105

Weir, A. A. S., Chappell, J., and Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science* 297, 981. doi: 10.1126/science.1073433

Wimpenny, J. H., Weir, A. A. S., Clayton, L., Rutz, C., and Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in New Caledonian Crows. *PLoS One* 4:e6471. doi: 10.1371/journal.pone.0006471

Xu, F., and Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition* 112, 97–104. doi: 10.1016/j.cognition.2009.04.006

Zucca, P., Milos, N., and Vallortigara, G. (2007). Piagetian object permanence and its development in Eurasian jays (*Garrulus glandarius*). *Anim. Cogn.* 10, 243–258. doi: 10.1007/s10071-006-0063-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 July 2020
doi: 10.3389/fpsyg.2020.01350

Spatial Learning in Japanese Eels Using Extra- and Intra-Maze Cues

Shigeru Watanabe*

Department of Psychology, Keio University, Tokyo, Japan

Japanese eels (*Anguilla japonica*) were trained on a spatial-learning paradigm in a pool placed in an experimental room where several extra-maze cues were present. Four tubes were placed in the pool, of which one was open and could be entered by the eels. The open tube was placed at a fixed position in the pool and contained a triangular block that served as an intra-maze cue. The eels learned to identify the open tube, and their performance was maintained when the pool was rotated. However, they were unable to maintain their performance in a dark room, which suggests that spatial learning is based on visual cues. To determine the influence of the extra- and intra-maze cues, the tube with the triangle was moved to a new position and another open tube was kept in its place. The eels chose either the tube at the original position or the tube with the triangle at its new position, suggesting that spatial discrimination may be based on either extra- or intra-maze cues. We thus conclude that the eels employed an adjunctive strategy of multiple cues. In the next experiment, the eels were trained to visually discriminate the position of the stimulus (triangle), which changed in every trial. After the training, the eels were submitted to a test in which, in addition to the triangular pattern, a rectangular pattern was introduced. The eels discriminated between the tubes with the triangular and rectangular patterns, suggesting that they had the ability to discriminate visual patterns.

Keywords: spatial cognition, visual discrimination, intra-maze cue, extra-maze cue, attention

OPEN ACCESS

Edited by:
Michael Colombo,
University of Otago, New Zealand

Reviewed by:
Damian Scarf,
University of Otago, New Zealand
Jonas Rose,
Ruhr University Bochum, Germany

***Correspondence:**
Shigeru Watanabe
swat@flet.keio.ac.jp

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *21 February 2020*
Accepted: *20 May 2020*
Published: *16 July 2020*

Citation:
Watanabe S (2020) Spatial Learning in Japanese Eels Using Extra- and Intra-Maze Cues. Front. Psychol. 11:1350. doi: 10.3389/fpsyg.2020.01350

INTRODUCTION

Animals living in the hydrosphere are important subjects for studies on comparative psychology because more than 26000 species of the 45000 vertebrates on earth live underwater. To find a general rule of learning or to find divergence in learning abilities, studying the behavior of animals in the hydrosphere is imperative. MacPhail (1985) hypothesized that there is no difference between the intellects of non-human vertebrates. Recent studies on fish demonstrated a variety of higher cognitive abilities comparable to those of mammals, including self-recognition in mirrors (Kohda et al., 2019), sense of numbers (Agrillo et al., 2012), human face recognition (Newport et al., 2018), transitive inference (Grosenick et al., 2007), or episode-like memory (Hamilton et al., 2016). Some fish also displayed complex architecture comparable to that of bower birds (Matsuura, 2015).

Some fish show remarkable orientation and navigating abilities for migration (Dodson, 1988; Braithwaite and Burt de Perera, 2006) and others live in complex environments that require them to have considerable spatial memory (Brown, 2015). The jumping goby (*Bathygobius soporator*), for instance, swims over tidal areas during high tides and apparently learns the topography of the area. Aaronson (1971) constructed an artificial tide pool and confirmed that this species learned spatial configuration during high tides and used it to jump to safe tide pools during low tides. This study thus demonstrated that fish are capable of spatial learning.

The radial arm maze was designed by Olton and Samuelson (1976) to measure spatial learning and memory in rats. Roitblat et al. (1982) trained Siamese fighting fish (*Betta splendens*) in an eight-armed maze, and they found that the fish made an average of 6.63 successful choices in eight attempts; they also inserted a 5-min delay between the fourth and fifth choice, which reduced the overall success rate. Hughes and Blight (1999) performed an experiment using two intertidal species, *Spinachia spinachia* and *Crenilabrus melops*, which belong to different families but live in similar habitats. Without any extra-maze spatial cues, both species showed a fixed pattern while visiting the arms; however, they followed a spatial memory-related strategy when visual intra-maze cues (i.e., colored tiles on the floor) were present. Rotating the maze did not impair their foraging behavior, but rearranging the tiles did; therefore, Hughes and Blight (1999) suggested that the fish used tile configuration as a cue for spatial learning. These species also associated visual cues with food location in the radial maze (Hughes and Blight, 2000).

Another common apparatus used in spatial memory experiments on rodents is the Morris water maze (Morris, 1981). An apparatus for spatial learning in goldfish, similar to a dry version of the Morris maze, was developed by Saito and Watanabe (2005). This maze had 16 small holes, one of which was baited. The study showed that, after the goldfish learned the position of the baited hole, rotating the maze did not affect the fish's performance, whereas covering the maze with a curtain disrupted their performance. Moreover, sectioning the fish's olfactory nerves did not impair performance, but eye enucleation did. These results led the authors to conclude that goldfish used visual extra-maze cues for spatial learning. In another experiment (Saito and Watanabe, 2004), a landmark was placed in the maze. The study showed that, even when the landmark and food positions were changed every day, as long as the spatial relation of the two with reference to each other was fixed, the fish were able to learn this task using the visual intra-maze cue. Durán et al. (2008) also reported intra-maze cue-learning in goldfish. These results demonstrated that fish have spatial-learning abilities comparable to those of rodents.

Eels have an outstanding migratory ability. Tsukamoto et al. (2011) found hatched eggs and Japanese eel larvae in the west Mariana Ridge, which indicates that larvae are able to migrate from this location to Japan and that adult eels can swim thousands of kilometers back to the west Mariana Ridge. Studies have examined the sensory systems of eels, such as olfaction (Westin, 1990; Tesch et al., 1991; Barbin et al., 1998; Huertas et al., 2010; Churcher et al., 2015; Schmucker et al., 2016), magnetic sense (Nishi et al., 2005; Cresci et al., 2017; Naisbett-Jones et al., 2017), and vision (Omura et al., 1997; Byzov et al., 1998). The retina of Japanese eel larvae has cones (Omura et al., 1997), and Byzov et al. (1998) reported yellow-sensitive and green-sensitive cones in European eels (*Anguilla anguilla*), indicating that these animals have color vision. However, except for the study by Watanabe and Shinozuka (2020), the spatial-learning ability of eels has not yet been examined in a laboratory.

Setting up a suitable apparatus to examine spatial learning in a new species is crucial. A possible apparatus that could be used for eels is the one used for octopi. Boal et al. (2000) released an octopus in a pool with five closed and one open burrow; the octopus learned the position of the open one. This apparatus is functionally similar to the Morris maze. As eels prefer to hide in small holes such as tubes, we examined spatial learning in eels using this habit as reinforcement (Watanabe and Shinozuka, 2020). Four tubes were placed in a round pool within an experimental chamber with several extra-maze cues. One tube, placed at a fixed position, was open, and the other tubes were closed. When the eel reached the open tube, it was allowed to stay in there for 10 min; thus, they learned the position of the open tube. The eels were unable to maintain their discriminative behavior when the test was performed in a dark room. These results demonstrated that spatial learning in eels was based on extra-maze visual cues. Natural settings commonly have both extra- and intra-maze cues. Therefore, in the present study, eels were trained in a pool with both extra- and intra-maze cues, and their spatial discrimination based on these cues was examined.

EXPERIMENT I: MAZE WITH EXTRA- AND INTRA-MAZE CUES

Materials and Methods

Study Subjects

Nine Japanese eels (*Anguilla japonica*), obtained from Omori-Tansui Co., Ltd. (Miyazaki, Japan), were used in this experiment. The length of the eels was 22–35 cm. The eels were housed individually in aquaria (39.8 × 25.4 × 28 cm) that were filled with dechlorinated tap water and fitted with an air pump. Sand was placed on the floor of each aquarium, and a gray vinyl chloride tube (inner diameter, 1.5 cm; length, 24 cm) was added to each aquarium. The experiments were initiated 2 weeks after the eels were brought to the laboratory. A 13L:11D artificial illumination cycle was used, but the racks holding the aquariums were covered with a gray vinyl curtain. Earthworms were provided once a week, but most of the eels did not eat them.

Apparatus

Figure 1 shows the experimental setup and apparatus. The experimental maze consisted of a white polypropylene circular pool (diameter, 100 cm; depth, 38 cm) filled with dechlorinated tap water (**Figure 1**). The water level was 5 cm from the bottom of the pool. The water temperature was maintained at 25°C, and the water was changed every fifth day. The experimental room was illuminated with fluorescent lamps, and there were several extra-maze cues in the room (**Figure 1A**). The pool contained four gray vinyl chloride tubes (inner diameter, 1.6 cm; length, 24 cm), and each tube had four lead weights (20 mm^3) attached to it to affix it to the floor. A transparent acyl cylinder (diameter, 15 mm; length, 30 mm) was inserted into both ends of the tubes. For the three closed tubes, acyl screws fixed the acyl cylinders so that the eels could not enter the tube. For the open tube, the acyl cylinders were not fixed with screws so that the eels would be able to enter the tube. A gray vinyl chloride equilateral triangle

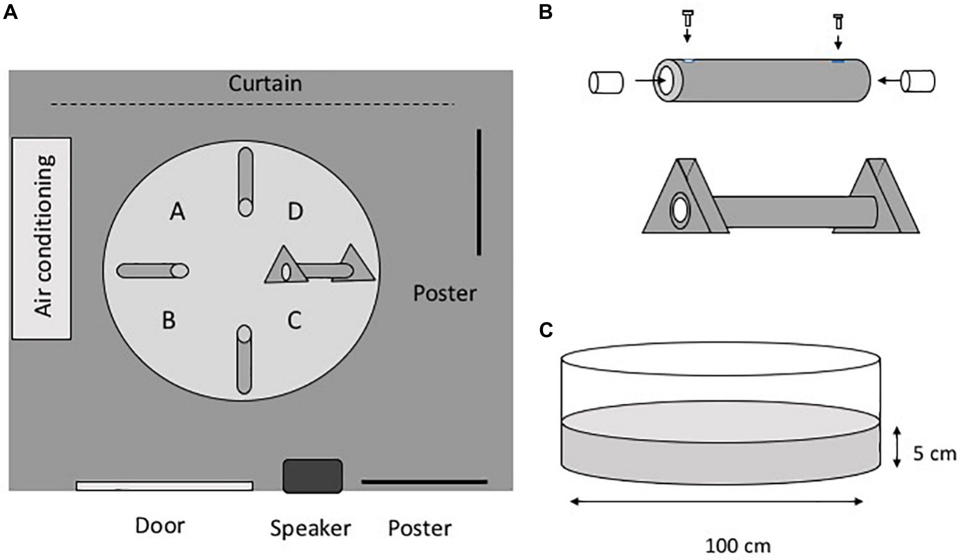


FIGURE 1 | Apparatus. **(A)** Plain view of the experimental room. The room has several extra-maze cues, such as posters and furniture. A–D indicate the place of release of the subjects. **(B)** Tube. To prepare the closed tubes, a small acyl cylinder was fixed at both ends of the tube. The distance from the entrance of the tube to the cylinder is 1.0 cm. The cylinders are movable in the open tubes. A triangular block was fixed at both ends of the open tube. **(C)** Pool.

(base, 80 mm; thickness, 30 mm) was fixed at both ends of the open tube. Eel behavior was monitored using a CCD camera (G100, NEC Avio, Tokyo, Japan) connected to a computer. In one of the tests, which was carried out in a dark room, a night scope (Super Night Compact 1000 NDX; Kenko Tokina Co., Ltd., Tokyo, Japan) was also used to observe the eels.

Habituation to the Apparatus

Each eel was individually habituated to the apparatus. An eel was first transferred from its aquarium to a bucket using a nylon net. Then, it was gently released from the bucket into the pool. During the habituation phase, all the tubes were open and no triangles were attached to any tubes. The eel was allowed to move around the pool for 10 min during which time it typically selected one tube and stayed inside it. At the end of the 10 min, the eel was returned to the aquarium. All the tubes were cleaned with a brush before the next eel was released into the apparatus. This procedure was repeated for 2 days.

Spatial Discrimination Training

During the spatial discrimination training phase, the treatment of the eels was identical to that during the habituation phase, except that only one tube with a triangle at a fixed position was open. An eel could visit the tubes until it reached the open tube. When the eel entered the open tube, it was allowed to stay there for 10 min. The first choice was recorded as its response in that trial. If the eel did not enter the tube, the eel was retrieved after 10 min. Because catching and releasing causes stress in the eels, all eels underwent only one training trial per day, and the position at which they were released was randomly assigned (**Figure 1**). The eels had to undergo at least 10 training trials, and the criterion used to determine if the eel was trained was three correct responses within four successive trials (P = 0.004, binominal test). Once trained, the following tests were performed.

Rotation Test (Three Trials)

The eels might use unknown visual cues inside the pool to learn the position of the correct tube. To prevent the eels from using such cues, the pool was rotated 90°, 180°, or 270° at each trial. Each eel underwent only one trial at each of the three rotation positions. The open tube remained at its original position relative to the room. Again, each eel underwent one trial test per day.

Dark Room Test (Four Trials)

To eliminate visual cues, a test similar to that of the spatial discrimination training was carried out in a dark room. The leaked illumination was measured from nine positions inside the pool. Mean illuminance was 0.10 Lx in the dark room, whereas that in the illuminated room was 368 Lx. The test was repeated four times for each eel, with one trial per day.

Cue Separation Test (Three Trials)

In this test, the local cue (triangle) and global cue (the position of the open tube) were separated. The tube with the triangle was moved to one of three new positions. At its original position, another tube was placed. The tube with the triangular cue and the newly placed tube were both open so that eels could enter either of them. The test procedure was similar to that of the spatial discrimination training, except that there were two open tubes. The test was repeated three times; at each trial, the position of the tube with the triangular cue was changed.

Bidimensional (2D) Triangle Test (Four Trials)

In this test, the triangular blocks had tridimensional (3D) stimuli; i.e., not only visual cues but also tactile cues could be used. In this test, a triangular pattern, similar in shape, color, and size to that of the original triangular block, was pasted on an 80 × 80 mm transparent acyl board, and the board was fixed at both ends of the open tube. Transparent acyl boards without patterns were fixed on the closed tubes. The test consisted of four trials.

Statistical Procedures

A single sample *t*-test was used for the analysis of cumulative correct responses in each test. One-factor ANOVA was used for comparison among tests, and Shaffer's modified sequentially rejective Bonferroni procedure was used for *post hoc* multiple comparisons.

Results

Figure 2A shows the averaged forward-learning curve. The vertical axis indicates the cumulative number of correct responses. The eels that achieved the discrimination criterion (i.e., three correct responses in four consecutive trials) were not considered for averaging the calculations of cumulative correct trials. All eels reached the discrimination criterion. The fastest an eel took to reach the criterion was after 10 trials, whereas the slowest was after 18 trials (average: 14.3 trials). **Figure 2B** shows the averaged latency. The latency to reach the open tube decreased throughout the trials. The means of the first and last four trials were 140 s (sd = 83.4) and 66.1 s (sd = 35.8), respectively, and they were significantly different [two-tailed paired *t*-test, $t(8) = 2.56$, $P = 0.03$]. Thus, the eels learned how to detect the correct tube quickly.

Figure 3 shows the results of the four tests that were performed. As four trials were performed for the dark test and three for the rotation and 2D tests, the chance of the eels having a successful performance was 1.25 and 1.0, respectively. As there were two correct (i.e., open) tubes (i.e., one with the correct pattern and one at the correct position) in the cue separation test, the chance of having a successful performance in three trials was 1.5. The subjects clearly maintained their discriminative behavior in all tests, except in the dark room test. In the total of 36 dark test trials, a subject did not reach the open tube within 10 min in only 1 trial. The subjects demonstrated perfect performance in the cue separation test. Therefore, the eels detected the open tubes when the local pattern cue and global position cue were separated. Single sample *t*-tests revealed a significant difference in the chance level of the rotation [$t(8) = 7.0$, $P < 0.001$] and 2D tests [$t(8) = 5.57$, $P < 0.001$] but no significant difference in the dark test [$t(8) = 0.43$, $P = 0.68$]. They did not use possible visual cues, such as small scratches on the wall of the pool, to identify the open tube, and they used visual rather than tactile cues to identify the triangular block. The results of the dark test demonstrated that the eels were unable to maintain their discriminative behavior without visual information.

Figure 4A shows the differences in the chance level for each test. One-factor ANOVA indicated the significant effects of the tests [$F(3,31) = 12.90$, $P = 0.0001$]. *Post hoc* multiple comparison showed a significant difference between the dark test and the rotation [$t(7) = 4.66$, $P = 0.007$], 2D test [$t(7) = 3.64$, $P = 0.025$], and cue separation tests [$t(7) = 15.0$, $P < 0.0001$]. No significant difference was found for the other combinations.

Figure 4B shows the mean latency to reach the open tube in the four tests. Analysis of latency demonstrated that the eels required a longer time to reach the open tube in the dark than under light conditions. One-factor ANOVA indicated the significant effects of the tests [$F(3,31) = 10.81$, $P = 0.0002$]. The *post hoc* multiple comparison showed a significant difference between the dark and rotation tests [$t(7) = 4.03$, $P = 0.01$], between the dark and cue separation tests [$t(8) = 3.64$, $P = 0.025$], and between the dark and 2D tests [$t(7) = 5.00$, $P = 0.009$]. No significant difference was found for the other combinations.

Figure 5 shows the individual data from the cue separation test. One individual selected the tube with a triangle in all the test trials regardless of the tube's position, and four other eels chose the tube with the triangular cue in only two trials. The other four eels chose the tube at the correct position in two trials and that with the triangle in one trial. Thus, there was no consistent

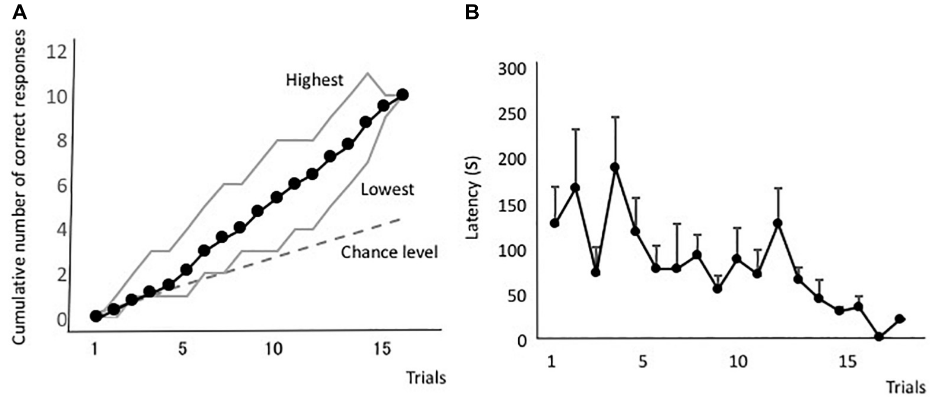


FIGURE 2 | Forward-learning curves. **(A)** The vertical axis indicates the cumulative number of correct responses. "Highest" and "Lowest" indicate the highest and lowest scores at each trial. The black line indicates the average value, the broken line indicates the expected cumulative number, and gray lines indicate the highest and lowest number at each trial. **(B)** Average time taken from release to the point of entry into the open tube. Latency of subjects that did not enter the tube is assigned as 600 s.

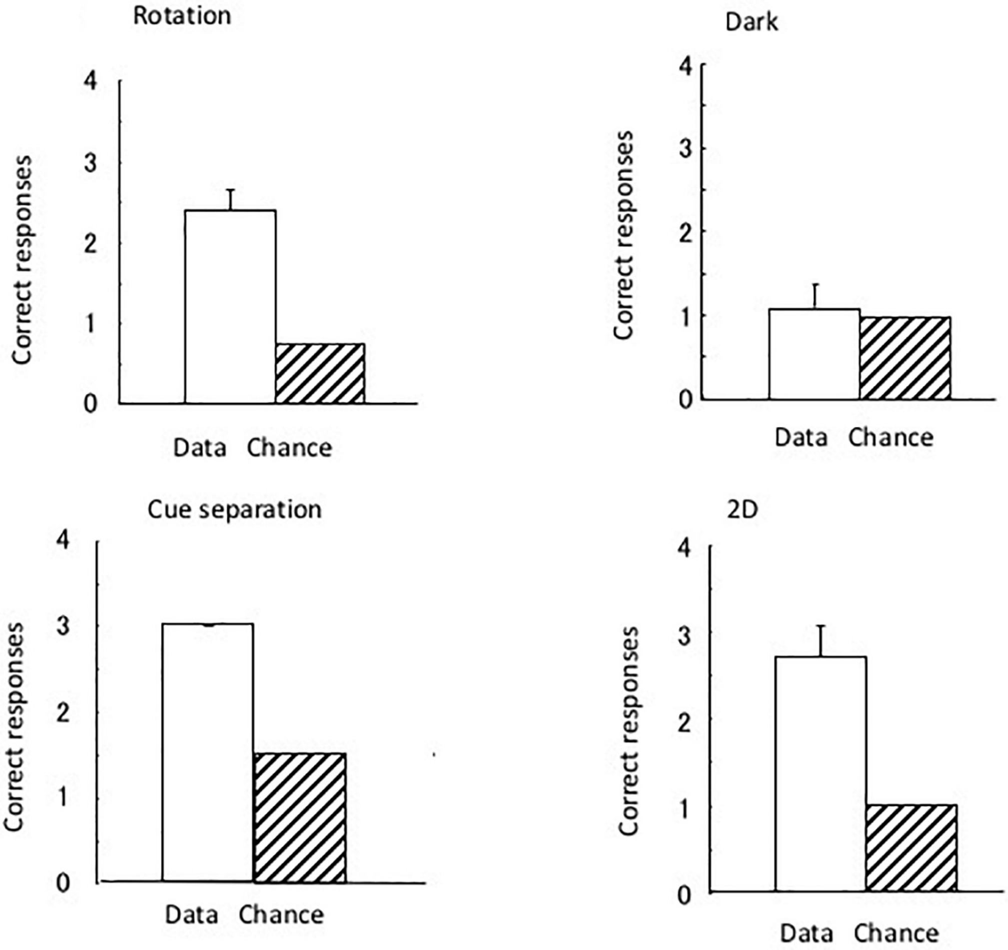


FIGURE 3 | The results of the four tests conducted under Experiment I. Open bars and hashed bars indicate the data and chance level in each test. Small bars indicate the standard error. None of the subjects showed an error in the separation test.

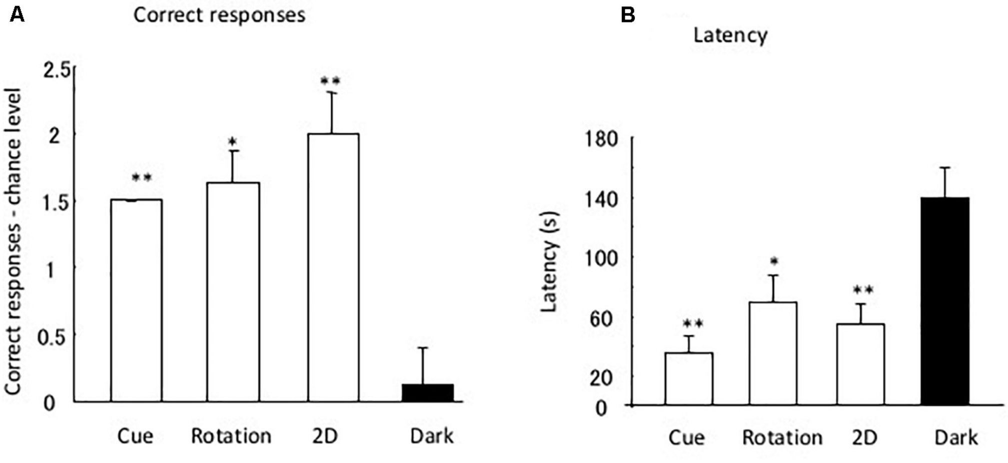


FIGURE 4 | Difference between data and chance level. (A,B) Indicate the correct response and latency, respectively. Small bars indicate the standard error. $*P < 0.05$, $**P < 0.01$.

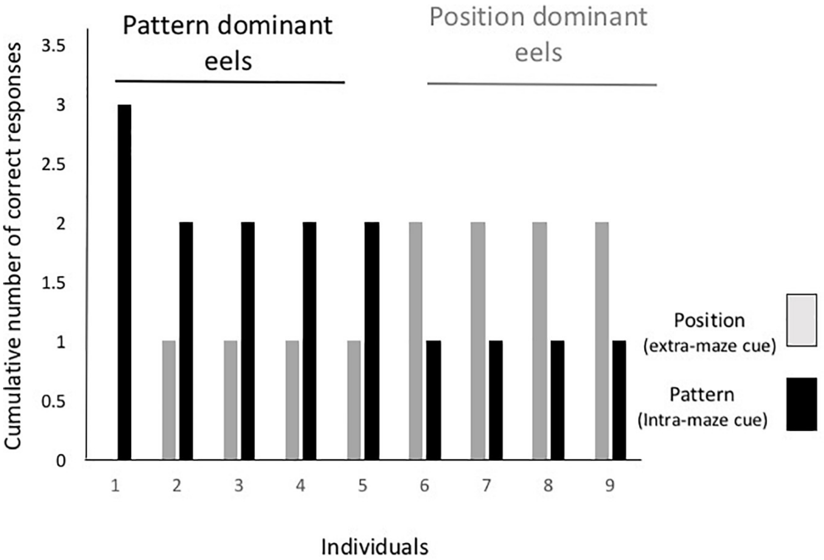


FIGURE 5 | Individual results from the cue separation test. Gray and black bars indicate the number of trials required until either the tube at the correct position or the tube with the correct pattern was chosen.

dominant cue among individuals, and cue selection varied with individual and trial.

In the first trial of the test, five eels selected the spatial cue and four selected the pattern cue. Five eels chose the same cue in the first and second trials, and three eels changed the cue in the first and second trials. Thus, no tendency regarding discriminative behavior from the first trial to the last trial was detected.

EXPERIMENT II: VISUAL DISCRIMINATION

Experiment I demonstrated that eels used not only the global positional cue but also the local visual cue for spatial discrimination. In Experiment II, the eels were trained to visually discriminate the position of the 2D stimulus (triangle), which changed in every trial. Thus, they had to learn the correct tube by the intra-maze cue alone. After the discriminative training, the eels were subjected to a test in which a triangular and a rectangular pattern were introduced. If the eels identify the correct tube as a tube with a darker board, they should choose either the tube with the triangular pattern or that with the rectangular pattern. Contrastingly, if the stimulus control by the triangle is visual pattern discrimination, the eels should choose the tube with the triangular pattern.

Materials and Methods

Subjects

Five eels from Experiment I were used for this experiment.

Apparatus

This experiment used the same apparatus as in Experiment I. The open tube had triangular boards at both ends. The other three tubes had transparent boards.

Procedure

The subjects were trained on visual discrimination using a tube with a triangular board and closed tubes with transparent acyl boards without a triangle. The tube with the triangular cue was placed at a different position in each trial. The following procedure was identical to that performed in Experiment I. The subjects were trained for at least 10 trials. The subjects that made three correct choices in four consecutive trials were used for the test, whereas those that failed to reach the criterion in 10 trials received additional training until the criterion was reached.

Triangle Versus Rectangle Test

A 2D triangle was fixed to the open tube, whereas one of the three closed tubes had a 2D rectangular pattern. For the latter, a gray rectangular pattern (80 × 80 mm) was pasted on an acyl board, and this board was fixed at each end of the closed tube. The other two closed tubes had transparent boards. The test consisted of four trials, and the positions of the open tube with the triangular cue and the closed tube with the rectangular cue were quasi-randomly assigned. There was no repetition of the same arrangement for both tubes.

Results

Figure 6A shows the forward-learning curves resulting from this test. Three eels reached the criterion in 10 trials, one in 13 trials, and one in 18 trials. Thus, the eels were able to learn visual discrimination. Latency during the discriminative training is shown in **Figure 6B**. There was no clear improvement in latency. The mean of the first and last four trials was 88.8 s (sd = 52.7) and 110.3 s (sd = 41.9), respectively. There was no statistically significant difference between them [two-tailed paired *t*-test, $t(4) = 0.10$, $P = 0.93$].

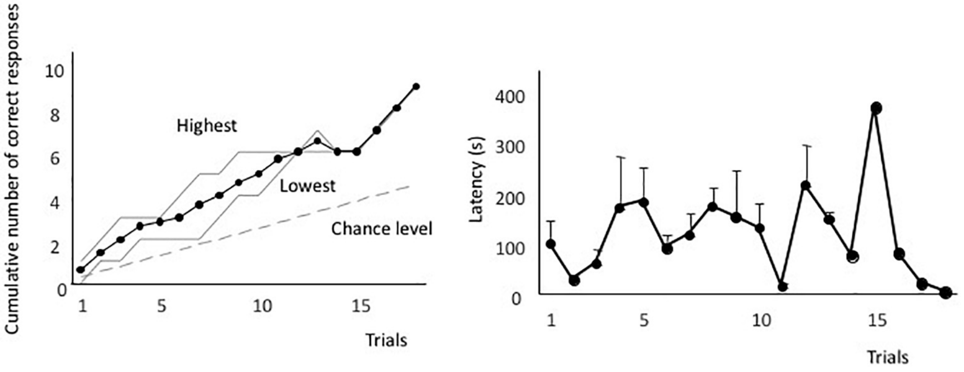


FIGURE 6 | Forward-learning curve for Experiment II **(A)**, and forward mean of the latency **(B)**.

The test results are shown in **Figure 7**. The eels chose the tube with the triangular cue. The single sample *t*-test showed a significant difference in chance level [$t(4) = 5.88$, $P < 0.005$]. The eels chose a total of four incorrect tubes in the 20 test trials, and two of the incorrect choices were regarding the tube with the rectangular cue. Thus, they were slightly distracted by the rectangular pattern. In other words, the shape of the stimulus (triangle) controlled the discriminative behavior of the eels.

DISCUSSION

The results of the present study show that: (1) the eels used visual cues for spatial learning; (2) they used extra- and intra-maze cues; and (3) they were able to learn visual discrimination without spatial cues.

Comparison With Other Studies

The eels' performance in the experiments show the effectiveness of using their behavior of hiding in a shelter tube as reinforcement, and these results confirm previous findings (Watanabe and Shinozuka, 2020). A previous experiment showed that the eels' spatial-learning ability was based only on extra-maze cues (16.4 trails, on average). We herein show that the eels' spatial-learning ability was based on both extra- and intra-maze cues (14.3 trials, on average). Although the eels seem to have had more success in the present study than in the previous one, no statistically significant difference was found between these results [two-tailed *t*-test, $t(19) = 1.13$, $P = 0.27$]. Therefore, adding intra-maze cues did not significantly improve the acquisition of spatial learning.

Because of the different methodologies used, a comparison of the present results from experiments on eels with the results from other studies on other fish species is rather difficult. Goldfish, for instance, demonstrated spatial-learning ability in a Morris-type maze with extra- or intra-maze cues after 35.2 trials on average and 22.4 trails (four trials per day), respectively (Saito and Watanabe, 2004). Moreover, the goldfish completed the Morris-like maze task with both extra- and intra-maze cues faster than that without intra-maze cues (Saito and Watanabe, 2005). Taniuchi et al. (2013) also reported the slight positive effect of intra-maze cues on the performance of goldfish in a radial

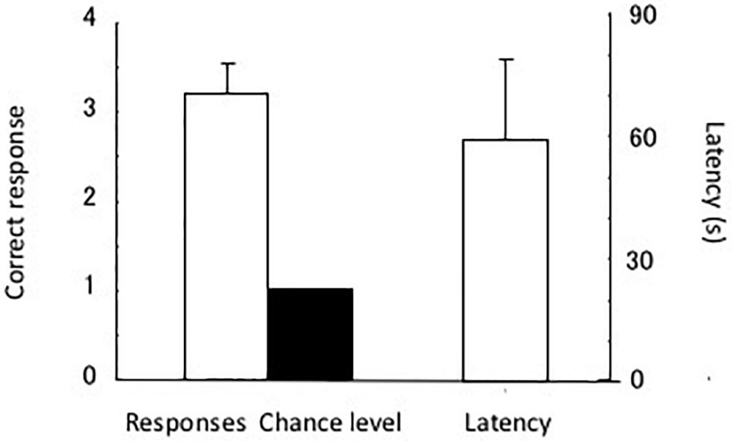


FIGURE 7 | Results of the test. Correct response **(left)** and latency **(right)**. Small bars indicate the standard error.

maze task. These results suggest the additive effect of intra- and extra-maze cues in goldfish.

Octopi trained to find an open burrow showed a clear decrease in the distance traveled to find the open burrow within six trials (Boal et al., 2000). Although the data on the success rate is not available and the number of choices is different from that in the present experiment, the performance of the octopi seems comparable to that of fish. Zebra finches trained to choose one of four feeders in a flying cage acquired spatial learning in nine to twelve trials (Watanabe and Bischof, 2002, 2004). Although differences in the reinforcement used and the number of trials per day must be considered, this suggests that spatial learning in eels is comparable to that of zebra finches. C57/BL mice learned the standard Morris water maze in 15–20 trials (Yoshida et al., 2001). They also learned a dry-type Morris maze with auditory extra-maze cues (Watanabe and Yoshida, 2007) and one with airflow cues (Bouchekioua et al., 2014) in 15–20 trials. Therefore, successful learning of similar spatial tasks in eels, goldfish, octopus, zebra finch, and rodents supports the null hypothesis proposed by MacPhail (1985).

Visual Information and Spatial Learning

The results of the rotation and dark tests suggest that eels learned spatial discrimination based on visual cues, which confirms our previous findings (Watanabe and Shinozuka, 2020). The eels in the present experiment showed a slow response in the dark test. Although the swimming speed might be reduced in the dark, visiting incorrect tubes was the main factor prolonging the latency. In 35 of 36 cases, the eels finally reached the correct tube after visits to the wrong tubes. Thus, the chance level performance in the dark test can be attributed to a lack of discriminative behavior rather than motor deficits. Goldfish also maintained spatial discrimination after the rotation of the experimental pool, but spatial learning was lost after enucleation of the eyes (Saito and Watanabe, 2004).

The 2D test performed in Experiment I suggests that the visual cue was more influential than the tactile cue. As vision is a sense that detects distance, visual cues may be more effective in the detection of hiding places from a distance. Since the work performed by Sutherland (1969), many other studies on fish visual discrimination have been performed and have found that some fish species are capable of fine visual discrimination (e.g., human face discrimination performed by Archer fish, *Toxotes chatareus*; Newport et al., 2018). Unfortunately, no other data on eel visual discrimination is available. However, the results of Experiment II demonstrated that eels can learn to visually discriminate shapes.

Cue Selection in Spatial Learning

In this study, the experiments provided the eels with redundant extra- and intra-maze cues to detect the correct tube. Selective attention occurs when an animal uses one particular cue and ignores the other cues (Reynolds, 1961), whereas divided attention occurs when an animal processes two or more elements of compound stimuli (see Zentall, 2005 for review). The results of the present study suggest that most of the eels used both extra- and intra-maze cues rather than selective attention. The dominancy of a cue depends on its salience or on the discriminability of the cues; therefore, it is premature to presume the total absence of selective attention in eels based on the present experiment alone. Although the intra-maze cues did significantly affect the eels' discriminative behavior, the eels' perfect performance in the separation test suggests that they learned the tasks as an adjunctive of the extra- and intra-maze cues rather than as a conjunctive of them. They learned both cues during the training, and either cue type may have provided them with enough information to detect the correct tube. When the subject finds the intra-maze cue first, it is expected to choose the correct pattern tube, and when it identifies enough extra-maze cues first, it is expected to choose the correct position tube. The perfect performance and the maintenance of the latency in the cue separation test also support this adjunctive strategy.

Limitation and Future Studies

The present experiment used hiding behavior as a reinforcement of training. Thus, the spatial learning here might be a result of a particular form of training from a particular reinforcement and not a general ability of eels.

Eels are a novel subject in comparative psychology. Basic research, such as psychophysics and general-learning ability, should be done to understand their behavior. Eels are easy to maintain in a laboratory, and they may be used as possible experimental models for studies on spatial cognition. As their navigational abilities in natural settings are outstanding, behavioral studies on eels performed in a laboratory can help bridge our understanding of spatial cognition not only in the laboratory but also in field studies. Furthermore, eels are also important in fishery science. Obtaining an understanding of eel behavior will contribute to the preservation of this animal.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

Ethical review and approval was not required for the animal study because Fish are not target of our animal experiment committee. There was no deprivation nor physical invasion at all.

AUTHOR CONTRIBUTIONS

SW designed the study and conducted the experiments, collected the data, analyzed the data, and wrote the manuscript.

FUNDING

This research was supported by a Grant-in-Aid for Scientific Research by the Japanese Society for the Promotion of Science.

ACKNOWLEDGMENTS

The author expresses his gratitude to Omori-Tansui Co., Ltd., for providing the experimental animals and Katsumi Tsukamoto for his advice on the experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01350/full#supplementary-material

REFERENCES

Aaronson, L. R. (1971). Orientation and jumping behavior in the gobiid fish, *Bathygobius soporator*. *Am. Mus. Novitiates* 1486, 1–22.

Agrillo, C., Piffer, L., Bisazza, A., and Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One* 7:e31923. doi: 10.1371/journal.pone.0031923

Barbin, G. P., Parker, S. J., and McCleave, J. M. (1998). Olfactory clues play a critical role in the estuarine migration of silver-phase American eels. *Environ. Biol. Fishes* 53, 283–291.

Boal, J. G., Dunhum, A. W., Williams, K. T., and Hanlon, R. T. (2000). Experimental evidence for spatial learning in octopuses (*Octopus bimaculoides*). *J. Comp. Psychol.* 114, 246–252. doi: 10.1037/0735-7036.114.3.246

Bouchekioua, B., Mimura, M., and Watanabe, S. (2014). Air flow cued spatial learning in mice. *Anim. Cogn.* 18, 339–404.

Braithwaite, V. A., and Burt de Perera, T. (2006). Short-range orientation in fish: how fish map space. *Mar. Freshw. Behav. Physiol.* 39, 37–47. doi: 10.1080/10236240600562844

Brown, C. (2015). Fish intelligence, sentience and ethics. *Anim. Cogn.* 18, 1–17. doi: 10.1007/s10071-014-0761-0

Byzov, A. L., Damjanovic, I., Utina, I. A., Mickovic, B., Gracic, Z., and Andjus, R. K. (1998). Electrophysiological and spectral properties of second-order retinal neurons in the eel. *Comp. Biochem. Physiol. A* 121, 197–208. doi: 10.1016/s1095-6433(98)10122-8

Churcher, A. M., Hubbard, P. C., Marques, J. P., Canard, A. V. M., and Hubertas, M. (2015). Deep sequencing of the olfactory epithelium reveals specific chemosensory receptors are expressed at sexual maturity in the European eel Anguilla anguilla. *Mol. Ecol.* 24, 822–834. doi: 10.1111/mec.13065

Cresci, A., Paris, C. B., Duriff, C. M., Shema, S., Bjelland, R. M., Skiftesvik, A. B., et al. (2017). Glass eels (Anguilla anguilla) have a magnetic compass linked to the tidal cycle. *Sci. Adv.* 3:e1602007. doi: 10.1126/sciadv.1602007

Dodson, J. J. (1988). The nature and role of learning in the orientation and migratory behavior of fishes. *Environ. Biol. Fishes* 23, 161–182. doi: 10.1007/bf00004908

Durán, E., Ocaña, F. M., Gómez, A., Jiménez-Moya, F., Broglio, C., Rodríguez, F., et al. (2008). Telencephalon ablation impairs goldfish allocentric spatial learning in a hole-board task. *Acta Neurobiol. Exp.* 68, 519–525.

Grosenick, L., Clement, T. S., and Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature* 445, 429–432. doi: 10.1038/nature05511

Hamilton, T. J., Myggland, A., Dupperreaut, E., May, Z., Gallup, J., Powell, R. A., et al. (2016). Episodic-like memory in zebrafish. *Anim. Cogn.* 19, 1071–1079.

Huertas, M., Hagey, L., Hofmann, A. F., Cerda, J., Canario, A. V. M., and Hubbard, P. C. (2010). Olfactory sensitivity to bile fluid and bile salts in the European eel (*Anguilla anguilla*), goldfish (*Carassius auratus*) and Mozambique tilapia (*Oreochromis mossambicus*) suggests a 'broad range' sensitivity not confined to those produced by conspecifics alone. *J. Exp. Biol.* 213, 308–317. doi: 10.1242/jeb.033142

Hughes, R., and Blight, C. M. (1999). Algorithmic behaviour and spatial memory are used by two intertidal fish species to solve the radial maze. *Anim. Behav.* 58, 601–613. doi: 10.1006/anbe.1999.1193

Hughes, R., and Blight, C. M. (2000). Can zebrafish learn spatial tasks? An empirical analysis of place and single CS-US associative learning. *Behav. Brain Res.* 233, 415–421. doi: 10.1016/j.bbr.2012.05.024

Kohda, M., Hotta, T., Takeyama, T., Awata, S., Tanaka, H., Asai, J.-Y., et al. (2019). If a fish can pass the mark test, what are implications for consciousness and self-awareness testing in animals? *PLoS Biol.* 17:e3000021. doi: 10.1371/journal.pbio.3000021

MacPhail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B* 308, 37–51. doi: 10.1098/rstb.1985.0008

Matsuura, K. (2015). A new pufferfish of the genus Torquigener that builds "mystery circles" on sandy bottoms in the Ryukyu Islands. Japan (Actinopterygii: Tetraodontiformes: Tetraodontidae). *Ichthyol. Res.* 62, 207–212. doi: 10.1007/s10228-014-0428-5

Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learn. Motiv.* 12, 239–260. doi: 10.1016/0023-9690(81)90020-5

Naisbett-Jones, L. C., Putman, N. F., Stephenson, J. F., Ladak, S., and Young, K. A. (2017). A magnetic map leads juvenile European Eels to the Gulf Stream. *Curr. Biol.* 27, 1236–1240. doi: 10.1016/j.cub.2017.03.015

Newport, C., Wallis, G., and Siebeck, U. E. (2018). Object recognition in fish: accurate discrimination across novel views of an unfamiliar object category (human faces). *Anim. Behav.* 145, 39–49. doi: 10.1016/j.anbehav.2018.09.002

Nishi, T., Kawamura, G., and Sannomiya, S. (2005). Anosmic Japanese eel *Anguilla japonica* can no longer detect magnetic fields. *Fish. Sci.* 71, 101–106. doi: 10.1111/j.1444-2906.2005.00936.x

Olton, D., and Samuelson, R. J. (1976). Remembrance of places passed: spatial memory in rats. *J. Exp. Psychol: Anim. Learn. Cogn.* 2, 96–96.

Omura, Y., Uematsu, K., Tachiki, H., Furukawa, K., and Satoh, H. (1997). Cone cells appear also in the retina of eel larvae. *Fish. Sci.* 63, 1052–1053. doi: 10.2331/fishsci.63.1052

Reynolds, G. S. (1961). Attention in the pigeon. *J. Exp. Anal. Behav.* 4, 203–208.

Roitblat, H. L., Tham, W., and Golub, L. (1982). Performance of Betta splendens in a radial arm maze. *Anim. Learn. Behav.* 10, 108–114. doi: 10.3758/bf03212055

Saito, K., and Watanabe, S. (2004). Spatial learning deficits after the development of dorsomedial telencephalon lesions in goldfish. *Neuroreport* 15, 2695–2699.

Saito, K., and Watanabe, S. (2005). Experimental analysis of spatial learning in goldfish. *The Psychol. Rec.* 55, 647–662. doi: 10.1007/bf03395532

Schmucker, A. K., Johnson, N. S., Galbraith, H. S., and Li, W. (2016). Glass-eel-stage American eels respond to conspecific odor as a function of concentration. *Trans. Am. Fish. Soc.* 145, 712–722. doi: 10.1080/00028487.2016.1146164

Sutherland, N. S. (1969). Shape discrimination in rat, octopus, and goldfish: A comparative study. *J. Comp. Physiol. Psychol.* 67, 160–176. doi: 10.1037/h0026798

Taniuchi, T., Washizuka, K., and Ueno, R. (2013). Effects of intra-maze and extra-maze cues on radial maze performance in goldfish (*Carassius auratus*). *Jpn. J. Anim. Psychol.* 63, 79–85. doi: 10.2502/janip.63.1.2

Tesch, F. W., Westerberg, H., and Karlsson, L. (1991). Tracking studies on migrating silver eels in the Central Baltic. *Meeresforsch* 33, 183–196.

Tsukamoto, K., Chow, S., Otake, T., Kurogi, H., Mochioka, N., Miller, M. J., et al. (2011). Oceanic spawning ecology of freshwater eels in the western North Pacific. *Nat. Comm.* 2:179. doi: 10.1038/ncomms1174

Watanabe, S., and Bischof, H.-J. (2002). Spatial learning in songbird. *Curr. Psychol. Lett.* 3, 71–80.

Watanabe, S., and Bischof, H.-J. (2004). Effects of hippocampal lesions on acquisition and retention of spatial learning in zebra finches. *Behav. Brain Res.* 155, 147–152. doi: 10.1016/j.bbr.2004.04.007

Watanabe, S., and Shinozuka, K. (2020). Spatial learning in Japanese eels (*Anguilla japonica*). *Anim. Cogn.* 23, 233–236. doi: 10.1007/s10071-019-01320-y

Watanabe, S., and Yoshida, M. (2007). Auditory cued spatial learning in mice. *Physiol. Behav.* 92, 906–910. doi: 10.1016/j.physbeh.2007.06.019

Westin, L. (1990). Orientation mechanisms in migrating European silver eel (*Anguilla anguilla*): temperature and olfaction. *Mar. Biol.* 106, 175–179. doi: 10.1007/bf01314798

Yoshida, M., Goto, T., and Watanabe, S. (2001). Task-dependent strain differences of spatial learning in C57/BL and BALB/c mice. *Physiol. Behav.* 73, 37–42. doi: 10.1016/s0031-9384(01)00419-x

Zentall, T. R. (2005). Selective and divided attention in animals. *Behav. Proc.* 69, 1–15. doi: 10.1016/j.beproc.2005.01.004

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 July 2020
doi: 10.3389/fpsyg.2020.01631

Human Choice Predicted by Obtained Reinforcers, Not by Reinforcement Predictors

Jessica P. Stagner*, Vincent M. Edwards, Sara R. Bond, Jeremy A. Jasmer, Robert A. Southern and Kent D. Bodily*

Department of Psychology, Georgia Southern University, Statesboro, GA, United States

Macphail (1985) proposed that "intelligence" should not vary across vertebrate species when contextual variables are accounted for. Focusing on research involving choice behavior, the propensity for choosing an option that produces stimuli that predict the presence or absence of reinforcement but that also results in less food over time can be examined. This choice preference has been found multiple times in pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Laude et al., 2014) and has been likened to gambling behavior demonstrated by humans (Zentall, 2014, 2016). The present experiments used a similarly structured task to examine adult human preferences for reinforcement predictors and compared findings to choice behavior demonstrated by children (Lalli et al., 2000), monkeys (Smith et al., 2017; Smith and Beran, 2020), dogs (Jackson et al., 2020), rats (Chow et al., 2017; Cunningham and Shahan, 2019; Jackson et al., 2020), and pigeons (Roper and Zentall, 1999; Stagner and Zentall, 2010). In Experiment 1, adult human participants showed no preference for reinforcement predictors. Results from Experiment 2 suggest that not only were reinforcement predictors not preferred, but that perhaps reinforcement predictors had no effect at all on choice behavior. Results from Experiments 1 and 2 were further assessed using a generalized matching equation, the findings from which support that adult human choice behavior in the present research was largely determined by reinforcement history. Overall, the present results obtained from human adult participants are different than those found from pigeons in particular, suggesting that further examination of Macphail (1985) hypothesis is warranted.

Keywords: suboptimal choice, choice, matching law, preference, comparative psychology

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Michael Colombo,
University of Otago, New Zealand
Katja Liebal,
Freie Universität Berlin, Germany

****Correspondence:***
Jessica P. Stagner
jbodily@georgiasouthern.edu
Kent D. Bodily
kdbodily@georgiasouthern.edu

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *01 March 2020*
Accepted: *16 June 2020*
Published: *24 July 2020*

Citation:
Stagner JP, Edwards VM, Bond SR, Jasmer JA, Southern RA and Bodily KD (2020) Human Choice Predicted by Obtained Reinforcers, Not by Reinforcement Predictors. Front. Psychol. 11:1631. doi: 10.3389/fpsyg.2020.01631

INTRODUCTION

Macphail (1985) argued that comparative psychologists should adopt the assumption of general processes of learning. That is, despite the common notion that learning capacities vary between species and that species may be ranked by these capacities, the null hypothesis in the comparison of behavioral traits across species must be that there are no differences. Macphail specified that research had found no cross-species differences with regard to either qualitative (i.e., differences in mechanism) or quantitative (i.e., differences in efficiency of a shared mechanism). According to Macphail, the failure to convincingly rule out the null hypothesis was due to an absence of systematic replications to rule out contextual variables (e.g., motivating operations, stimulus characteristics, etc.). Parsimony, Macphail stated, requires any apparent differences must first be ascribed to contextual variables. Macphail's only exception was that verbal humans were to be excluded from this argument. But what if humans were tested in procedures that were analogous to those applied to non-humans? Here we reviewed selected systematic replications of choice behavior in human and non-human animals and presented two experiments with human participants to test Macphail's Null Hypothesis, which stated that there should be no cross-species differences in general learning processes.

Past research has demonstrated preference for an alternative which produces stimuli that signal the future presence or absence of reinforcement over an alternative which does not produce reinforcement-predictive stimuli. This preference has been found in multiple species, including capuchins, rhesus macaques, pigeons, and rats (Roper and Zentall, 1999; Smith et al., 2017; Cunningham and Shahan, 2019; Smith and Beran, 2020). For example, when Roper and Zentall (1999) presented a choice between two alternatives which produced reinforcement equally 50% of the time, pigeons' choices were more frequently allocated to the alternative which provided reinforcement-predictive stimuli (see **Figure 1A**). This preference is interesting in that pigeons did not obtain any additional food by choosing the predictive alternative. Additionally, the preference for predictive stimuli increased when the likelihood of food reinforcement for each alternative was reduced from 50% to only 12.5%. Thus, it appears that predictive stimuli might have been more valued when overall probability of reinforcement was low (Roper and Zentall, 1999).

To test the strength of the preference for predictive stimuli, Stagner and Zentall (2010) used a similar procedure, but the two choice options did not payout equally. The "suboptimal" alternative produced predictive stimuli but only resulted in reinforcement 20% of the time whereas the "optimal" alternative was not followed by predictive stimuli but resulted in reinforcement 50% of the time (see **Figure 1B**). If pigeons' choices were determined by obtained reinforcement, they should have shown a clear preference for the "optimal" option that provided two and half times more food. However, if predictors of reinforcement were valued, particularly when reinforcement was somewhat scarce due to the lean schedule, then pigeons should have shown a preference for the "suboptimal" alternative that resulted in predictive stimuli but paid out less. Stagner and Zentall (2010) found that pigeons demonstrated a near exclusive preference for the "suboptimal" option that provided predictive stimuli despite the fact that it yielded much less food over time. Pigeons were also given the same task, but choosing both the "optimal" alternative and the "suboptimal" alternative produced non-predictive stimuli. In this case, pigeons chose the "optimal" alternative that yielded more food (50%). Thus, it seems that the predictive nature of the stimuli used in this task was the mechanism for the "suboptimal" preference found by pigeons (Stagner and Zentall, 2010).

Considering these findings (Stagner and Zentall, 2010), the preference found for a "suboptimal" alternative was not due to pigeons' inability to detect the difference in yield between the two alternatives. Rather, this preference provides more support that reinforcement predictors were valued such that pigeons were foregoing food to choose the option that provided them. This finding has been suggested to be analogous to human gambling behavior in that humans that gamble incur losses but continue to engage in the behavior (Stagner and Zentall, 2010; Zentall, 2014; Zentall, 2016).

Rats have also been tested for a preference for predictive stimuli using a two lever task in which the "suboptimal" lever produced predictive stimuli but less food over time, while the "optimal" lever resulted in more food over time but no reinforcement predictors (Cunningham and Shahan, 2019). Interestingly, few rats preferred the "suboptimal" alternative that produced predictive stimuli when those stimuli were presented for the duration of 10 s that has been effective in many previous studies (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Stagner et al., 2012; Chow et al., 2017). Moreover, the duration for these predictive stimuli had to be extended to at least 30 s before eight of the nine rats in the task preferred the "suboptimal" predictive alternative relative to chance (Cunningham and Shahan, 2019). A similar study also done with rats found a preference for predictive stimuli despite that it resulted in half as much food over time when compared to its non-predictive counterpart (Chow et al., 2017). Within this study, two initial alternatives required nose-poke responses to indicate a choice selection. Following selection of the "suboptimal" alternative that produced predictive stimuli, terminal predictive stimuli were either a 10 s light presentation or a 10 s lever presentation to signal reinforcement while a 10 s blackout was used to signal no reinforcement. Rats that received a lever as a terminal stimulus signaling reinforcement chose the predictive suboptimal alternative whereas rats that received a light signaling reinforcement did not (Chow et al., 2017). However, when rats were tested with this two-alternative task but with odors used as predictive terminal stimuli, rats showed no preference for reinforcement predictors (Jackson et al., 2020). Taken together, these studies demonstrate that under certain conditions (i.e., contextual variables) rats may prefer predictive stimuli, even if it comes at the cost of food resources. It is important to note that between the two of these studies that found evidence of a preference for predictive stimuli by rats (Chow et al., 2017; Cunningham and Shahan, 2019), neither found the same degree of preference for predictive stimuli that has been consistently found by pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Stagner et al., 2012; Laude et al., 2014). That is, despite the systematic variations of contextual variables, these results may demonstrate a quantitative difference in preference for reinforcement predictors between rats and pigeons, which should not occur according to Macphail (1985).

When a similar study was conducted with rhesus macaques, subjects were more likely to choose a risky option which gave reinforcement less often if the outcomes were signaled (Smith et al., 2017). That is, with experience, macaques chose suboptimally more often if stimuli predictive of reinforcement outcomes followed those choices. Interestingly, macaques chose suboptimally around 64% of the time after several sessions

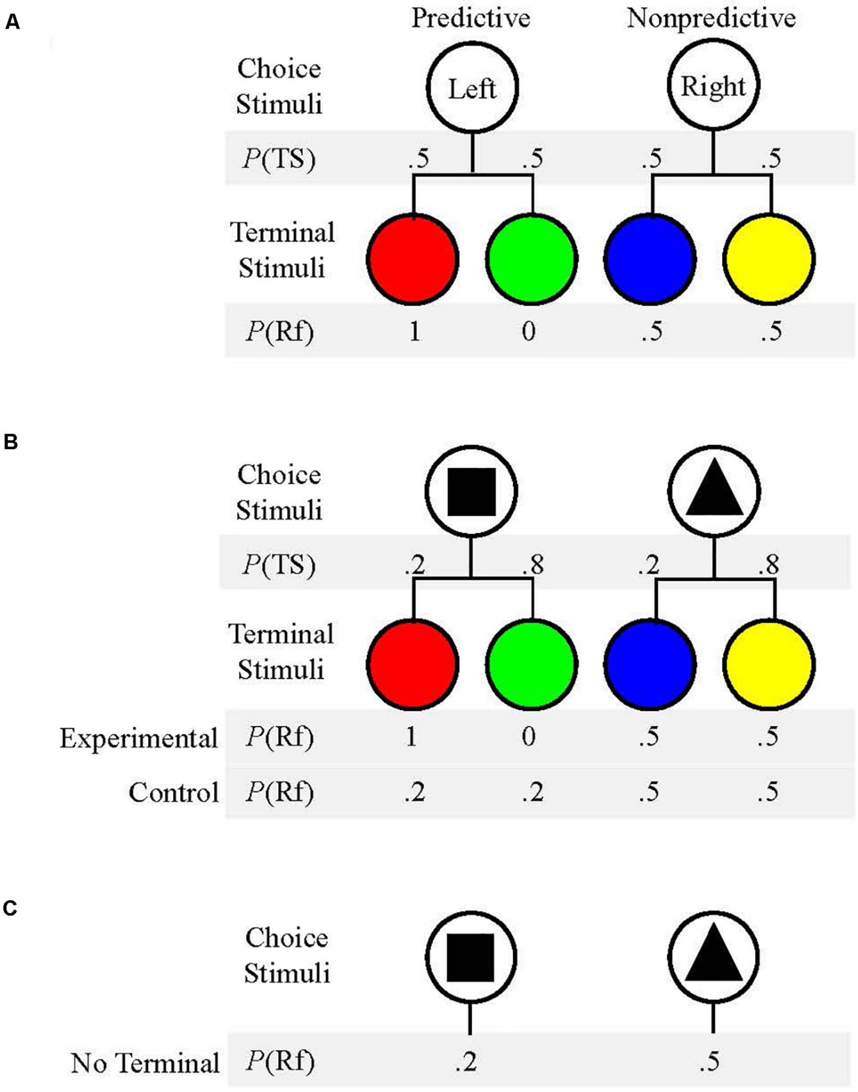


FIGURE 1 | (A) Schematic of the procedure used by Roper and Zentall (1999). Pigeons chose between the Left and Right keys (Choice Stimuli). *P*(TS) is the probability that each Terminal Stimulus would occur following a response. *P*(Rf) is the probability that reinforcement would follow each Terminal Stimulus. **(B)** Schematic of the procedure used by Stagner and Zentall (2010). Choice Stimuli were shapes presented on left or right sides with equal likelihood. The probability of each Terminal Stimulus, *P*(TS), and probability of reinforcement following each Terminal Stimulus, *P*(Rf), are given for each condition. **(C)** Schematic of procedure used in Experiment 2. Choice Stimuli were presented without any Terminal Stimuli. P(Rf) is the probability that reinforcement followed each Choice Stimulus.

of experience (Smith et al., 2017). This finding suggests a quantitative difference when comparing the preference for predictive stimuli by macaques (Smith et al., 2017) to the preference for predictive stimuli by pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Stagner et al., 2012; Laude et al., 2014).

Human participants have also been tested with similar procedures to assess preference for reinforcement predictive stimuli. Lalli et al. (2000) used a two-alternative task to assess if children with developmental delays would demonstrate a preference for reinforcement predictors. Children were given an initial choice between two black boxes. Choice of the optimal box always produced a colored block followed by reinforcement after a 30 s terminal duration. Choice of the suboptimal box produced either a 30 s colored block that signaled reinforcement or a 30 s colored block that was predictive of non-reinforcement. Lalli et al. (2000) found that children chose the box that resulted in less reinforcement as long as the colored blocks predicted the reinforcement outcome. In a condition in which the colored blocks were not predictive of reinforcement, children began to choose the optimal box option. Interestingly, when 10 s terminal durations were used for the colored block stimuli, children chose optimally. In a second experiment replicating a pigeon procedure that inserted a 10 s delay between the choice of box and the presentation of the 30 s colored block stimulus duration (Belke and Spetch, 1994), children chose the optimal box alternative that provided reinforcement 100% of the time (Lalli et al., 2000).

Taken together, these findings support that children with developmental delays, like pigeons, rats, and macaques, will choose suboptimally, under certain conditions, if stimuli are provided that predict reinforcement outcomes. However, it is important to note that children (Lalli et al., 2000) did not have the same strength in preference for predictive stimuli that has been consistently found with pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Stagner et al., 2012; Laude et al., 2014). Thus, there appears to be more similarity in choice allocation between developmentally delayed children (Lalli et al., 2000) and rhesus macaques (Smith et al., 2017) within this two-alternative task, while a quantitative difference seems to be present when comparing children and macaques to pigeons.

Further exploration employed adult human participants and aimed to observe how preference for predictive outcomes might influence human gambling behavior. A similar two-alternative task was used but was presented to adult human participants in the format of a video game (Molet et al., 2012). During choice trials, participants were allowed to select one of two planetary systems to kill as many generals as possible. Choosing the "suboptimal" system produced stimuli which predicted the number of generals that would be killed but resulted in fewer generals killed over time. Choosing the "optimal" system produced non-predictive stimuli but resulted in more generals killed over time. Participants in this study were selected based on their responses to a survey they completed in a screening during an introductory psychology course. Specifically, participants that reported that they engaged in gambling behaviors were assigned to the "gambling habit" group, whereas those that reported no such engagement were assigned to the "non-gambling habit" group. Participants in the "gambling habit" group selected the "suboptimal" planetary system 56.5% of the time on average, whereas participants in the "non-gambling habit" group selected the "suboptimal" system only 23% of the time on average. This was taken as support that self-reported gamblers made more "suboptimal" choices. It is important to note that while self-reported gamblers chose less optimally than self-reported non-gamblers in this task, neither group chose the "suboptimal" predictive alternative as often as has been demonstrated by pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011).

Recently, McDevitt et al. (2019) attempted to replicate previous findings in pigeons and adult humans using a two-alternative task. While pigeons performed similarly to past studies in that they preferred a suboptimal alternative that produced reinforcement-predictive stimuli, adult human participants showed no such preference (McDevitt et al., 2019). Specifically, adult human participants demonstrated a clear preference for an optimal alternative that did not produce predictive stimuli. When an unsignaled condition was employed in which neither alternative produced predictive stimuli, adult human participants also chose the optimal alternative (McDevitt et al., 2019). Interestingly, these results are similar to those found by Molet et al. (2012) with adult human participants, but contrasts with results from developmentally delayed children (Lalli et al., 2000), rhesus macaques (Smith et al., 2017), rats (Chow et al., 2017; Cunningham and Shahan, 2019), and pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011). Collectively, within the range of contextual variables that have been systematically investigated, it seems that there is a difference between vertebrate species when observing preference for reinforcement-predictive stimuli. This finding contrasts with the notion that these phenomena would not differ in nature across vertebrate species (Macphail, 1985).

The present experiments were conducted to continue to explore the preference for predictive stimuli in adult human participants. In Experiment 1, participants were presented with a computer task that replicated the method of Stagner and Zentall (2010, see **Figure 1B**). To assess if terminal stimulus duration would affect choice behavior as has been found with rats (Cunningham and Shahan, 2019), the terminal stimulus duration was systematically manipulated across three conditions: 2, 8, and 20 s terminal durations. This procedure was chosen to allow for direct comparison to other studies conducted with pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011) and with rats (Chow et al., 2017; Cunningham and Shahan, 2019). If human participants prefer predictive stimuli as pigeons, and sometimes rats, do within this task structure, then participants should select the alternative that provides those stimuli but that pays out less often. This finding would also support the hypothesis that this preference/susceptibility is the same across vertebrate species (Macphail, 1985). Alternatively, if participants' choices do not correspond to predictive stimuli, it might suggest that there is a difference between these species.

EXPERIMENT 1

Method

Participants

Participants (N = 73) were undergraduate students over the age of 18 who were recruited from a subject pool of the Department of Psychology at Georgia Southern University. They selected to participate using SONA Systems[1] and received course credit for their participation. Participants signed a consent form before beginning an experimental session. Once the experimental session was completed, participants were debriefed. No deception was used in Experiment 1. Data from six participants were excluded; three were excluded for incomplete data, and three were excluded because they exhibited a side bias (a preference for one side that was greater than two standard deviations away from the group mean). This left a final number of 67 participants, 55 females and 12 males. Participants in the Experimental Condition received reinforcement-predictive terminal stimuli whereas participants in the Control Condition received terminal stimuli that did not predict reinforcement. Within both the Experimental and Control Conditions there were three conditions which had terminal stimuli durations of 2, 8, and 20 s, respectively. Participants were randomly assigned to one of six conditions resulting in the following compositions: Experimental Condition 2 s (n = 14, 9 female and 5 male),

---

[1] www.sona-systems.com

Experimental Condition 8 s (n = 10, 7 female and 3 male), Experimental Condition 20 s (n = 9, 6 female and 3 male), Control Condition 2 s (n = 17, 11 female and 6 male), Control Condition 8 s (n = 9, 6 female and 3 male), and Control Condition 20 s (n = 8, 6 female and 2 male).

Apparatus

Experimental tasks were run using Windows 10 and presented on a ThinkVision L2250p 22in monitor with a resolution of 1,024 pixels × 768 pixels. All procedures were programed with OpenSesame version 3.2.8 (OpenSesame, RRID:SCR_002849, Mathôt et al., 2012). Participants experienced individual sessions in separate rooms with a researcher seated outside of that room in the waiting area of the lab.

Stimuli were presented 14.5 cm down the screen, and 10.5 cm across the screen from the left for the left location position or 38 cm across the screen from the left for the right position location. The square shape stimulus (2.61 cm × 3.16 cm) and the triangle shape stimulus (3.62 cm × 4.49 cm) appeared in either of these two locations. During forced trials, a gray circle (17.58 cm in circumference and 5.6 cm in diameter; rgb: 128, 128, and 128) was displayed in the side location that did not have a shape stimulus. Terminal stimuli appeared in the same side locations as the shape stimuli and measured 17.58 cm in circumference and 5.6 cm in diameter. The terminal stimuli displayed were colored either red (rgb: 255, 0, and 0), green (rgb: 0, 255, and 0), blue (rgb: 0, 0, and 255), or yellow (rgb: 255, 255, and 0). Reinforcement was an image of a gold coin (14.44 cm in circumference and 4.6 cm in diameter) that was centrally located 24 cm down the screen. The image of this coin was paired with an auditory stimulus similar to the sound of an old cash register ("*ka-ching!*"). In the top right corner of the screen (5.3 cm from the top of the screen and 34 cm from the left side of the screen) a green rectangle (0.89 cm × 3.52 cm) counted the reinforcers obtained. Directly above the counter were the words "Coins Received."

Procedure

The procedures used in Experiment 1, which were approved by the Institutional Review Board at Georgia Southern University, were similar to those previously used with non-human subjects (see **Figure 1B**) but modified for adult human participants (see **Figure 2**). An image of a coin and an auditory stimulus were used instead of food grain or pellets. Rather than responding on keys in an operant chamber, participants responded with the left and right arrow keys of a computer keyboard. Before the initiation of a session, participants were instructed to read along while a research assistant read the following directions on the computer screen aloud:

> Thank you for agreeing to participate in this study.
> Your goal is to earn as many coins as you can.
> Items will appear on the left and/or right of the screen.
> Press "LEFT ARROW" to choose the left option.
> Press "RIGHT ARROW" to choose the right option.
> Get as many coins as you can. The number of coins you've received appears at the top-right corner.
> Press any key to begin!

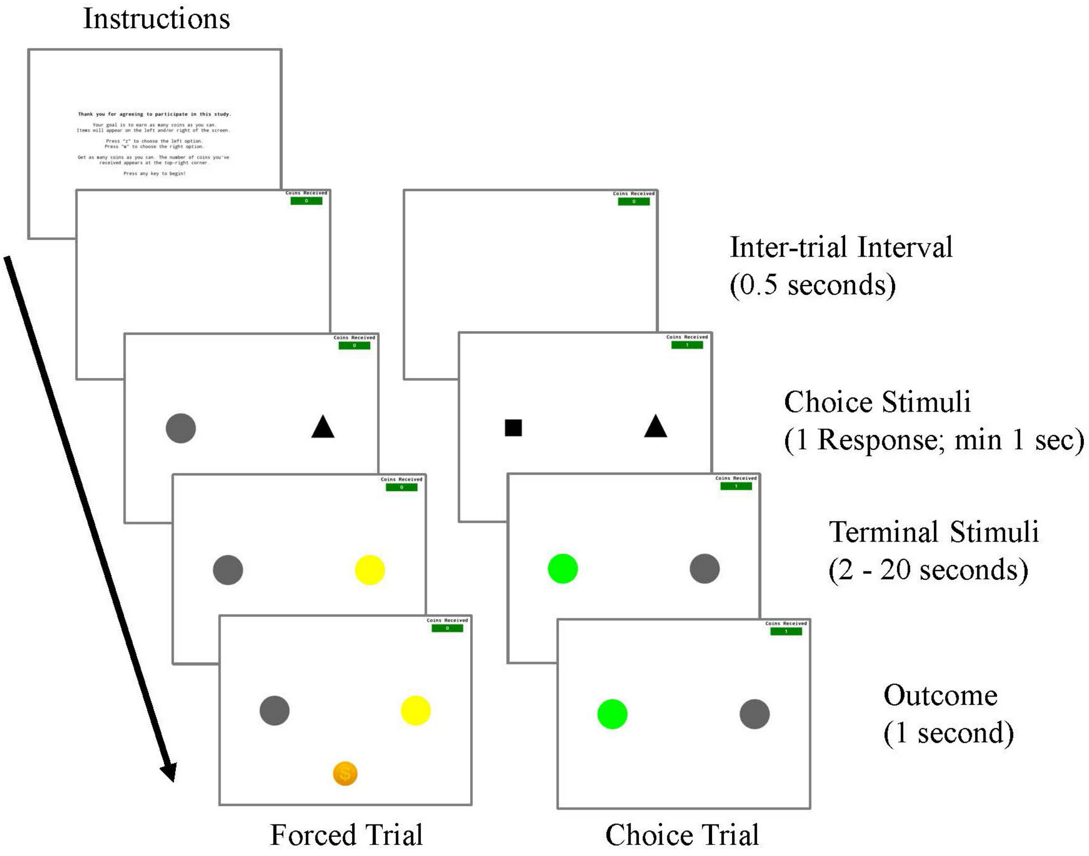
A session began once a key response was made. Forced trials were initiated by the presentation of either a square or a triangle choice stimulus in the left or the right side position. Following a single key response to the shape stimulus (FR1), the shape stimulus offset and a terminal color stimulus appeared in the same position. For Experimental Conditions, one choice stimulus was associated with two terminal stimuli, one of which was presented on 20% of the choices and was always followed by reinforcement, whereas the other was presented on 80% of these trials and was never followed by reinforcement. The other choice stimulus was associated with two terminal stimuli that occurred with equal frequency (50% of these choice trials), and were each followed by reinforcement 50% of the time. For Control Conditions, each choice stimulus was associated with two non-predictive terminal stimuli. The probability of appearance of these terminal stimuli was equated to those in the Experimental conditions. Specifically, for one choice stimulus, a terminal stimulus appeared on 20% of these shape trials while another terminal stimulus appeared on 80% of these choice trials and were each followed by reinforcement 20% of the time. For the other choice stimulus, both terminal links appeared with equal probability (50% of these choice trials) and were followed by reinforcement 50% of the time. The durations of the terminal stimuli were 2, 8, or 20 s, depending on the participant's assigned condition. After the terminal stimulus duration elapsed a reinforcer was presented for 1 s according to the schedule, followed by a 0.5 s inter-trial interval. The inter-trial interval consisted of a blank screen with no stimuli present aside from the reinforcement counter. There were a total of 40 forced trials per session.

Choice trials were initiated with the simultaneous presentation of the shape stimuli in the left and right positions. A single key response to either shape stimulus was followed by the offset of both shape stimuli and the presentation of a terminal stimulus associated with the chosen shape. Probabilities associated with both terminal stimuli appearance and reinforcement were the same as in forced trials. There were a total of 20 choice trials per session.

A complete session consisted of 60 trials, with two blocks of 20 forced trials and 10 choice trials each. A greater number of forced trials were used to ensure that participants had ample experience with each alternative, and to replicate previous procedures that have also used more forced trials than choice trials (Stagner and Zentall, 2010; Zentall and Stagner, 2011). Trial presentation was randomized, but it was ensured that there would be one choice trial for every three trials and that there was never more than one choice trial presented consecutively. Both shape and terminal stimuli were counterbalanced for side presentation, and the shape stimuli that terminal stimuli were associated with were counterbalanced across conditions.

Results

Statistical analyses were conducted using JASP 0.11.1 (JASP Team, 2019, RRID:SCR_015823). Means are reported with 95% confidence intervals and all significant effects are reported at a $p < 0.05$. **Figure 3** shows the mean proportion of choosing



Thank you for agreeing to participate in this study.

Your goal is to earn as many coins as you can.
Items will appear on the left and/or right of the screen.

Press "z" to choose the left option.
Press "m" to choose the right option.

Get as many coins as you can. The number of coins you've
received appears at the top-right corner.

Press any key to begin!

FIGURE 2 | Procedure for Experiment 1.

the higher probability of reinforcement, or "optimal," alternative plotted across blocks of 10 choice trials. A mixed Analysis of Variance (ANOVA) was used to determine the effects of Terminal Stimulus Type (predictive or non-predictive) and Terminal Stimulus Duration (2, 8, and 20 s) on mean proportion of optimal choices across Blocks 1 and 2. There was no main effect of Block [$F(1) = 0.909$, $p = 0.344$], no main effect of Terminal Stimulus Type [$F(1) = 0.042$, $p = 0.839$], and no main effect of Terminal Stimulus Duration [$F(2) = 0.324$, $p = 0.725$]. Due to the absence of any main effects, the data were collapsed across all conditions for further analysis. One sample two-tailed t-tests were run to detect any difference between chance performance (50%) and Block 1 [$t(66) = 1.631$, $p = 0.108$ ($M = 0.54$, 95% CI = ± 0.04)], and between chance performance and Block 2 [$t(66) = 2.312$, $p = 0.024$ ($M = 0.56$, 95% CI = ± 0.05)]. The t-test comparing Block 1 to chance was not significant but the t-test for Block 2 was, indicating that there was a change in participants' choice behavior as a function of experience with the task.

Discussion

In Experiment 1, participants allocated more choices to the optimal alternative that was followed by non-predictive stimuli during Block 2. This result indicates that participants developed a slight preference for the optimal alternative over the course of a session. This preference was not influenced by manipulations of Terminal Stimulus Type (non-predictive or predictive) or Terminal Stimulus Durations (2, 8, and 20 s) within this task.

Interestingly, not only did participants show no preference for predictive terminal stimuli, there was no clear indication that any of the terminal stimuli played a role in participants' choice behavior. Considering this finding, it is possible that participants' choice allocation in Experiment 1 was dependent on reinforcement history associated with each shape stimulus rather than the terminal stimuli that predicted reinforcement outcomes. Thus, if these terminal stimuli were removed all together, then it should not have an effect on how participants allocate their choices. To examine this, in Experiment 2, participants in the No Terminal Stimulus condition received the outcome of each trial directly after the shape stimulus/stimuli offset. The preference of participants in the No Terminal Stimulus condition was compared to the preference of new participants run in the Experimental and Control conditions. Additionally, because terminal stimulus duration did not play a role in participants' choices in Experiment 1, the 2 s terminal stimulus duration was used in Experiment 2. This shortened the duration of each trial and allowed for more trials within a session, increasing the number of experiences that participants received with the contingencies in the task.

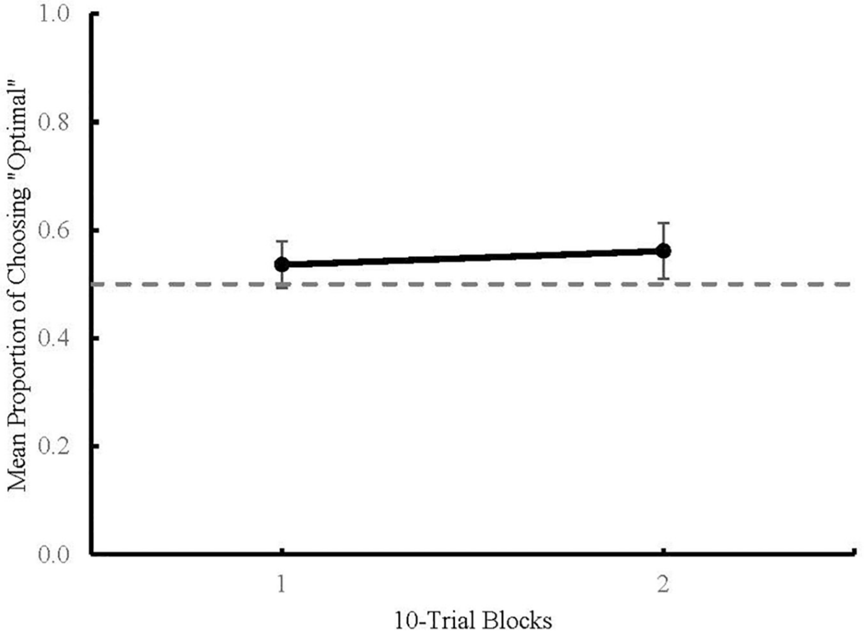


FIGURE 3 | Mean proportion of optimal choices across 10-trial blocks in Experiment 1. Dashed line represents chance (0.5). Error bars represent 95% Confidence Intervals.

EXPERIMENT 2

Method

Participants

Participants ($N = 47$) were undergraduate students over the age of 18 and were recruited as described in Experiment 1. Participants signed a consent form before beginning an experimental session. Once the experimental session was completed, participants were debriefed. No deception was used in Experiment 2. Data from three participants were excluded; one was excluded for incomplete data, and two were excluded because they exhibited a side bias (using the same criteria as in Experiment 1). Thus, there was a final total of 44 participants, 29 females and 15 males. Participants were randomly assigned to either Experimental Condition 2 s (n = 11, 6 females and 5 males) or Control Condition 2 s (n = 10, 6 females and 4 males) utilized in Experiment 1, or the No Terminal Stimulus Condition (n = 23, 17 females and 6 males).

Apparatus

The apparatus and stimuli for Experiment 2 were identical to those in Experiment 1, save the exclusion of the “Coins Received” counter and the inclusion of a progress bar. A progress bar (13.5 cm in length) was located at the top of the computer screen (positioned 2 cm down the screen and 23.5 cm toward the middle of the screen). The shading on the progress bar increased every time the participant obtained a reinforcer, filling the bar through the task. Participants were informed that the progress bar filled as they grew closer to completing the task. Completion of the task was not contingent on the participant completely filling the progress bar.

Procedure

The procedures used in Experiment 2 were approved by the Institutional Review Board at Georgia Southern University. Before the initiation of a session, participants were instructed to read along while a research assistant read the following directions on the computer screen aloud:

> Thank you for agreeing to participate in this study.
> Your goal is to earn as many coins as you can.
> Items will appear on the left and/or right of the screen.
> Press “LEFT ARROW” to choose the left option.
> Press “RIGHT ARROW” to choose the right option.
> Earn as many coins as you can. Earning coins fills the Progress bar and moves you closer to completion.
> Press any key to begin!

A session began once a key response was made. Contingencies for Experimental Condition 2 s and Control Condition 2 s were the same as in Experiment 1. For the No Terminal Stimulus Condition, trial outcomes were delivered immediately following the offset of a shape stimulus on forced trials or following the offset of both shape stimuli on choice trials. For the No Terminal Stimulus Condition, as was the case for both Experimental Condition 2 s and Control Condition 2 s (see **Figure 1C**), one shape stimulus produced reinforcement 20% of the time whereas the other shape produced reinforcement 50% of the time. There were 120 forced trials and 60 choice trials for all three conditions, for a total of 180 trials for a complete session. There were six total trial blocks consisting of 20 forced trials and 10 choice trials, and trial type presentation was randomized in the same manner as was used in Experiment 1. The progress bar was centered at the top of the screen and remained visible for the entirety of the experiment.

Results

Figure 4 graphs the proportion of mean optimal responses across choice trials in Blocks 1 through 6, each consisting of 10 choice trials with 95% confidence intervals. An ANOVA was used to determine the effect of Terminal Stimulus Condition (predictive, non-predictive, or none) on the mean proportion of optimal choice across six blocks. Results revealed no main effect of block $[F(5,205) = 2.086, p = 0.069]$; the assumption of sphericity was violated and was adjusted with the Greenhouse–Geisser correction $[F(3.69,151.38) = 2.086, p = 0.091]$. There was also no main effect of condition observed $[F(2,41) = 0.83, p = 0.443]$. Due to the lack of main effects, the data were collapsed across conditions for further analysis. **Table 1** shows results from one-sample t-tests which revealed that Blocks 2 through 6 differed significantly from chance ($M = 0.591–0.641$, 95% CI = $\pm 0.07 - \pm 0.08$). The significant t-tests for Blocks 2–6 indicate that participants' choice behavior changed over the course of a session. Specifically, participants in all conditions preferred the optimal choice stimulus.

Discussion

Choice data from the Experimental and Control conditions in Experiment 2 show an absence of preference for predictive terminal stimuli. Further, participants in all the three conditions in Experiment 2 demonstrated a propensity to select the optimal shape stimulus. This preference was not affected by the type of terminal stimuli (or lack thereof) that were associated with the optimal shape stimulus. This finding is consistent with the conclusion from Experiment 1 that reinforcement history associated with the shape stimuli, rather than the predictive nature of the terminal stimuli, was the mechanism responsible for choice behavior in this task.

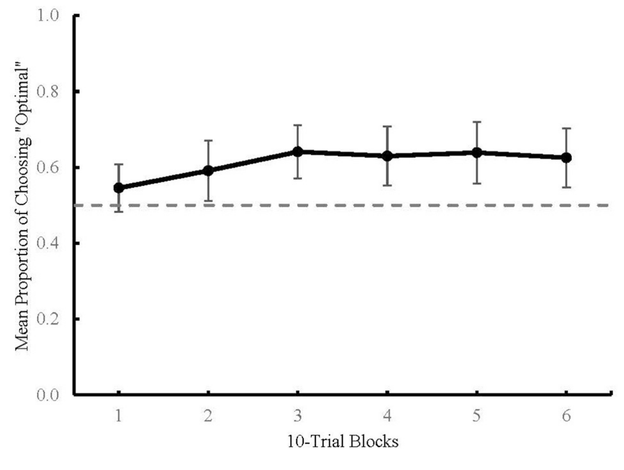


FIGURE 4 | Mean proportion of optimal choices across 10-trial blocks in Experiment 2. Dashed line represents chance (0.5). Error bars represent 95% Confidence Intervals.

TABLE 1 | Values for one sample t-tests from Experiment 2.

Experiment 2 one sample *T*-test comparison against chance

Block	*t*-score	df	*p*-value	Mean	95% CI
1	1.431	43	0.16	0.545	[0.485, 0.605]
2	2.254	43	0.029*	0.591	[0.511, 0.671]
3	3.935	43	< 0.001*	0.641	[0.571, 0.711]
4	3.269	43	0.002*	0.63	[0.550, 0.710]
5	3.348	43	0.002*	0.639	[0.559, 0.719]
6	3.127	43	0.003*	0.625	[0.545, 0.705]

*CI, confidence interval. All conditions were collapsed. Alternative hypothesis specifies the mean is different from chance (0.5). *p < 0.05.*

GENERAL RESULTS

The absence of an effect of Block on the group means indicating no significant change in choice allocation across a session could be interpreted as a lack of learning or preference acquisition. However, since reinforcers followed both alternatives, it was also possible that obtained reinforcement outcomes shaped individual participants' choices across the duration of a session in both Experiments 1 and 2. That is, it is possible that selecting the suboptimal alternative was serendipitously reinforced, thereby increasing the likelihood of selecting it again. Specifically, the reinforcement outcomes following choice behavior may have been what determined future choices (Herrnstein, 1970). To determine whether obtained reinforcement might account for choice allocation, the generalized matching equation was applied (Baum, 1974) to the results of the first and last blocks of Experiments 1 and 2. The generalized matching equation is traditionally applied to choice data when the probabilities of reinforcement have been manipulated for an individual subject. However, Vollmer and Bourret (2000) applied the generalized matching equation to aggregated data, which enabled them to determine whether the choice between two alternatives collectively fit with predictions from obtained reinforcements.

To test the extent to which reinforcement history could account for participants' choice behavior, we summed choice-allocation and obtained-reinforcement data within the First Block and Last Block for each participant. Then, the number of choices to the optimal-shape stimulus was divided by the number of choices to the suboptimal-shape stimulus (response ratio) and the number of reinforcements obtained following the optimal-shape stimulus was divided by the number of reinforcements obtained following the suboptimal-shape stimulus (reinforcement ratio). Data from participants that did not receive any reinforcers after choosing the suboptimal alternative were excluded from these analyses, resulting in the exclusion of six participants from the First Block and one participant from the Last Block in Experiment 1, and two participants from the First Block and three participants from the Last Block of Experiment 2.

Figure 5 shows each participant's response and reinforcement ratios as one data point for the First and Last Blocks of Experiments 1 and 2. Response and reinforcement ratios were

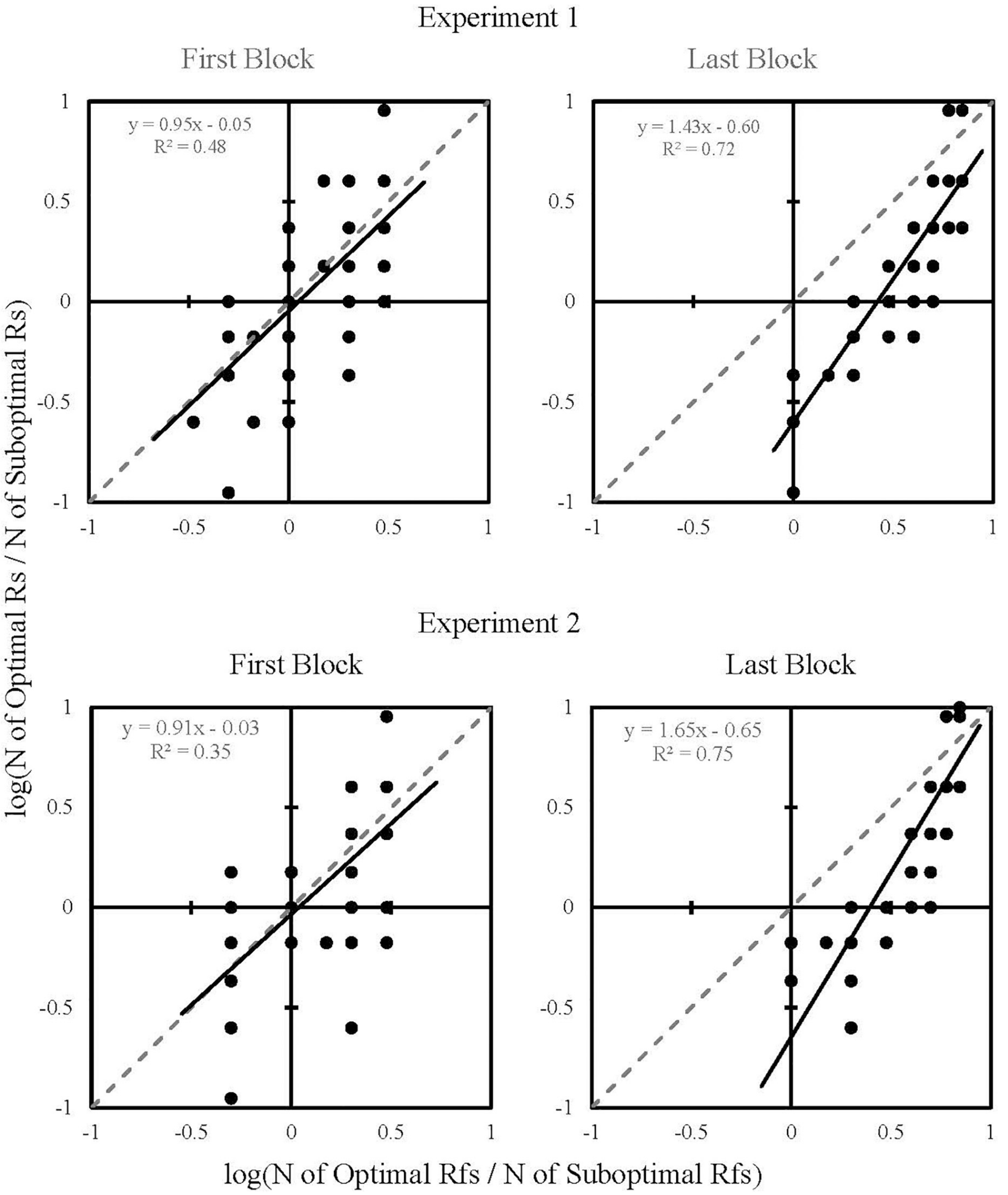


FIGURE 5 | Log response ratios plotted as a function of log reinforcement ratios for First Block and Last Block of Experiments 1 and 2. Each dot represents one participant. The regression line equation and corresponding R^2 values are in the upper-left quadrant of each plot.

logged (base 10) to allow linear regression analysis. The data were fitted to regression lines, the equations of which appear in the upper-left quadrants of each plot. Response ratios significantly correlated with reinforcement ratios in Experiment 1 [First Block, $r(59) = 0.69$, $p < 0.001$; Last Block, $r(64) = 0.85$, $p < 0.001$] and Experiment 2 [First Block, $r(40) = 0.59$, $p < 0.001$; Last Block, $r(39) = 0.87$, $p < 0.001$]. Referring to goodness of fit, it appears that obtained reinforcement became a better predictor of response allocation from First Block [Experiment 1: $R^2 = 0.48$ (RMSE = 0.26); Experiment 2: $R^2 = 0.35$ (RMSE = 0.33)] to Last

Block [Experiment 1: $R^2 = 0.72$ (RMSE = 0.23); Experiment 2: $R^2 = 0.75$ (RMSE = 0.25)], providing further evidence that the participant's choices were influenced by obtained reinforcers, not by the predictiveness of the terminal stimuli.

GENERAL DISCUSSION

The results obtained from Experiments 1 and 2 contribute to the expansion of literature on adult human choice behavior and how it compares to that of non-humans. Perhaps most interesting is that reinforcement predictors seemed to play no role in adult human choice behavior within the scope of this task, which is different than what has been found in pigeons (Stagner and Zentall, 2010). Similar results were found recently with adult human participants in that they also did not show any preference for reinforcement predictors (McDevitt et al., 2019). A closer look at the present results illuminates that adult human choice in this task was driven by reinforcement history rather than reinforcement predictors. Additionally, removing predictive stimuli from the task all together in Experiment 2 had no effect on participants' choice behavior. These findings suggest that reinforcement predictors within this two-alternative task did not play a role in participants' choice allocation.

When comparing the absence of a preference for reinforcement predictors from adult humans in the present studies as well as McDevitt et al. (2019) to that of developmentally delayed children, a clear difference can be observed in choice behavior. That is, adult humans do not show a preference for predictive stimuli (McDevitt et al., 2019) but, under certain conditions, developmentally delayed children do (Lalli et al., 2000). When given a similar two-alternative task, rhesus macaques (Smith et al., 2017) showed a similar preference to that of developmentally delayed children (Lalli et al., 2000). From consideration of these findings from higher-order primates, a quantitative difference in preference for reinforcement predictors emerges across adult humans, developmentally delayed children, and rhesus macaques.

Previous observations from rats suggest that they may choose a suboptimal alternative that provides reinforcement predictors under certain conditions (Chow et al., 2017; Cunningham and Shahan, 2019). However, they do not do so, at least certainly not to the extent to which pigeons do, when given the equivalent task that has been used when observing this behavior in pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011). More recently, pigeons, rats, and dogs were presented with a two-alternative task in which the two alternatives paid out equally but one produced predictive terminal stimuli (Jackson et al., 2020), much like the procedure used by Roper and Zentall (1999). Just as Roper and Zentall (1999) found, pigeons showed a preference for the alternative that produced predictive stimuli. Interestingly, dogs and rats showed no such preference (Jackson et al., 2020). The absence of a preference for predictive stimuli from dogs and sometimes from rats (Jackson et al., 2020) is similar to the present studies' findings as well as those from McDevitt et al. (2019) with human participants.

Taken together, previous findings as well as the present results help to illuminate the value of reinforcement predictors within a two-alternative choice task. Adult human participants (McDevitt et al., 2019) and dogs (Jackson et al., 2020) do not seem to prefer predictive stimuli over non-predictive stimuli. Human children with developmental delays (Lalli et al., 2000), rhesus macaques (Smith et al., 2017), and rats (Chow et al., 2017; Cunningham and Shahan, 2019) show preference for predictive stimuli under certain contextual variables. Pigeons show a strong and clear preference for predictive stimuli (Roper and Zentall, 1999; Stagner and Zentall, 2010; Zentall and Stagner, 2011; McDevitt et al., 2019; Jackson et al., 2020). All of these studies contribute evidence that suggests that there are differences across vertebrate species with respect to preference (or lack thereof) for reinforcement predictors, which is in contrast to Macphail (1985) hypothesis. The greatest disparity is perhaps between the strong preference that pigeons (Stagner and Zentall, 2010; Zentall and Stagner, 2011) show for predictive stimuli and the lack of such a preference observed from adult humans (McDevitt et al., 2019).

The present studies observed adult human preferences for predictive stimuli using a two-alternative task. This task was very similar in format to what pigeons have received in past research (Stagner and Zentall, 2010; Zentall and Stagner, 2011; Stagner et al., 2012; Laude et al., 2014). Adult human participants, unlike pigeons, did not show a preference for stimuli that predict reinforcement. This finding suggests a possible qualitative difference between pigeons and adult humans within this type of two-alternative task. Differences between the two species have also been observed when measuring preferences for stimuli that look like they would produce reinforcement. These stimuli can be thought of as "near hits" in that visually they look very similar to stimuli that are predictive of a win. Slot machines, for example, produce some turns (or trials) that may visually look more like a win than a loss. For example, if there are three reels in a slot machine, two of the three reels would stop on matching stimuli during a "near-hit" trial. This might visually appear closer to a win than if all three reels produce different stimuli—a clear loss. Although these "near-hit" trials are equivalent to clear losses, adult human participants show a preference for slot machines that produce these trials more frequently under conditions in which they have experienced losses (MacLin et al., 2007).

To test for this preference in pigeons, Stagner et al. (2015) gave pigeons a two-alternative task. Choice of one alternative sometimes yielded "near-hit" trials in which a positive stimulus that signaled food would appear but then would change to a negative stimulus signaling the absence of food. Choice of the other alternative resulted in the same amount of food overall, but did not produce "near-hit" trials. Pigeons did not prefer the alternative that produced "near-hit" trials (Stagner et al., 2015) which is different from what was found with human participants (MacLin et al., 2007). Two-alternative tasks like the one used by Stagner et al. (2015) have been suggested to be analogous to human gambling procedures. However, when considering the differences between pigeons and adult human participants on tasks such as this, it must be noted that perhaps this task does not produce behavior that is analogous to human gambling or that it only does so under specific conditions with specific subjects.

When considering Macphail (1985) proposal that there should not be differences between vertebrates, the findings from pigeons and humans with respect to preferences for predictive stimuli are in sharp contrast. When considering the present findings, adult human choice allocation seems more driven by reinforcement history. Conversely, the predictive nature of stimuli seems to be valuable for pigeons at the expense of how often food is actually presented.

Some evidence for preference for predictive stimuli rather than overall amount of reinforcement by pigeons has been observed in the same type of two-alternative task (Stagner et al., 2012). When both alternatives sometimes produced a terminal stimulus that predicted reinforcement 100% of the time, pigeons were indifferent between the two alternatives. What is interesting about this is that while both alternatives were associated with a stimulus that predicted reinforcement, a reinforcement predictor occurred more frequently following choice of one alternative (50% of the time) than following choice of the other alternative (20% of the time). Similar results were also found by Zentall et al. (2015) in that pigeons showed preference for the reliability of a reinforcement predictor independent of its frequency. Considering the findings from Stagner et al. (2012) and Zentall et al. (2015), pigeons appeared to be selecting the alternative that produced the best stimulus predictive of reinforcement within the tasks. However, if pigeons' choice within these tasks was determined by overall reinforcement associated with each alternative, then there would have been more choice allocation to the optimal alternative that provided more reinforcement. The present findings suggest that adult human participants might take a more global view when allocating choices, which is in contrast to the findings from pigeons (Stagner et al., 2012; Zentall et al., 2015).

In addition, future research with adult human participants focusing on the effects of deprivation and depletion may also provide further insight into predictive stimuli preferences. Motivating operations could be explored by systematically manipulating contextual variables, such as food deprivation and social enrichment. Both have been found to have an effect on pigeons' preference for predictive stimuli. Specifically, more food-deprived pigeons chose an alternative that produced predictive stimuli but less reinforcement over time whereas less food-deprived pigeons chose optimally (Laude et al., 2012). Additionally, pigeons that received social enrichment were much slower to show a preference for a predictive suboptimal alternative than were their control counterparts that received no such social enrichment (Pattison et al., 2013). The performance by less food-deprived pigeons (Laude et al., 2012), and early performance by socially enriched pigeons (Pattison et al., 2013), more closely resemble the data that has been collected using adult human participants in the present studies and from McDevitt et al. (2019). In the future, motivating operations could be examined in adult humans to observe if similar results are found.

Considering the present results, pigeons appear to perform differently than rats (Chow et al., 2017; Cunningham and Shahan, 2019), dogs (Jackson et al., 2020), rhesus macaques (Smith et al., 2017), developmentally delayed children (Lalli et al., 2000), and especially adult humans (McDevitt et al., 2019) within this task. When considering the performance of adult humans and pigeons specifically, the extant difference in preference for reinforcement-predictive stimuli could indicate that the two species are fundamentally different. This pervasive finding is in contradiction to the notion that vertebrate species should not differ (Macphail, 1985), and suggests that further examination into factors which account for the choice-allocation differences across vertebrate species is warranted.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board Georgia Southern University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JS, KB, SB, VE, JJ, and RS were part of the development of all procedures used in Experiments 1 and 2. KB, SB, VE, JJ, and RS ran statistical analyses and constructed figures used. KB programed procedures for Experiments 1 and 2 using OpenSesame. SB, VE, JJ, and RS collected the data. JS wrote this manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors would like to acknowledge Erin Baum, Melvin Marsh, Steven Nichols, and Austin Smith for their help with data collection.

REFERENCES

Baum, W. M. (1974). On two types of deviation from the matching law: bias and undermatching. *J. the Exper. Analy. Behav.* 22, 231–242. doi: 10.1901/jeab.1974.22-231

Belke, T. W., and Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: preference for unreliable reinforcement. *J. Exper. Analys. Behav.* 62, 353–366. doi: 10.1901/jeab.1994.62-353

Chow, J. J., Smith, A. P., Wilson, A. G., Zentall, T. R., and Beckmann, J. S. (2017). Suboptimal choice in rats: incentive salience attribution promotes maladaptive decision-making. *Behav. Brain Res.* 320, 244–254. doi: 10.1016/j.bbr.2016.12.013

Cunningham, P. J., and Shahan, T. A. (2019). Rats engage in suboptimal choice when the delay to food is sufficiently long. *J. Exper. Psychol. Anim. Learn. Cogn.* 45, 301–310. doi: 10.1037/xan0000211

Herrnstein, R. J. (1970). On the law of effect. *J. Exper. Analys. Behav.* 13, 243–266. doi: 10.1901/jeab.1970.13-243

Jackson, S. M., Gordon, K. L., MacPherson, K., Martin, G. K., and Roberts, W. A. (2020). Information preferences across species: pigeons, rats, and dogs. *Behav. Process.* 170, 1–8. doi: 10.1016/j.beproc.2019.104016

JASP Team (2019). *JASP (Version 0.11.1) [Computer Software]. RRID:SCR_015823.* Available online at: http://jasp-stats.org (accessed February 24, 2020).

Lalli, J. S., Mauro, B. C., and Mace, F. C. (2000). Preference for unreliable reinforcement in children with mental retardation: The role of conditioned reinforcement. *J. Appl. Behav. Analys.* 33, 533–544. doi: 10.1901/jaba.2000.33-533

Laude, J. R., Pattison, K. F., and Zentall, T. R. (2012). Hungry pigeons make suboptimal choices, less hungry pigeons do not. *Psychonom. Bull. Rev.* 19, 884–891. doi: 10.3758/s13423-012-0282-2

Laude, J. R., Stagner, J. P., and Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *J. Exper. Psychol. Anim. Learn. Cogn.* 40, 12–21. doi: 10.1037/xan0000010

MacLin, O. H., Dixon, M. R., Daugherty, D., and Small, S. L. (2007). Using a computer simulation of three slot machines to investigate a gambler's preference among varying densities of near-miss alternatives. *Behav. Res. Methods* 39, 237–241. doi: 10.3758/BF03193153

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. B* 308, 37–51. doi: 10.1098/rstb.1985.0008

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324. doi: 10.3758/s13428-011-0168-7

McDevitt, M. A., Diller, J. W., and Pietrzykowski, M. O. (2019). Human and pigeon suboptimal choice. *Learn. Behav.* 47, 334–343. doi: 10.3758/s13420-019-00391-8

Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., and Zentall, T. R. (2012). Decision making by humans in a behavioral task: do humans, like pigeons, show suboptimal choice? *Learn. Behav.* 40, 439–447. doi: 10.3758/s13420-012-0065-7

Pattison, K. F., Laude, J. R., and Zentall, T. R. (2013). Environmental enrichment affects suboptimal, risky, gambling-like choice in pigeons. *Anim. Cogn.* 16, 429–434. doi: 10.1007/s10071-012-0583-x

Roper, K. L., and Zentall, T. R. (1999). Observing behavior in pigeons: the effect of reinforcement and probability and response cost using a symmetrical choice procedure. *Learn. Motiv.* 30, 201–220. doi: 10.1006/lmot.1999.1030

Smith, T. R., and Beran, M. J. (2020). Outcome expectancy and suboptimal risky choice in nonhuman primates. *Learn. Behav.* doi: 10.3758/s13420-019-00406-4 [Epub ahead of print].

Smith, T. R., Beran, M. J., and Young, M. E. (2017). Gambling in rhesus macaques (*macaca mulatta*): the effect of cues signaling risky choice outcomes. *Learn. Behav.* 45, 288–299. doi: 10.3758/s13420-017-0270-5

Stagner, J. P., Case, J. P., Sticklen, M. F., Duncan, A. K., and Zentall, T. R. (2015). Do pigeons prefer alternatives that include near-hit outcomes? *J. Exper. Psychol. Anim. Learn. Cogn.* 41, 247–254. doi: 10.1037/xan000069

Stagner, J. P., Laude, J. R., and Zentall, T. R. (2012). Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and the number of presentations of the conditioned reinforcer. *J. Exper. Psychol. Anim. Behav. Process.* 38, 446–452. doi: 10.1037/a0030321

Stagner, J. P., and Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonom. Bull. Rev.* 17, 412–416. doi: 10.3758/PBR.17.3.412

Vollmer, T. R., and Bourret, J. (2000). An application of the matching law to evaluate the allocation of two- and three-point shots by college basketball players. *J. Appl. Behav. Analys.* 33, 137–150. doi: 10.1901/jaba.2000.33-137

Zentall, T. R. (2014). Suboptimal choice by pigeons: an analog of human gambling behavior. *Behav. Process.* 103, 156–164. doi: 10.1016/j.beproc.2013.11.004

Zentall, T. R. (2016). An animal model of human gambling. *Intern. J. Psychol. Res.* 9, 96–112. doi: 10.21500/20112084.2284

Zentall, T. R., Laude, J. R., Stagner, J. P., and Smith, A. P. (2015). Suboptimal choice by pigeons: evidence that the value of the conditioned reinforce rather than its frequency determines choice. *Psychol. Record* 65, 223–229. doi: 10.1007/s40732-015-0119-2

Zentall, T. R., and Stagner, J. P. (2011). Maladaptive choice behaviour by pigeons: an animal analogue and possible mechanism for gambling (sub-optimal human decision-making behaviour). *Proc. R. Soc. B* 278, 1203–1208. doi: 10.1098/rspb.2010.1607

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

MINI REVIEW
published: 06 August 2020
doi: 10.3389/fpsyg.2020.01954

Working Memory as an Indicator for Comparative Cognition – Detecting Qualitative and Quantitative Differences

Lukas Alexander Hahn and Jonas Rose*

Neural Basis of Learning, Institute of Cognitive Neuroscience, Faculty of Psychology, Ruhr University Bochum, Bochum, Germany

Working memory (WM), the representation of information held accessible for manipulation over time, is an essential component of all higher cognitive abilities. It allows for complex behaviors that go beyond simple stimulus-response associations and inflexible behavioral patterns. WM capacity determines how many different pieces of information (items) can be used for these cognitive processes, and in humans, it correlates with fluid intelligence. As such, WM might be a useful tool for comparison of cognition across species. WM can be tested using comparatively simple behavioral protocols, based on operant conditioning, in a multitude of different species. Species-specific contextual variables that influence an animal's performance on a non-cognitive level are controlled by adapting the WM paradigm. The neuronal mechanisms by which WM emerges in the brain, as sustained neuronal activity, are comparable between the different species studied (mammals and birds), as are the areas of the brain in which WM activity can be measured. Thus WM is comparable between vastly different species within their respective niches, accounting for specific contextual variables and unique adaptations. By approaching the question of "general cognitive abilities" or "intelligence" within the animal kingdom from the perspective of WM, the complexity of the core question at hand is reduced to a fundamental memory system required to allow for complex cognitive abilities. This article argues that measuring WM can be a suitable addition to the toolkit of comparative cognition. By measuring WM on a behavioral level and going beyond behavior to the underlying physiological processes, qualitative and quantitative differences in cognition between different animal species can be identified, free of contextual restraints.

Keywords: Macphail, null-hypothesis, quantifying cognition, models of working memory, comparative cognition

Edited by:
Thomas Zentall,
University of Kentucky, United States

Reviewed by:
William Albert Roberts,
The University of Western Ontario, Canada
Mathias Osvath,
Lund University, Sweden

****Correspondence:***
Jonas Rose
jonas.rose@rub.de;
jonas.rose@ruhr-uni-bochum.de

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *29 February 2020*
Accepted: *15 July 2020*
Published: *06 August 2020*

Citation:
Hahn LA and Rose J (2020) Working Memory as an Indicator for Comparative Cognition – Detecting Qualitative and Quantitative Differences. Front. Psychol. 11:1954. doi: 10.3389/fpsyg.2020.01954

INTRODUCTION

In his article on vertebrate intelligence, Macphail (1985) argues that there are no qualitative differences between vertebrate species when it comes to their cognitive abilities. His major line of reasoning builds on the success of the "Columban Simulation Project" to reproduce experiments performed with primates, using pigeons (Epstein, 1981, 1986; Epstein et al., 1984).

While quantitative differences can be described, e.g., by inspecting the required amount of training to master a task (Scarf and Colombo, 2020), the claim for a lack of qualitative difference appears to be more robust. Macphail identifies contextual variables [species-specific experimental conditions, a notion also investigated by Bitterman (1975)] as the source of perceived qualitative differences amongst vertebrates. Neutralizing them would reveal cognitive abilities to be identical.

A prominent approach for comparative cognition features experiments that have shown the impressive abilities of higher cognition of primates, and apply these experiments to other species. Amongst those are protocols that train the animal to apply an abstract rule to a novel set of stimuli. This is the case for numerical competence (Brannon and Terrace, 1998) and orthographic processing (Grainger et al., 2012), to name but two examples. Other experiments focus on an innate cognitive ability that should be present without training a specific response. Famous examples are the mirror self-recognition test (Gallup, 1970) and experiments testing “theory of mind” (Hare et al., 2001). While it is not impossible to train animals other than primates on these tasks (Reiss and Marino, 2001; Plotnik et al., 2006; Clayton et al., 2007; Prior et al., 2008; Scarf et al., 2011, 2016), the experiments might contain insurmountable methodological hurdles for some vertebrate species (e.g., the task may require more training than is possible in animals with short life-span or hands to manipulate objects or operate a touch-screen). These hurdles may prevent an animal from performing successfully in the task, sometimes leading to ambivalent results (Call and Tomasello, 2008; de Waal, 2019). Bitterman focused on qualitative differences of cognition regarding learning (Bitterman, 1975). For example, monkeys and rats maximized reward in probability learning, pigeons and fish showed probability matching (Bitterman et al., 1958; Bullock and Bitterman, 1962b; Wilson et al., 1964; Woodard and Bitterman, 1973). For habit reversal pigeons and rats showed progressive improvement while fish did so only under specific circumstances (Bitterman et al., 1958; Bullock and Bitterman, 1962a; Engelhardt et al., 1973). Together these tests represent a large family (also including tests of inhibitory control, object permanence, and social cognition) that are a set of tools with which complex cognitive behavior can be described and its intricacies can be investigated and disentangled. We suggest adding a different approach to this family. One that investigates a fundamental trait of cognition on a physiological level that can be compared between species. To that end, we suggest working memory (WM) as a critical component of many higher cognitive functions. This addition is warranted by findings of comparative behavioral science and neuroscience. Despite their long independent evolution (Hedges, 2002) resulting in vastly different brain architectures, the cognitive abilities of mammals and birds are very similar, a case of convergent evolution (Colombo and Scarf, 2012; Güntürkün, 2012; Güntürkün and Bugnyar, 2016; Nieder, 2017). WM has been investigated on a behavioral and a physiological level in birds and compared to mammalian WM (Veit and Nieder, 2013; Ditz and Nieder, 2015; Balakhonov and Rose, 2017; Rinnert et al., 2019; Fongaro and Rose, 2020). The bottom line under all these investigations is that besides the different organizations of their brains the same fundamental processes take place. A comparison of different species at this physiological level of WM would widen the scope of comparative cognition and would allow testing Macphail’s idea that focusses primarily on learning, in a new way. Differences and similarities in WM (e.g., in its capacity) may offer insights into why some animals may be (un-)able to display certain cognitive behaviors. Macphail’s null-hypothesis can thus be investigated in the light of potentially qualitative, and quantitative differences of a fundamental trait of cognition.

Using WM comes with its own set of challenges: a precise definition, its concrete link with higher cognition, a precise measurement of the process, and control of the testing environment are all required to ensure a comparable metric. If those challenges are met, WM can be a powerful tool to determine quantitative or even qualitative differences in cognition amongst vertebrates.

A DEFINITION OF WORKING MEMORY

The concept of WM was originally devised by Baddeley and Hitch (1974), developed from the earlier models of short term memory of Broadbent (1958) and Atkinson and Shiffrin (1968). Fundamentally, WM is the process of holding limited information accessible for a limited time. Importantly, this maintenance is controlled by executive processes that also allow for the manipulation of this information (Honig, 1978; Baddeley, 2000; Cowan, 2014). A common test in humans is the n-back task (Conway et al., 2005), here participants memorize sequentially presented numbers and indicate if a number is the repetition of the number presented n numbers earlier. A typical test in the animal literature is the delayed match to sample task: An animal has associated two stimuli, for instance, the colors red and blue, with food. At the start of a trial, it sees the color red, which then disappears from view. After a delay, the animal has the choice between a red and a blue food bowl. It chooses the red bowl (using WM) because the condition of the trial (the red color) determines which bowl is baited. The blue bowl, equally associated with food, is not chosen because the information held in WM allows for a goal selection based on the task demands on the current trial. This general protocol has been used in numerous experiments and variations with different species (e.g., Weinstein, 1941; Lu et al., 1993; Zhang et al., 2005; Bloch et al., 2019). The neuroscientific literature interprets physiologically sustained activity during the delay as a correlate of WM, other physiological processes notwithstanding. This concept of “active memory” has shaped our understanding of this cognitive process (Fuster and Alexander, 1971; Funahashi et al., 1989; Goldman-Rakic, 1995; Miller et al., 1996; Fuster, 2000). Sustained activity, of which persistent neuronal spiking is a simple form, differentiates between samples and thereby encodes them as information. The presence or decay of this information during the delay is correlated with correct and incorrect matching of the sample, respectively. The amount of information encoded by neuronal activity is quantifiable, and significant differences between correct and error trials exist that indicate information loss (Brody et al., 2003; Nieder, 2012; Jacob and Nieder, 2014; Veit et al., 2014; Moll and Nieder, 2015). For this article, we will consider the active physiological process of sustained activity during the delay to be the "fingerprint of WM" (**Box 1**). Its presence can be unequivocally detected by neuronal recordings from animals performing a task. Such sustained activity has been recorded in mammals and birds (e.g., Fuster and Alexander, 1971; Funahashi et al., 1989; Diekamp et al., 2002; Baeg et al., 2003; Veit and Nieder, 2013; Tsutsui et al., 2016; Wu et al., 2020). The similarity of the WM fingerprint in these different species is indicative of a common mechanism (Veit and Nieder, 2013). This definition of WM narrowly focusses on one aspect of WM, it can only explain WM as an effect under equally narrow experimental circumstances (Miller et al., 2018). We chose to focus on this simple physiological definition, with all its limitations, to facilitate comparability between species. The definition is limited to maintenance of information for a short time, and cannot be used to differentiate the many possibilities of how successful behavioral performance emerges in a WM task (Zentall and Smith, 2016), nor can it tease apart the many intricacies of WM function at the physiological level (Miller et al., 2018). But by using an appropriate experimental setup, the physiological measure at the center of our definition is robust and can be controlled for influences on WM (**Box 2**). This

BOX 1 | Different approaches of measuring WM.

There are two different approaches for measuring WM in behavioral experiments, focusing on different aspects of the concept. In the first approach, WM is operationalized as both temporary storage of information (in a range from seconds to hours) and executive functions for retrieval and manipulation of this information (Baddeley, 1992; Floresco and Phillips, 2001; Phillips et al., 2004). This means WM contains accessible information, up to hours after the information has been initially available, and even if the animal, at acquisition, did not know it would require the information later on. The approach is commonly tested in rodents (most prominently rats) and on occasion in other species (e.g., fish and pigeons). Animals are usually tested in a spatial context and over retention intervals ranging from a few seconds to hours (e.g., Olton and Samuelson, 1976; Roitblat et al., 1982; Roberts and Van Veldhuizen, 1985). The tasks the animals perform often consist of navigation in an open area or a radial arm maze in search of food. Efficient navigation of the maze requires memorizing which parts have already been visited (and hence are depleted of food) and visiting only those that have not yet been visited still containing food.

In the second approach, WM is measured as a form of actively sustained short-term memory. In this case, the animal knows that the information will be behaviorally relevant soon and has to hold it accessible and subject it to manipulation, during and after a short delay, in the range of a few seconds (Goldman-Rakic, 1995). Monkeys and birds (but also rodents) are usually tested in tasks during which actively sustained WM can be measured, requiring the animal to attend to a stimulus, remember it and perform a discriminative choice based on the retained stimulus after a short delay of a few seconds (e.g., monkeys: (Fuster and Alexander, 1971), pigeons: (Diekamp et al., 2002), crows: (Veit et al., 2014), mice: (Wu et al., 2020), rats: (Baeg et al., 2003; Tsutsui et al., 2016). The active memory component in these tasks bridges the temporal gap in which the stimulus is not present and holds the information accessible. The absence of the stimuli defines the duration of WM.

These approaches can be challenging to disentangle at the behavioral level since they build on the same definition of working memory introduced by Baddeley and Hitch (1974). However, it is important to keep in mind that these approaches can imply different neural mechanisms. Active maintenance was never demonstrated, and seems counterintuitive, for maintenance that lasts several hours, but the term WM has been used for such long delays. Likewise, many animal studies of WM utilizing a delayed matching to sample task do not directly demonstrate an executive component, which makes it difficult to distinguish behaviorally between short term and working memory. We favor a definition of WM as the cognitive effort of actively keeping stimulus information in an accessible state that can be manipulated for cognitive processes while the stimulus is not physically present. This definition implies a testable neural fingerprint (sustained physiological activity), short duration, susceptibility to interference (**Box 2**), and limited capacity. These are all aspects that allow for a quantification of WM in different species while the definitive neural fingerprint, active maintenance, can provide a qualitative test if *this WM* is present in a given species.

BOX 2 | Influences on WM.

Proactive and retroactive interference:

When testing WM, both proactive (e.g., Grant, 1975) and retroactive interference can occur. Our definition makes the measuring of WM robust to such interference. The physiological trace is informative about which stimulus sample is encoded, thus stimuli from both sources of interference can be differentiated in terms of cause for memory failure.

Encoding of preceding sample stimuli is reflected in the activity of individual neurons that show specific sustained activity during the delay that corresponds to correct and incorrect choice behavior. For example, neuron N has elevated sustained activity during the delay following the presentation of sample A and baseline activity following the presentation of sample B. Another neuron AN shows sustained activity following B but only baseline activity following A. On correct trials with sample A, N has sustained activity, AN does not, and the animal matches the sample. On error trials, with sample A, N has baseline activity and AN has sustained activity, subsequently the animal mismatches the sample by choosing B. This has been shown in both monkeys and crows (Brody et al., 2003; Nieder, 2012; Veit et al., 2014; Moll and Nieder, 2015). The same holds on a more abstract level for neurons encoding different behavioral rules like "match" or "non-match," instead of purely sensory stimuli (Wallis et al., 2001; Veit and Nieder, 2013). In this way, if e.g., trial one was correct with sample A, trial two was correct with sample B, and trial three was incorrect with sample A, the possible interference of sample B with WM of sample A can be detected. This also fits the conclusion of Grant about prospective interference that "the retention deficit in pigeon STM is the product of competition between the prior, conflicting memory and the current memory at the time of the Trial 2 test" (Grant, 1975, p. 214).

The effect of retroactive interference on WM during the delay (in the form of distractors) has been detected in the physiological trace of neurons. During the presentation of the distractor, information about the sample (i.e., differential neuronal activity) diminishes, and the distractor is encoded in neuronal activity (Miller et al., 1996; Jacob and Nieder, 2014). In the following delay, sample information recovers, while distractor information is not sustained (Miller et al., 1996; Jacob and Nieder, 2014). Memory failure (mismatching the sample) correlates with the decay of sample information following the distractor presentation, while the information about the distractor does not have any influence (Jacob and Nieder, 2014). Additionally, if instead of a distractor the sample is repeated, information about the sample increases (Jacob and Nieder, 2014).

Differential behavior during the delay:

Differential behavior (e.g., Zentall et al., 1978) can be used as a strategy to avoid the use of WM as we define it. Its presence should thus result in the absence of the WM trace (sustained activity), while behavioral performance solves the task. Subsequently, if there is no WM trace during the delay of the task, the experimental parameters must be adjusted to prohibit alternative behavior-mediated strategies.

is essential for the comparative approach, to ensure that the measurements in different animal species are always of the raw information-storage, quantifying WM abilities.

How is the definition of WM, and its measurement, informative about general cognitive abilities? Cognition requires a system that processes information to produce behavior. We argue that "the neuron remains the important unit of function for developing a rational account of how behavior is generated" (Barlow, 1995). The mechanisms underlying WM (Miller et al., 2018) are essential for the maintenance and processing of sensory stimuli and the generation of action plans ("executive control") that are the foundation of flexible behavior. One key aspect of WM is its capacity, as it determines how much information is simultaneously available for processing. By comparing this capacity, we aim to understand, if cognitive abilities are based on the same fundamental resources, or if already on this basic level a divergence occurs that limits some species' cognitive abilities. Measuring the capacity and complexity of WM could serve as a proxy for measuring the complexity of cognition in general. Indeed, using a battery of different tests, Kolata et al. (2005) quantified the learning and success rate of mice and found that "general cognitive ability" co-varied with spatial WM capacity. In humans too, WM capacity is correlated with fluid intelligence (Cowan et al., 2005; Fukuda et al., 2010; Luck and Vogel, 2013).

THE ROLE OF CONTEXTUAL VARIABLES IN WORKING MEMORY TASKS

A precise measurement of the underlying process is required to compare the cognitive abilities of different animal species. This implies that "contextual variables" must not influence the results. A contextual variable is the result of an interaction between species and the test environment. Examples are the motivational state of the animal for an available reward, sensory demands of relevant stimuli, and motoric demands of the behavioral task (Bitterman, 1965; Macphail, 1985). Removing the influence of contextual variables on an animal's performance is, therefore, essential for comparative research. For example, a food reward used to motivate a pigeon may not motivate a monkey to work in an otherwise identical task, leading to vastly different performances (Macphail, 1985) argued that such contextual variables are the parsimonious explanation for observed performance differences, instead of underlying cognitive differences. Hence, in this example, the appropriate food reward for each animal is required to motivate the same behavioral and cognitive process. Saturation is another variable that determines the motivational drive (Bitterman et al., 1958) found that for their fish, food intake and the number of days of food deprivation are positively correlated, with a few days of deprivation already showing a strong effect on intake, comparable to what has been found for water deprivation in rats (Stellar and Hill, 1952). By using non-differential reward, motivation should only be affected in a general matter (e.g., by saturation). Unfortunately, it's not always as simple as controlling saturation or switching the food reward. Animals whose ecology is based around active foraging for long periods can be trained comparatively easy because the task design is similar to naturally occurring foraging. Pigeons, corvids, rats, and monkeys are part of this group. Other animals cannot be motivated as readily because testing them in isolation, in what is an ecologically untypical task for them, is prohibitive for any performance in the task. Rewards commonly used in DMTS tasks, small amounts of food or water, might be unsuitable to elicit any kind of motivation (a snake who may only eat once a month and actively hunts for its food might be a striking example for such issues). This means that the reward for matching behavior needs to be adapted to elicit a motivational drive in the tested species, ideally in multiple trials back-to-back. Social or environmental variables that are rewarding to the animal might be an alternative to food-based rewards (e.g., for a snake, escaping a cold place to enter a warm place might be rewarding; for a fish leaving a current and entering still water might be rewarding, etc.). This requires precise knowledge of the animal's ecological background and creative experimental design to ensure that the animal can only use the sample information to solve the task.

Many tasks that test cognitive abilities were designed for primates and make use of their specific abilities (e.g., manipulations of objects, touchscreens, long periods of training, etc.). Due to contextual variables, the translation of such task designs to the needs of other species can be very challenging. Pigeons, for instance, require many more learning trials than monkeys until they perform equally well on many tasks (Scarf and Colombo, 2020), even though a behavioral protocol might be well established for pigeons. It might, therefore, prove to be virtually impossible to train species like fish on such tasks, entirely because a non-cognitive trait, like a limited trial number due to quick saturation, prevents task acquisition. These tests of impressive higher cognitive abilities are, therefore, often difficult to compare between different species, simply because the tested animal does not produce the cognitive trait at all. The issue might even arise at a more basic level. The lack of a hand to manipulate things, or eyes to see with, will dramatically alter the way an animal cognitively engages with the environment. This raises the question if the WM measures of DMTS tasks, performed with samples of different sensory modalities, can even be compared in a meaningful way. Sensory specialization is commonplace and testing an animal within a sensory dimension it is adapted to is a prerequisite to investigate its WM capacity. This is exemplified by using different modalities in the same species that will yield different performances in DMTS tasks. These differences can be explained by sensory discriminability that is required for differential encoding. A pigeon might not be able to differentiate between two odors, while a rat easily can. However, no meaningful information about the sample can be memorized when the upcoming choice between alternatives does not allow for differentiation. By investigating WM, as code for information about samples, the cognitive process is reduced to the differential translation of sensory input into a neuronal representation. This is a very simple form of cognition that we assume to be present irrespective of sensory-motor adaptation. How much information can be encoded at the same time thus becomes a "pure" capacity for comparison. Such information can even be independent of its sensory origin, exemplified by the neuronal activity of monkeys and crows, where neurons encode the same meaning, rather than just the sensory identity, of stimuli (Wallis et al., 2001; Veit and Nieder, 2013; Moll and Nieder, 2017). By measuring WM abilities, we might be able to quantify differences in cognitive abilities, using a unified testing regime that overcomes hurdles imposed by contextual variables. Such context-free WM abilities allow testing Macphail's null-hypothesis (Macphail, 1985), supporting it with WM being similar in animals, or disproving it with differences in WM.

ADEQUATELY MEASURING WM

The delayed matching to sample (DMTS) task, originally introduced by Blough (1959), has become a benchmark for investigating memory processes (Zentall and Smith, 2016). In DMTS tasks, an animal has to attend to a behaviorally relevant stimulus ("sample"), and following a delay, in which the sample is not present, select a matching stimulus from an array of multiple stimuli. Alternatively, it is also possible that a non-matching stimulus has to be selected (delayed non-match to sample, DNMTS) for an emphasis on control processes (manipulation of information). Both experimental designs require the animal to encode the sample, actively maintain its representation throughout the delay, and choose the matching stimulus (**Box 3**). Therefore, WM can be measured in the delay period, when the stimulus is not present, isolating the process of active maintenance from sensory processes. This describes the concept of DMTS. The way this task is implemented can differ based on species-specific requirements (see below), making the DMTS suited for many different species. For instance, it can be spatially distributed, with the sample in one location and the match in another, where the delay is defined by the animal covering the distance between sample stimulus and choice array. Or on a touchscreen by attending to the sample and after a delay, matching it in a stimulus-array. The DMTS task alone primarily tests the short term memory of an animal. A component guiding behavior may add the requirement for manipulation of this information in combination with the stimulus, like a rule instructing a response to "same" or "different," depending on stimulus identity.

We suggest two measures to quantify WM. The capacity of WM (how many items can be stored simultaneously) is indicative of how much available information can be integrated by a cognitive process. This measure can be influenced by chunking individual pieces of information into larger units (Miller, 1956; Cowan, 2001; Cowan et al., 2005). Our approach seeks to quantify WM at the level of differential items to elucidate if differences in capacity exist. Thus, experimental conditions that prohibit chunking are required. A possible outcome of testing raw capacity is to find great similarity between species, as is the case for monkeys and crows (Buschman et al., 2011; Balakhonov and Rose, 2017). Further experiments with conditions that allow chunking may then reveal if all animals can make use of this cognitive strategy. However, successfully testing chunking in non-human animals comes with its own set of challenges (Terrace, 2001; Mathy and Feldman, 2012), especially if a comparative approach is aimed for. Information decay is an indicator of WM performance. This can be measured directly by investigating the neuronal signal, comparing information at the start of the delay (after stimulus offset) to the end of the delay (Buschman et al., 2011). The retention of the sample through a delay needs to be part of the initial training to ensure that the animal learns to maintain the information (Dorrance et al., 2000). Short delays (one to 2 s) facilitate testing of different species because the burden on WM is kept as small as possible while still allowing for meaningful quantification of information. The key to comparability is that sustained activity is present. Short delays that facilitate successful performance, allow for species comparison of the amount of retained information at the end of the delay. Delays can subsequently be prolonged to see how information decays, resulting in task failure. If prolonged retention (e.g., Grant, 1976) is based on sustained activity, a comparison of information at different time points of the delay is possible between animals. If the delay activity is not sustained (but performance is), this would be indicative of other processes not covered by our narrow definition of WM. This can also be tested by adding components to the task that interfere with WM but not with, for example, long term memory (e.g., familiarity Brady and Hampton, 2018). Additional experiments investigating this different process would then be required. By appropriately adapting a DMTS task for different animals, quantification of WM can become a suitable tool to compare cognitive abilities between individuals of one species [as has been shown to work in mice by Kolata et al. (2005)] and ultimately between different species. Measuring WM with a DMTS task is very suitable to neutralize the role contextual variables may play. This is based on three attributes of the test regime: (1) The DMTS task is usable with stimuli of different modalities (visual, olfactory, auditory), thereby eliminating species-specific sensory demands.

BOX 3 | The "nature" of sustained activity.

Reward coding signals, motor preparation signals, and retrospective (encoding the memorandum) or prospective (e.g., encoding the identity of the response stimulus, or the response location) content can be encoded by sustained activity. These different interpretations can be dissociated through the experimental design of the DMTS task on both a behavioral level (Zentall and Smith, 2016) and on a physiological level. Motor related signals (i.e., planning on where to respond) can be controlled for by randomizing the response location or limiting it to one kind of response. To avoid that the sustained delay activity reflects differential reward [as it does for monkeys and pigeons (Watanabe, 1996; Browning et al., 2011; Johnston et al., 2017a)], a common-outcome procedure is sufficient, for which sustained neuronal activity then encodes the sample stimulus (Johnston et al., 2017b). If every memorandum of the DMTS task is represented with the same amount of information (i.e., by the same type of representation and all stimuli are equally easy to differentiate from one another), comparing WM capacity between species can be based on this information being encoded in the neuronal representation of WM. Knowing how this information is represented, as sample identity, reward code, or any other form of code, is not required for the measurement of capacity. Decoding neuronal activity during the delay reveals the amount of information in WM (Buschman et al., 2011) and this measure is to be compared between species.

(2) The task duration can be adapted to match the ecological and ethological time scale of the tested animal, which overcomes hurdles of temporal scaling (e.g., a bird pecking on a touchscreen responds a lot faster than a fish swimming to an answer location). (3) The choice the animal has to make during the matching can be adapted according to the animal's abilities to indicate a decision (e.g., by touching a choice key, navigating to a choice position, etc.). By adapting the DMTS task to a species specialized adaptations (e.g., using rats' excellent sense of smell, instead of their comparatively poor vision), tests can make use of the ethological repertoire of the animal instead of forcing it into producing disjointed conditioned responses.

WORKING MEMORY OF DIFFERENT SPECIES: SIMILARITIES, DIFFERENCES, AND COMPARISONS

There is an abundance of studies investigating WM in different species, Lind et al. (2015), list such studies. The primary concern of this article is how the results of the individual studies may support a claim for Macphail's null-hypothesis in principle. To that end, studies have been selected that fulfill the following criteria. The WM tested in the animal is considered "active" in the sense of this text (i.e., the maintained information is accessible for manipulation while it is not physically available). The experimental design is suitable for adaptation for other animal species (i.e., contextual variables can be neutralized). Alternative explanations of performance can be excluded (e.g., associative memory, stereotypical responses, etc.).

Due to the relative ease of adapting the task design, the WM abilities of monkeys are easiest to compare to humans'. Quantification of WM performance can be assessed with delay length, the number of training trials, and capacity (Weinstein, 1941) showed that rhesus monkeys successfully perform a DMTS, using objects at delay lengths of 5, 10, and 15 s, for a sample size of one. This study simultaneously also tested two young, pre-verbal children in the same task. Both species learned to perform at virtually the same level for all delay lengths, but humans took far fewer trials to learn the procedure. This might reflect a quantitative difference between the species. The number of trials to reach a defined performance threshold in a DMTS task is a good measure for this quantitative difference (Scarf and Colombo, 2020) have suggested the same metric when comparing the performance of monkeys to pigeons. The capacity of WM at stable delay length was investigated by Buschman et al. (2011), who showed in a DNMTS task that macaques perform successfully for short delay lengths of 800–1000 ms with up to five samples. A marked drop in performance occurred at five items, indicating a capacity of about four items, strikingly similar to the famous "magical number four" of human WM (Cowan, 2001; Buschman et al., 2011).

The DMTS and DNMTS protocols are also used in rats and mice. Rats have been shown to successfully learn to discriminate and match stimuli in the visual domain (Mumby et al., 1990; Prusky et al., 2004), but in these experiments, additional factors may interfere with the measurement of WM. Unlike monkeys, rats and mice have poor vision, so visual stimuli are most likely not adequate for testing WM in comparable ways. Additional features of task design (the novelty of stimuli, olfactory cues, and object recognition processes) require special attention to ensure that WM is adequately measured and compared to what is being measured in other species (Ennaceur, 2010). Fortunately, adapting DMTS tasks from the visual to the olfactory domain resolves these issues. When different odors are used as sample and choice stimuli, both rats and mice can perform at very high levels of asymptotic performance and show degrading performance as a function of delay length (Lu et al., 1993; Liu et al., 2014; Roddick et al., 2014).

The range of species is not limited to mammals. Birds were successfully tested in WM paradigms. The physiological correlate of WM has been described in pigeons performing a visual go-no-go task (Diekamp et al., 2002), where pigeons had to maintain an instructive color across a delay and match a behavioral response. Direct comparisons of monkeys and pigeons performing the same task have been performed, using a change detect paradigm (Leising et al., 2013; Wright and Elmore, 2016). The results indicate that there is no major difference between the species. A second bird species adds to these findings. Crows' WM has been investigated in the visual domain in combination with abstract rules, here too the physiological process of WM has been recorded (Veit and Nieder, 2013; Veit et al., 2014; Balakhonov and Rose, 2017) have explicitly compared WM capacity in crows to monkeys reproducing the task of Buschman et al. (2011) and were able to show that crows and monkeys show the same capacity dependent function of performance, reaching a plateau performance at about four items (Buschman et al., 2011; Balakhonov and Rose, 2017). Overall, the results of WM studies in birds and monkeys indicate virtually identical physiological processes and behavioral performance amongst these two groups of animals (Colombo and Scarf, 2012; Güntürkün, 2012; Güntürkün and Bugnyar, 2016), a result congruent with behavioral observations of cognitive abilities and indicative of convergent evolution (Emery and Clayton, 2004; Jarvis et al., 2005).

Amongst "fish" (a paraphyletic group of animals), few species were tested on their WM abilities. Recently, zebrafish (*Danio rerio*) have been tested in a DMTS task, using different colors with a delay of three and 4 s (Bloch et al., 2019). With their study, Bloch and colleagues established an experimental setup that allows fish to be tested, overcoming contextual hurdles for an animal species that is notoriously difficult to train on a behavioral task. A substantial constraint is the meager amount of trials that can be performed by the fish in a session [in the study of Bloch et al. (2019), only ten trials]. The fish were considered to have learned the task only if they performed at a level of $\geq$70% correct in a session, for three consecutive sessions, to account for this low number of trials. This study nicely overcomes contextual variables impeding comparative cognition between fish and other animals, by using the DMTS task. Similarly, WM abilities can be quantified even outside of the vertebrate clade. Bees (*Apis mellifera*) can learn to match visual and odor samples after a delay and are able to successfully transfer to novel stimuli (Giurfa et al., 2001; Zhang et al., 2005). Performance of fish and bees is reported to be lower than what can be found in other vertebrates and the length of the delay strongly influences the performance, nonetheless, the principle holds.

The results of mammals, birds, fish, and bees show that WM can be comparably measured across species. Even when they have vastly different organizations of their brains, vastly different ecological niches, and vastly different contextual specialties. Comparisons across species have to be considered carefully nonetheless. WM capacity, retention time, and length of training (measured in trials to criterion) are valuable indicators that allow us to compare the vastly different species sensibly. But these metrics are themselves not completely context-free. The capacity of WM is subject to two competing models of resource allocation, discrete (Luck and Vogel, 1997; Awh et al., 2007) and continuous (Alvarez and Cavanagh, 2004; Bays and Husain, 2008). Depending on which model is being tested in a given experiment, the capacity estimate of WM might vary, based on the applied method of measurement (Fukuda et al., 2010; Luck and Vogel, 2013). Similarly, the quantified amount of training can only be compared if species-specific attributes are taken into account. For example, a singular trial for the pigeon (who performs several hundred per session) is relatively speaking less relevant for performance than a singular trial for a fish (with only ten trials per session). Careful normalization within species may resolve such issues. A final open issue that remains is that WM has not (yet) been shown on the physiological level for all species and thus can only be inferred from the task design if the same "active memory" system is tested in all instances. Ultimately WM can help us resolve the difficulties of application of the famous studies developed for primates, for other vertebrates. With WM and the DMTS task to measure it, we can actively quantify differences in a basic component of all higher cognition and add to the analysis of the principles of animal cognition.

CONCLUSION

To understand animal cognition and to investigate Macphail's null-hypothesis, different approaches can offer insight. Tackling higher cognitive abilities with complex tasks can produce milestones indicative of qualitative differences, while a focus on fundamental aspects of cognition, like WM, with simple tasks, allows us to recognize the quantitative scaling of abilities. The relative simplicity of WM allows us to quantify an animal's cognitive ability with a unified testing paradigm (the DMTS/DNMTS) that is adaptable to the species of interest, overcoming methodological and contextual hurdles imposed by the complex tasks. Many vastly different species from different classes of vertebrates (mammals, birds, and ray-finned fish) have been successfully tested in the DMTS task, and it is even applicable for an invertebrate like the bee. This is a vital step toward the goal of comparative cognition. A physiological definition, such as active memory, can offer an additional tool beyond behavior to quantify cognition. Importantly, this physiological approach offers precise criteria for comparison along with tools to analyze the underlying processes not only qualitatively but also quantitatively. There is still a lack of physiological evidence of WM in many species (like fish and bees) that unequivocally shows this basic neural process. However, the addition of avian electrophysiology concerning WM has produced results that strongly support the idea that active WM is a universal neural process amongst vertebrates. Concerning Macphail's null-hypothesis, we conclude that, on the level of WM ability, there does not appear to be a qualitative difference between different vertebrate species. On the quantitative side, differences between species are detectable. The WM of fish and bees seems to be more limited when compared to mammals and birds. This is indicated by the relative difficulty that comes with training them. Data is, however, still lacking and a comprehensive evaluation of WM in different groups of vertebrates, using the same tests and systematically measuring WM capacity, and retention decay of information, along with the physiological correlates of WM, needs to be performed to conclude whether Macpahil's null-hypothesis can be disproven on this most basic level of cognition.

AUTHOR CONTRIBUTIONS

LH and JR conceptualized and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

Freigeist Fellowships awarded to JR by the Volkswagen Foundation.

REFERENCES

Alvarez, G. A., and Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15, 106–111. doi: 10.1111/j.0963-7214.2004.01502006.x

Atkinson, R. C., and Shiffrin, R. M. (1968). "Human memory: a proposed system and its control processes," in *Psychology of Learning and Motivation*, eds K. W. Spence and J. T. Spence (Cambridge, MA: Academic Press), 89–195. doi: 10.1016/s0079-7421(08)60422-3

Awh, E., Barton, B., and Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* 18, 622–628. doi: 10.1111/j.1467-9280.2007.01949.x

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *Psychology of Learning and Motivation*, ed. G. H. Bower (Cambridge, MA: Academic Press), 47–89.

Baeg, E. H., Kim, Y. B., Huh, K., Mook-Jung, I., Kim, H. T., and Jung, M. W. (2003). Dynamics of population code for working memory in the prefrontal cortex. *Neuron* 40, 177–188. doi: 10.1016/S0896-6273(03)00597-X

Balakhonov, D., and Rose, J. (2017). Crows rival monkeys in cognitive capacity. *Sci. Rep.* 7:8809. doi: 10.1038/s41598-017-09400-0

Barlow, H. (1995). "The neuron doctrine in perception," in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Cambridge, MA: The MIT Press), 415–435.

Bays, P. M., and Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science* 321, 851–854. doi: 10.1126/science.1158023

Bitterman, M. E. (1965). The evolution of intelligence. *Sci. Am.* 212, 92–101.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science* 188, 699–709. doi: 10.1126/science.188.4189.699

Bitterman, M. E., Wodinsky, J., and Candland, D. K. (1958). Some comparative psychology. *Am. J. Psychol.* 71, 94–110. doi: 10.2307/1419199

Bloch, S., Froc, C., Pontiggia, A., and Yamamoto, K. (2019). Existence of working memory in teleosts: establishment of the delayed matching-to-sample task in adult zebrafish. *Behav. Brain Res.* 370, 111924. doi: 10.1016/j.bbr.2019.111924

Blough, D. S. (1959). Delayed matching in the pigeon. *J. Exp. Anal. Behav.* 2, 151–160. doi: 10.1901/jeab.1959.2-151

Brady, R. J., and Hampton, R. R. (2018). Nonverbal working memory for novel images in Rhesus monkeys. *Curr. Biol.* 28, 3903.e–3910.e. doi: 10.1016/j.cub.2018.10.025

Brannon, E. M., and Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by Monkeys. *Science* 282, 746–749. doi: 10.1126/science.282.5389.746

Broadbent, D. E. (1958). *Perception and Communication*. Elmsford, NY: Pergamon Press.

Brody, C. D., Hernández, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cereb. Cortex* 13, 1196–1207. doi: 10.1093/cercor/bhg100

Browning, R., Overmier, J. B., and Colombo, M. (2011). Delay activity in avian prefrontal cortex – sample code or reward code? *Eur. J. Neurosci.* 33, 726–735. doi: 10.1111/j.1460-9568.2010.07540.x

Bullock, D. H., and Bitterman, M. E. (1962a). Habit reversal in the pigeon. *J. Comp. Physiol. Psychol.* 55, 958–962. doi: 10.1037/h0041070

Bullock, D. H., and Bitterman, M. E. (1962b). Probability-matching in the Pigeon. *Am. J. Psychol.* 75, 634–639. doi: 10.2307/1420288

Buschman, T. J., Siegel, M., Roy, J. E., and Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11252–11255. doi: 10.1073/pnas.1104666108

Call, J., and Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192. doi: 10.1016/j.tics.2008.02.010

Clayton, N. S., Dally, J. M., and Emery, N. J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 507–522. doi: 10.1098/rstb.2006.1992

Colombo, M., and Scarf, D. (2012). Neurophysiological studies of learning and memory in pigeons. *Comp. Cogn. Behav. Rev.* 7, 23–43. doi: 10.3819/ccbr.2012.70002

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/BF03196772

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114. doi: 10.1017/S0140525X01003922

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educ. Psychol. Rev.* 26, 197–223. doi: 10.1007/s10648-013-9246-y

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cogn. Psychol.* 51, 42–100. doi: 10.1016/j.cogpsych.2004.12.001

de Waal, F. B. M. (2019). Fish, mirrors, and a gradualist perspective on self-awareness. *PLoS Biol.* 17:e3000112. doi: 10.1371/journal.pbio.3000112

Diekamp, B., Kalt, T., and Güntürkün, O. (2002). Working memory neurons in pigeons. *J. Neurosci.* 22:RC210.

Ditz, H. M., and Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7827–7832. doi: 10.1073/pnas.1504245112

Dorrance, B. R., Kaiser, D. H., and Zentall, T. R. (2000). Event-duration discrimination by pigeons: the choose-short effect may result from retention-test novelty. *Anim. Learn. Behav.* 28, 344–353. doi: 10.3758/BF03200268

Emery, N. J., and Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306, 1903–1907. doi: 10.1126/science.1098410

Engelhardt, F., Woodard, W. T., and Bitterman, M. E. (1973). Discrimination reversal in the goldfish as a function of training conditions. *J. Comp. Physiol. Psychol.* 85, 144–150. doi: 10.1037/h0034879

Ennaceur, A. (2010). One-trial object recognition in rats and mice: methodological and theoretical issues. *Behav. Brain Res.* 215, 244–254. doi: 10.1016/j.bbr.2009.12.036

Epstein, R. (1981). On pigeons and people: a preliminary look at the Columban simulation project. *Behav. Anal.* 4, 43–55. doi: 10.1007/BF03391851

Epstein, R. (1986). "Simulation research in the analysis of behavior," in *Research Methods in Applied Behavior Analysis: Issues and Advances*, eds A. Poling and R. W. Fuqua (Boston, MA: Springer), 127–155. doi: 10.1007/978-1-4684-8786-2_7

Epstein, R., Kirshnit, C. E., Lanza, R. P., and Rubin, L. C. (1984). 'Insight' in the pigeon: antecedents and determinants of an intelligent performance. *Nature* 308, 61–62. doi: 10.1038/308061a0

Floresco, S. B., and Phillips, A. G. (2001). Delay-dependent modulation of memory retrieval by infusion of a dopamine D1 agonist into the rat medial prefrontal cortex. *Behav. Neurosci.* 115, 934–939. doi: 10.1037/0735-7044.115.4.934

Fongaro, E., and Rose, J. (2020). Crows control working memory before and after stimulus encoding. *Sci. Rep.* 10, 1–10. doi: 10.1038/s41598-020-59975-4

Fukuda, K., Vogel, E., Mayr, U., and Awh, E. (2010). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychon. Bull. Rev.* 17, 673–679. doi: 10.3758/17.5.673

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349. doi: 10.1152/jn.1989.61.2.331

Fuster, J. M. (2000). Executive frontal functions. *Exp. Brain Res.* 133, 66–70. doi: 10.1007/s002210000401

Fuster, J. M., and Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science* 173, 652–654. doi: 10.1126/science.173.3997.652

Gallup, G. G. (1970). Chimpanzees: self-recognition. *Science* 167, 86–87. doi: 10.1126/science.167.3914.86

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., and Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature* 410, 930–933. doi: 10.1038/35073582

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., and Fagot, J. (2012). Orthographic processing in Baboons (Papio papio). *Science* 336, 245–248. doi: 10.1126/science.1218152

Grant, D. S. (1975). Proactive interference in pigeon short-term memory. *J. Exp. Psychol. Anim. Behav. Process.* 1, 207–220. doi: 10.1037/0097-7403.1.3.207

Grant, D. S. (1976). Effect of sample presentation time on long-delay matching in the pigeon. *Learn. Motiv.* 7, 580–590. doi: 10.1016/0023-9690(76)90008-4

Güntürkün, O. (2012). The convergent evolution of neural substrates for cognition. *Psychol. Res.* 76, 212–219. doi: 10.1007/s00426-011-0377-9

Güntürkün, O., and Bugnyar, T. (2016). Cognition without cortex. *Trends Cogn. Sci.* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Anim. Behav.* 61, 139–151. doi: 10.1006/anbe.2000.1518

Hedges, S. B. (2002). The origin and evolution of model organisms. *Nat. Rev. Genet.* 3, 838–849. doi: 10.1038/nrg929

Honig, W. K. (1978). "Studies of working memory in the Pigeon," in *Cognitive Processes in Animal Cognition*, eds S. H. Hulse, H. Fowler, and W. K. Honig (New Jersey, NJ: Lawrence Erlbaum Associates, Inc.), 211–248. doi: 10.4324/9780203710029-8

Jacob, S. N., and Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron* 83, 226–237. doi: 10.1016/j.neuron.2014.05.009

Jarvis, E. D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159. doi: 10.1038/nrn1606

Johnston, M., Anderson, C., and Colombo, M. (2017a). Neural correlates of sample-coding and reward-coding in the delay activity of neurons in the entopallium and nidopallium caudolaterale of pigeons (*Columba livia*). *Behav. Brain Res.* 317, 382–392. doi: 10.1016/j.bbr.2016.10.003

Johnston, M., Anderson, C., and Colombo, M. (2017b). Pigeon NCL and NFL neuronal activity represents neural correlates of the sample. *Behav. Neurosci.* 131, 213–219. doi: 10.1037/bne0000198

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., and Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiol. Learn. Mem.* 84, 241–246. doi: 10.1016/j.nlm.2005.07.006

Leising, K. J., Elmore, L. C., Rivera, J. J., Magnotti, J. F., Katz, J. S., and Wright, A. A. (2013). Testing visual short-term memory of pigeons (*Columba livia*) and a rhesus monkey (*Macaca mulatta*) with a location change detection task. *Anim. Cogn.* 16, 839–844. doi: 10.1007/s10071-013-0644-9

Lind, J., Enquist, M., and Ghirlanda, S. (2015). Animal memory: a review of delayed matching-to-sample data. *Behav. Process.* 117, 52–58. doi: 10.1016/j.beproc.2014.11.019

Liu, D., Gu, X., Zhu, J., Zhang, X., Han, Z., Yan, W., et al. (2014). Medial prefrontal activity during delay period contributes to learning of a working memory task. *Science* 346, 458–463. doi: 10.1126/science.1256573

Lu, X.-C. M., Slotnick, B. M., and Silberberg, A. M. (1993). Odor matching and odor memory in the rat. *Physiol. Behav.* 53, 795–804. doi: 10.1016/0031-9384(93)90191-H

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390:279. doi: 10.1038/36846

Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400. doi: 10.1016/j.tics.2013.06.006

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 37–51. doi: 10.1098/rstb.1985.0008

Mathy, F., and Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition* 122, 346–362. doi: 10.1016/j.cognition.2011.11.003

Miller, E. K., Erickson, C. A., and Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the Macaque. *J. Neurosci.* 16, 5154–5167. doi: 10.1523/JNEUROSCI.16-16-05154.1996

Miller, E. K., Lundqvist, M., and Bastos, A. M. (2018). Working memory 2.0. *Neuron* 100, 463–475. doi: 10.1016/j.neuron.2018.09.023

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158

Moll, F. W., and Nieder, A. (2015). Cross-modal associative mnemonic signals in crow endbrain neurons. *Curr. Biol.* 25, 2196–2201. doi: 10.1016/j.cub.2015.07.013

Moll, F. W., and Nieder, A. (2017). Modality-invariant audio-visual association coding in crow endbrain neurons. *Neurobiol. Learn. Mem.* 137, 65–76. doi: 10.1016/j.nlm.2016.11.011

Mumby, D. G., Pinel, J. P. J., and Wood, E. R. (1990). Nonrecurring-items delayed nonmatching-to-sample in rats: a new paradigm for testing nonspatial working memory. *Psychobiology* 18, 321–326. doi: 10.3758/BF03327250

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11860–11865. doi: 10.1073/pnas.1204580109

Nieder, A. (2017). Inside the corvid brain—probing the physiology of cognition in crows. *Curr. Opin. Behav. Sci.* 16, 8–14. doi: 10.1016/j.cobeha.2017.02.005

Olton, D. S., and Samuelson, R. J. (1976). Remembrance of places passed: spatial memory in rats. *J. Exp. Psychol. Anim. Behav. Process.* 2, 97–116. doi: 10.1037/0097-7403.2.2.97

Phillips, A. G., Ahn, S., and Floresco, S. B. (2004). Magnitude of dopamine release in medial prefrontal cortex predicts accuracy of memory on a delayed response task. *J. Neurosci.* 24, 547–553. doi: 10.1523/JNEUROSCI.4653-03.2004

Plotnik, J. M., de Waal, F. B. M., and Reiss, D. (2006). Self-recognition in an Asian elephant. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17053–17057. doi: 10.1073/pnas.0608062103

Prior, H., Schwarz, A., and Güntürkün, O. (2008). Mirror-induced behavior in the Magpie (Pica pica): evidence of self-recognition. *PLoS Biol.* 6:e202. doi: 10.1371/journal.pbio.0060202

Prusky, G. T., Douglas, R. M., Nelson, L., Shabanpoor, A., and Sutherland, R. J. (2004). Visual memory task for rats reveals an essential role for hippocampus and perirhinal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 101, 5064–5068. doi: 10.1073/pnas.0308528101

Reiss, D., and Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5937–5942. doi: 10.1073/pnas.101086398

Rinnert, P., Kirschhock, M. E., and Nieder, A. (2019). Neuronal correlates of spatial working memory in the endbrain of crows. *Curr. Biol.* 29, 2616.e4–2624.e4. doi: 10.1016/j.cub.2019.06.060

Roberts, W. A., and Van Veldhuizen, N. (1985). Spatial memory in pigeons on the radial maze. *J. Exp. Psychol. Anim. Learn. Cogn.* 11, 241–260. doi: 10.1037/0097-7403.11.2.241

Roddick, K. M., Schellinck, H. M., and Brown, R. E. (2014). Olfactory delayed matching to sample performance in mice: sex differences in the 5XFAD mouse model of Alzheimer's disease. *Behav. Brain Res.* 270, 165–170. doi: 10.1016/j.bbr.2014.04.038

Roitblat, H. L., Tham, W., and Golub, L. (1982). Performance of Betta splendens in a radial arm maze. *Anim. Learn. Behav.* 10, 108–114. doi: 10.3758/BF03212055

Scarf, D., Boy, K., Reinert, A. U., Devine, J., Güntürkün, O., and Colombo, M. (2016). Orthographic processing in pigeons (*Columba livia*). *Proc. Natl. Acad. Sci. U.S.A.* 113, 11272–11276. doi: 10.1073/pnas.1607870113

Scarf, D., and Colombo, M. (2020). Columban simulation project 2.0: numerical competence and orthographic processing in pigeons and primates. *Front. Psychol.* 10:3017. doi: 10.3389/fpsyg.2019.03017

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334, 1664–1664. doi: 10.1126/science.1213357

Stellar, E., and Hill, J. H. (1952). The rat's rate of drinking as a function of water deprivation. *J. Comp. Physiol. Psychol.* 45, 96–102. doi: 10.1037/h0062150

Terrace, H. S. (2001). "Chunking & serially organized behavior in pigeons, monkeys and humans," in *Avian Visual Cognition*, ed. H. G. Cook (Medford, MA: Comparative Cognition Press).

Tsutsui, K.-I., Oyama, K., Nakamura, S., and Iijima, T. (2016). Comparative overview of visuospatial working memory in Monkeys and Rats. *Front. Syst. Neurosci.* 10:99. doi: 10.3389/fnsys.2016.00099

Veit, L., Hartmann, K., and Nieder, A. (2014). Neuronal correlates of visual working memory in the corvid endbrain. *J. Neurosci.* 34, 7778–7786. doi: 10.1523/jneurosci.0612-14.2014

Veit, L., and Nieder, A. (2013). Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nat. Commun.* 4:11. doi: 10.1038/ncomms3878

Wallis, J. D., Anderson, K. C., and Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature* 411, 953–956. doi: 10.1038/35082081

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632. doi: 10.1038/382629a0

Weinstein, B. (1941). Matching-from-sample by rhesus monkeys and by children. *J. Comp. Psychol.* 31, 195–213. doi: 10.1037/h0063449

Wilson, W. A., Oscar, M., and Bitterman, M. E. (1964). Visual probability-learning in the monkey. *Psychon. Sci.* 1, 71–72. doi: 10.3758/BF03342795

Woodard, W. T., and Bitterman, M. E. (1973). Further experiments on probability learning in goldfish. *Anim. Learn. Behav.* 1, 25–28. doi: 10.3758/BF03198993

Wright, A. A., and Elmore, L. C. (2016). Pigeon visual short-term memory directly compared to primates. *Behav. Process.* 123, 84–89. doi: 10.1016/j.beproc.2015.09.002

Wu, Z., Litwin-Kumar, A., Shamash, P., Taylor, A., Axel, R., and Shadlen, M. N. (2020). Context-dependent decision making in a premotor circuit. *Neuron* 106, 316.e6–328.e6. doi: 10.1016/j.neuron.2020.01.034

Zentall, T. R., Hogan, D. E., Howard, M. M., and Moore, B. S. (1978). Delayed matching in the pigeon: effect on performance of sample-specific observing responses and differential delay behavior. *Learn. Motiv.* 9, 202–218. doi: 10.1016/0023-9690(78)90020-6

Zentall, T. R., and Smith, A. P. (2016). Delayed matching-to-sample: a tool to assess memory and other cognitive processes in pigeons. *Behav. Process.* 123, 26–42. doi: 10.1016/j.beproc.2015.07.002

Zhang, S., Bock, F., Si, A., Tautz, J., and Srinivasan, M. V. (2005). Visual working memory in decision making by honey bees. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5250–5255. doi: 10.1073/pnas.0501440102

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

MINI REVIEW
published: 07 August 2020
doi: 10.3389/fpsyg.2020.01843

Absolute Numerosity Discrimination as a Case Study in Comparative Vertebrate Intelligence

*Andreas Nieder**

Animal Physiology Unit, Institute of Neurobiology, University of Tübingen, Tübingen, Germany

The question of whether some non-human animal species are more intelligent than others is a reoccurring theme in comparative psychology. To convincingly address this question, exact comparability of behavioral methodology and data across species is required. The current article explores one of the rare cases in which three vertebrate species (humans, macaques, and crows) experienced identical experimental conditions during the investigation of a core cognitive capability – the abstract categorization of absolute numerical quantity. We found that not every vertebrate species studied in numerical cognition were able to flexibly discriminate absolute numerosity, which suggests qualitative differences in numerical intelligence are present between vertebrates. Additionally, systematic differences in numerosity judgment accuracy exist among those species that could master abstract and flexible judgments of absolute numerosity, thus arguing for quantitative differences between vertebrates. These results demonstrate that Macphail's Null Hypotheses – which suggests that all non-human vertebrates are qualitatively and quantitatively of equal intelligence – is untenable.

Keywords: monkey (*Macaca mulatta*), crow, number cognition, categorication, intelligence

OPEN ACCESS

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Martin Giurfa,
UMR 5169 Centre de Recherches sur la Cognition Animale (CRCA), France
Fuat Balcı,
Koç University, Turkey

****Correspondence:***
Andreas Nieder
andreas.nieder@uni-tuebingen.de

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *20 May 2020*
Accepted: *06 July 2020*
Published: *07 August 2020*

Citation:
Nieder A (2020) Absolute Numerosity Discrimination as a Case Study in Comparative Vertebrate Intelligence. Front. Psychol. 11:1843. doi: 10.3389/fpsyg.2020.01843

INTRODUCTION

Intelligence, broadly defined, is the general capacity to solve problems (Macphail, 1987). Whether non-human vertebrate species differ in intelligence remains hotly debated in comparative psychology. After a survey of experimental studies, Macphail (1985) adopted the "null hypothesis" and concluded that no intelligence difference, either qualitative or quantitative, had yet been demonstrated among non-human vertebrates. He argued that the alleged difference in intellect could instead be attributed to a difference in some extraneous "contextual variable," such as species-specific variability in perception, motivation, or motor skills (Macphail, 1985, 1987).

The current article re-examines Macphail's null hypothesis in the realm of numerical competence. Estimating numerosity, the number of items in a set, is a type of abstract categorization that is central to adaptive and intelligent behavior (Miller et al., 2003). In numerical categorization, the specific sensory features of objects or events are irrelevant since what matters is the sheer presence of elements in a set. Because humans and non-human animals share an approximate capability to estimate numerosity (Nieder, 2019) numerosity judgments offer a "window of opportunity" to gain insights into cognitive capabilities in a comparative way across phylogeny.

As pointed out by Macphail (1985, 1987) comparing the performances of different vertebrate species requires commensurable approaches and data sets in order to avoid methodological confounds. This article exploits one of the rare cases in which this requirement is fulfilled; it quantitatively explores absolute numerosity judgments that have been collected under virtually identical experimental conditions in three vertebrate species (humans, macaques, and crows). Evivalent computer-controlled visual task protocols were applied for all three species in the same laboratory environment, minimizing the variability due to task differences that usually hampers comparative behavioral research. Additionally, all three species share an acute visual sense, motivation to learn, drive to perform tasks, and comparable volitional motor dexterity (hand movements in primates, and beak/head movement in birds) that ensure analogous contextual variables. If performance differences surface under these conditions that rule out methodological and contextual variables, they can be explained by true quantitative differences in numerical capabilities as a type of intelligence. Moreover, if such absolute numerosity judgments are only mastered by certain cognitively advanced vertebrates, such as mammals and birds, it stands to reason that qualitative differences in intelligence also exist among vertebrates.

FROM RELATIVE TO ABSOLUTE NUMEROSITY JUDGMENTS

The most intensely studied form of numerical competence in animal cognition are "relative numerosity" judgments (sometimes also termed "numerousness" judgments). Here, an animal's often spontaneous ability to select the numerical quantity that is larger relative to another quantity is tested (Nieder, 2020a). For instance, when choosing between food items (Stancher et al., 2015) or seeking shelter among groups of conspecifics (Agrillo et al., 2008) animals tend to "go for more."

More advanced relative numerosity judgments have been explored in laboratory studies with trained animals. When macaques and pigeons were trained to sequentially choose numerosity displays according to ascending numerical values (e.g., 1–2–3), both species showed an ordinal understanding of numerical quantity by transferring their behavior to novel ranges of numerosities (Brannon and Terrace, 1998; Scarf et al., 2011; Scarf and Colombo, 2020). Nevertheless, judging relative numerosity is probably the simplest form of numerical competence because it does not require a representation of the absolute quantity values.

Many classic studies primarily using rodents trained these animals to detect one and the same specific numerosity as a rewarded conditioned stimulus. For instance, rodents were trained to discriminate two specific numbers of sensory signals (Fernandes and Church, 1982; Davis and Albert, 1986) or to produce one specific number of lever presses to receive a reward (Mechner, 1958; Meck and Church, 1983; Çavdaroğlu and Balcı, 2016). However, rodents and many other vertebrates so far have never been trained to flexibly detect any possible absolute numerosity in random trials. Only if animals can flexibly represent any specific numerosity from any other value do they show absolute numerosity representations. Besides humans, only simian primates (*chimpanzees:* Matsuzawa, 1985; Murofushi, 1997; *rhesus macaque:* Cantlon and Brannon, 2007a; Merten and Nieder, 2009) and selected bird species (*parrot:* Pepperberg, 1994; *pigeons:* Xia et al., 2001; *corvids:* Smirnova et al., 2000; Ditz and Nieder, 2016) have been shown to master flexible absolute numerosity judgments. This suggests qualitative differences in numerical intelligence between species.

Absolute numerosity discriminations have been investigated in different vertebrate species using a delayed match-to-numerosity task (DMNT) (**Figure 1A**; Nieder et al., 2002). In the DMNT, motivated subjects discriminate numerosities that are carefully controlled for non-numerical features for reward (**Figure 1B**). A typical trial in a visual DMNT begins when a variable target numerosity (the sample) is presented on a screen. The subject has to recognize and then memorize the numerosity over a brief delay period. If the same target numerosity (a match) is shown again in the subsequent test phase, the subject is required to respond. However, if a deviant (smaller or larger) numerosity (a non-match) is presented in the test phase, the subject must withhold responding and wait for the next test stimulus, which always is a match. Match and non-match are presented with equal probability of $p = 0.5$. The accuracy of numerosity discrimination performance is calculated by dividing the number of correct responses by the number of total responses (correct plus erroneous responses) for the match and all non-match test stimuli.

Using a DMNT with virtually identical experimental conditions, detailed psychophysical characterization of absolute numerosity representations have been obtained in humans (Merten and Nieder, 2009), rhesus macaques (Nieder and Miller, 2003; Nieder et al., 2006; Merten and Nieder, 2009; Nieder, 2012), and carrion crows (Ditz and Nieder, 2015, 2016, 2020). These data allow us to characterize the subjective representations of numerosity in detail. When both smaller and larger non-match numerosity displays are presented besides the matching target numerosity, the subjects' responses give rise to bell-shaped performance functions (or "probability density functions") (**Figure 1C**). These performance functions represent the likelihood that any number is perceived as being equal to a specific objective target number (typically located at the center of the function). For instance, two monkeys made most mistakes for non-match numerosity adjacent to the target numerosity; only with increasing numerical distance of the non-match numerosities from the target numerosity did the monkeys err less and less, which resulted in the slopes of the bell-shaped performance functions fading away (Nieder and Miller, 2003). Thus, the performance functions graphically indicate a subject's subjective numerical representation of objective numbers.

QUANTIFICATION OF NUMBER DISCRIMINATION ACCURACY

The finding that absolute numerosity discriminations result in performance distributions of some width clearly shows that the non-symbolic discrimination of numerical quantity is an approximate estimation process. Several psychophysical signatures of non-symbolic number representations can be

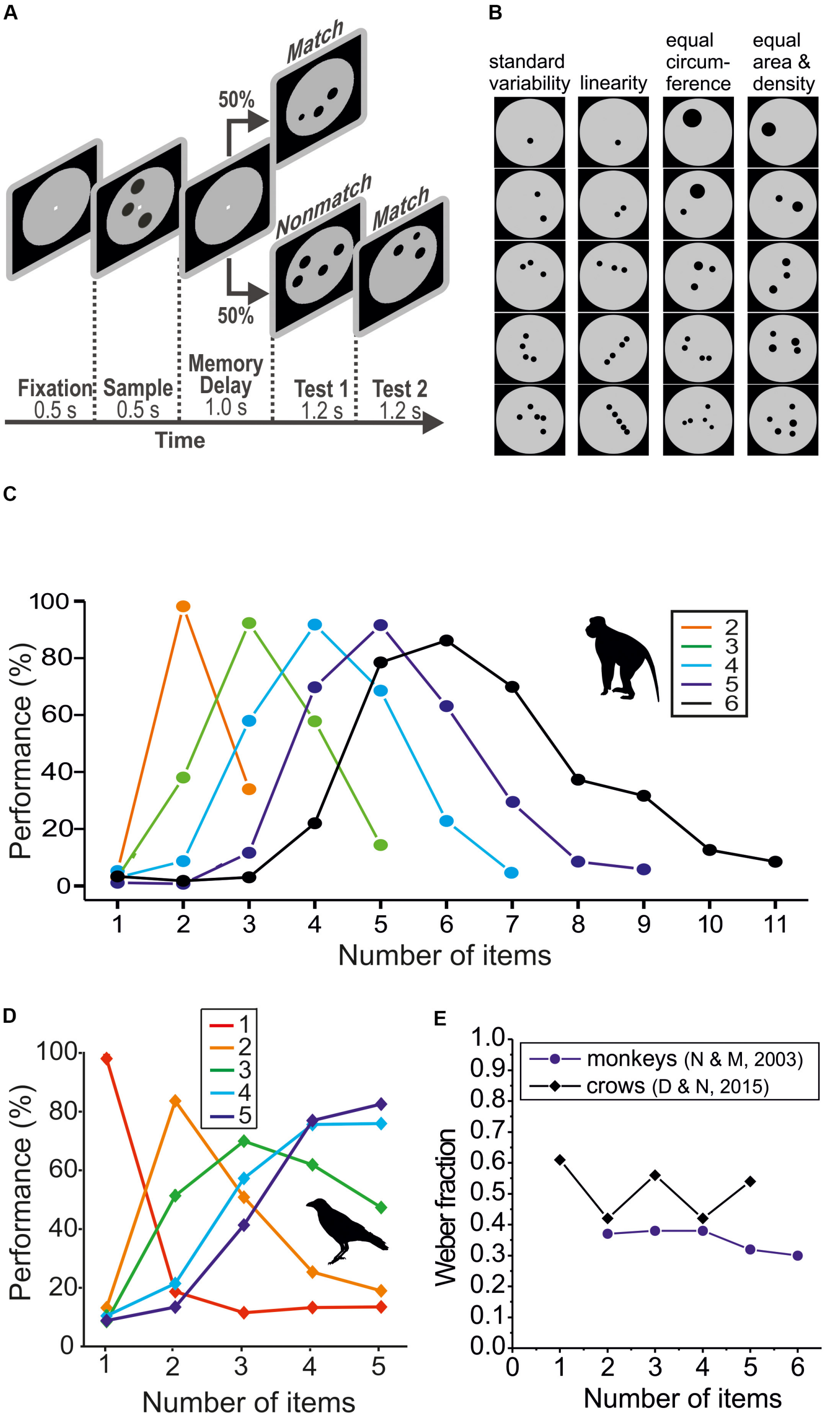


FIGURE 1 | Discrimination performance for simultaneously presented small numerosities. **(A)** Layout of the delayed match-to-numerosity task (DMNT) for dot arrays. **(B)** Example stimulus protocols for numerosity 1–5 that control for different non-numerical parameters. **(C)** Average numerosity performance functions of two rhesus macaques in the DMNT for target numerosity 2–6 (data from Nieder and Miller, 2003). **(D)** Average numerosity performance functions of two carrion crows in the DMNT for target numerosity 1–5 (data from Ditz and Nieder, 2015). **(E)** Weber fractions for small simultaneous-numerosity discriminations of two macaques and two crows. Weber fractions derived from the functions shown in **(C,D)**, respectively.

extracted from these performance functions. First, while similar numerical quantities are difficult to discriminate, discrimination performance systematically improves with increasing difference (or distance) between two quantities; this finding is called "numerical distance effect." Second, discrimination worsens at the same time with increasing magnitudes so that the numerical distance between numerosities must increase in proportion with the absolute magnitudes to enable discrimination; this phenomenon is called the 'numerical size effect.' Both numerical distance and size effects are captured by *Weber's law*. It states that the just-noticeable difference ("JND," ΔI, or "difference limen"; i.e., the stimulus difference that allows 50% correct discrimination) between two magnitudes divided by the reference magnitude, *I*, is a *constant* ($\Delta I/I = c$) (Weber, 1850). The widths of the resulting performance distributions reflect the numerical distance effect, while the progressive broadening of the functions in proportion to increasing magnitude mirror the numerical size effect (**Figure 1C**).

In addition, a third signature surfaces on top of Weber's law: relative to a given reference number, subjects find it easier to discriminate smaller numbers, and more difficult to discriminate larger number (**Figure 1C**). This effect results in performance functions being mildly asymmetric when plotted on a linear number scale (**Figure 1C**). This asymmetry of the performance functions is predicted by Fechner's law which states that the subjective sensation of number, *S*, is proportional to the logarithm of the objective stimulus magnitude, I [$S = k\ log(I)$] (Fechner, 1860). Both Weber's and Fechner's laws hold true in psychophysical assessments of numerosity discriminations across species (Nieder and Miller, 2003; Merten and Nieder, 2009; Ditz and Nieder, 2016). Signatures of Weber's law in numerosity discrimination are a clear sign of an internal "approximate number system" (ANS). The ANS has been found consistently for numerosity judgments in innumerate humans (Gordon, 2004; Pica et al., 2004; Frank et al., 2008) or humans prevented from counting (Whalen et al., 1999; Cordes et al., 2001; Merten and Nieder, 2009), as well as in a multitude of animal species (Nieder, 2020b) from primates (Nieder and Miller, 2003; Cantlon and Brannon, 2006) to bees (Dacke and Srinivasan, 2008; Howard et al., 2018).

To quantify discrimination accuracy, the Weber fraction is calculated. The Weber fraction expresses how much two stimuli need to differ in magnitude in order for a subject to be able to detect a difference between those two stimuli (i.e., "JND" or "difference limen"). Due to the logarithmic relationship that is stated by Fechner's law and has been confirmed experimentally for numerosity discriminations in humans, monkeys, and crows (Nieder and Miller, 2003; Merten and Nieder, 2009; Ditz and Nieder, 2016; Piantadosi and Cantlon, 2017) the JND (and thus the Weber fractions) for numerosities smaller and larger than the target numerosity differ (**Figure 2A**). The $\mathrm{JND_S}$ (n-n_S) for numerosities smaller (n_S) that the target (n) is smaller than the $\mathrm{JND_L}$ (n-n_R) for numerosities larger (n_R) that the target (n). Therefore, the left (toward smaller) and right (toward higher numbers) segments of the performance function need to be calculated separately when plotted on a linear number axis (van Oeffelen and Vos, 1982). Thus, the Weber fraction (W_S) for numerosities smaller than the target is

$$W_\mathrm{S} = (n - n_\mathrm{S})\,/n_\mathrm{S} \tag{1}$$

The Weber fraction ($\mathrm{W_L}$) for numerosities larger than the target is

$$W_\mathrm{L} = (n_\mathrm{L} - n)\,/n \tag{2}$$

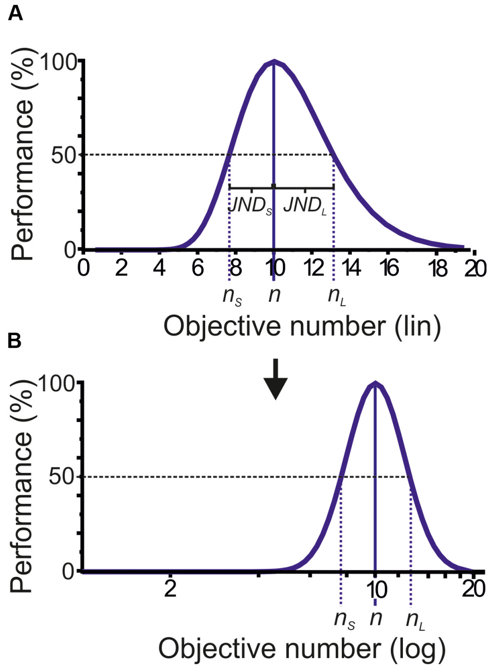


FIGURE 2 | Ideal numerosity performance function. **(A)** Ideal numerosity performance function for target numerosity 10 plotted on a linear number scale (*top graph*). The function shows a steeper slope toward smaller, and a shallower slope toward higher numerosity. As a result, the just-noticeable difference (JND, indicated by dotted colored lines) at which numerosities smaller (n_S) and larger numerosities (n_L) can be discriminated in 50% from the target (n) is smaller on the left compared to the right side of the function. **(B)** When the same function is plotted on a logarithmic number scale, the function becomes symmetric and the JNDs are equal on either side of the function (*bottom graph*).

To arrive at a single Weber-fraction value for a target numerosity, W_S and W_L need to be averaged. Alternatively, the data can be plotted on a logarithmic scale in agreement with Fechner's law, which renders the JND toward smaller and larger numerosities equal (**Figure 2B**). The smaller the Weber fraction, the higher is the discrimination accuracy. With the Weber fraction as an objective measure of discriminability, the judgment of absolute numerosities can be compared quantitatively.

NUMEROSITY DISCRIMINATION ACCURACY WITH SIMULTANEOUSLY PRESENTED ITEMS

By far most studies dealing with non-symbolic numerosity representations have employed item arrays as stimuli (i.e., ∴) (**Figure 1A**). Numerosity stimuli have to be carefully controlled for non-numerical variables because the number of items is intrinsically correlated with many other features of a physical stimulus. For instance, when the number of dots is increased, usually also the total amount of area covered by all dots and the density of the dots increases. Since primates and birds are sensitive to non-numerical magnitudes (Tudusciuc and Nieder, 2010; Moll and Nieder, 2014) a subject could in fact respond to changing item sizes or density rather than numbers. Unfortunately, it is physically not possible to control for all non-numerical factors simultaneously in a single stimulus display. The best way to tackle the problem of non-numerical cues is to control – unbeknown to the subject – one parameter after the other in separate stimulus configurations (**Figure 1B**). If a subject abstracts across these parameters and responds equally to these systematically varied numerosity stimuli, it is safe to conclude that the subject responds to number. The application of such control stimuli demonstrated that the subjects indeed responded to the number or items, not to non-numerical factors (Nieder et al., 2002; Merten and Nieder, 2009; Ditz and Nieder, 2015, 2016).

When simultaneously presented items are scattered across space, they can be assessed at one glance. This is evidenced by monkeys responding with similar reaction times to different simultaneously presented numerical values (Nieder and Miller, 2004b; Merten and Nieder, 2009). As an exception to this pattern, animals usually respond faster to very small numerosities 1 and 2 (Merten and Nieder, 2009). In addition, when the number of items in the displays increased, the monkeys showed the same number of eye movements prior to a decision; they did not scan individual items one after the other before responding. Both findings indicate that non-symbolic estimation of number in dot arrays is a parallel process because serial enumeration would require increasing reaction times with increasing numerical values (Mandler and Shebo, 1982). Thus, the simultaneous number estimation constitutes a specific type of enumeration that differs from a counting-like sequential process.

In initial studies, monkeys (**Figure 1C**) and crows (**Figure 1D**) were required to discriminate small sample numerosities (usually from 1 to 5) from other small numerosities. The average Weber fraction of two rhesus monkeys for sample numerosities 2–5 was 0.36 (+/— 0.03 std) (Nieder and Miller, 2003) which was significantly smaller than the average Weber fraction of 0.49 (+/— 0.07 std) of two carrion crows for the same numerosity range (Ditz and Nieder, 2015) ($p < 0.05$; one-tailed paired t-test; $n = 4$) (**Figure 1E**). Similar small Weber fractions were obtained for a third monkey (see **Figure 3B** in Merten and Nieder, 2009). Thus, for small numerosities, macaques discriminate more precisely than crows.

A similar advantage for primates emerged when larger sample numerosities ranging from 4 to 30 were applied (**Figures 3A–C**). While the performance of two macaques exhibited an average Weber fraction of 0.55 (+/— 0.04 std) (Merten and Nieder, 2009), crows showed a much higher Weber fraction of 1.42 (+/— 0.18 std) (Ditz and Nieder, 2016) ($p < 0.05$; two-tailed paired t-test; $n = 5$) (**Figure 3D**). The dramatically increased large-numerosity Weber fractions of the same two crows that showed smaller values when tested with small numerosities (Ditz and Nieder, 2015) may partly be explained by much larger numerical distances of the non-match numerosities relative to the sample numerosity. In other words, the crows were not forced to discriminate as precisely as in the previous study in which

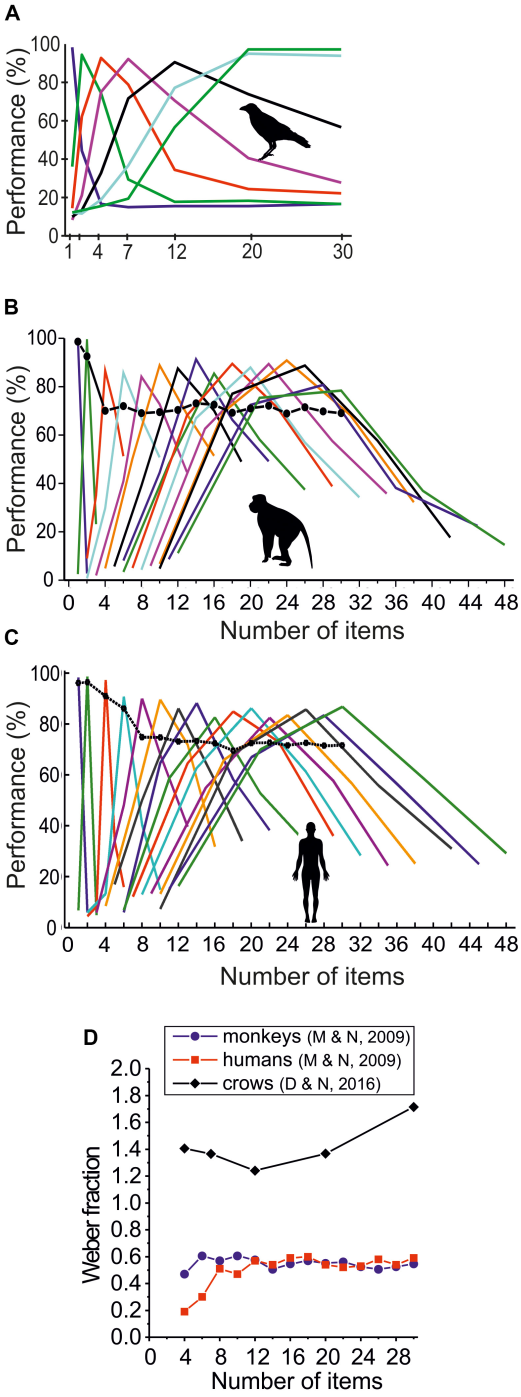


FIGURE 3 | Discrimination performance for simultaneously presented large numerosities. **(A)** Average numerosity performance functions of two carrion crows in the delayed match-to-numerosity task (DMNT) for target numerosity 1–30 (data from Ditz and Nieder, 2016). **(B)** Average numerosity performance functions of two rhesus macaques in the DMNT for target numerosity 1–30 (data from Merten and Nieder, 2009). **(C)** Average numerosity performance functions of 20 humans in the DMNT for target numerosity 1–30 (data from Merten and Nieder, 2009). **(D)** Weber fractions for large simultaneous-numerosity discriminations of two crows, two macaques, and 20 humans. Weber fractions derived from the functions shown in **(A–C)**, respectively.

minimal numerical distances of one between all numerosities were applied (Ditz and Nieder, 2015).

The same study that tested two rhesus macaques also tested 20 adult humans with the same stimuli, apparatus, and protocol (Merten and Nieder, 2009). Due to the rapid presentation of sample and test stimuli, humans were not able to count larger numbers of items symbolically (**Figure 3C**). Interestingly, humans showed the identical Weber fraction of 0.55 (+/− 0.12) as the two monkeys when non-symbolically discriminating numerosities 4–30 ($p > 0.05$; two-tailed paired t-test; $n = 14$). Overall, the data from both small and large numerosity discriminations argue that (human and non-human) primates are more precise when discriminating number in simultaneously presented item arrays.

NUMEROSITY DISCRIMINATION ACCURACY WITH SEQUENTIALLY PRESENTED ITEMS

The concept of numerosity does not only apply for item arrays, but also for items presented over time (**Figure 4A**). If items are presented one after the other in a temporal succession (i.e., ●- ●- ●, etc.), they need to be evaluated in sequence. Although only few studies tested sequential enumeration, it is not only more relevant for the auditory and tactile sense, but also more similar to actual counting, which is a sequential process.

Stimuli testing sequential enumeration need to be carefully controlled for temporal variables because it usually takes longer to present more items. The necessary stimulus configurations that control for a variety of temporal factors have been applied in studies with monkeys and crows. They show that the subjects indeed responded to the number of sequentially presented items, and not to temporal factors (Nieder et al., 2006; Nieder, 2012; Ditz and Nieder, 2020).

Detailed performance data for the enumeration of visual sequences of flashed dots are available for two monkeys (**Figure 4B**; Nieder, 2012) and two crows (**Figure 4C**; Ditz and Nieder, 2020). With an average Weber fraction of 0.31 (+/− 0.17), the two monkeys showed significantly better accuracy then the two crows with a Weber fraction of 0.59 (+/− 0.13) ($p < 0.05$; two-tailed paired t-test; $n = 4$) (**Figure 4D**). Just as with the simultaneous numerosity protocol, monkeys also outperformed crows in the sequential numerosity protocol.

The monkeys' performance is reminiscent of the performance of adult humans in non-symbolic sequential enumeration tasks. When human subjects produce target numbers of key presses at rates that made symbolic counting difficult or impossible, or by preventing them from counting by saying "the" at every press, similar precision was reported. In these human studies, the coefficient of variation (CV, the ratio of the standard deviation and mean) was used as a measure of number discriminability (Whalen et al., 1999; Cordes et al., 2001). On average, the CV of humans was around 0.2.

Even though the CV erroneously assumes symmetric performance distributions and is not directly related to the Weber fraction, we calculated the CV for the same monkey (Nieder, 2012) and crow data (Ditz and Nieder, 2020) from Gauss functions fitted to the sequential performance functions. For sequential enumeration, crows had a much larger average CV of 0.39. However, with a value of 0.19, the monkeys demonstrated a discrimination accuracy almost identical to humans. Just as with simultaneous numerosity protocols, the non-symbolic numerosity discrimination accuracy of humans and monkeys also matches for sequential protocols and surpassed those of crows.

FROM BEHAVIOR TO NEURONS

The controlled DMNT not only allows a detailed characterization of behavioral numerosity representations, but also offers the opportunity of combining behavioral and brain research. Not only does combining controlled behavior with simultaneous neurophysiological recordings give us a direct way to learn about how the brain gives rise to numerical competence, it also allows us a way to derive more objective signatures of cognitive capabilities at the level of the neural substrate.

The neuronal mechanisms of absolute numerosity representations in the endbrains of the three species show an impressive correspondence. A significant proportion of single neurons in the human medial temporal lobe (Kutter et al., 2018) the monkey frontal and parietal association cortices (Nieder et al., 2002; Nieder and Miller, 2004a) and the avian brain region "nidopallium caudolaterale" (NCL) (Ditz and Nieder, 2015, 2016) are tuned to individual preferred numerosities presented simultaneously in dot arrays. This approximate tuning results in peaked neuronal response functions that resemble behavioral performance functions. Just as the behavioral performance functions, the neuronal tuning functions show all the characteristics of the Weber–Fechner law: neurons best discriminate numerosities that are distant from the preferred numerosity (mirroring the distance effect), the neuronal tuning functions become broader with an increase of the neurons' preferred numerosity (a reflection of the size effect), and finally the neuronal tuning functions are best described (i.e., symmetric) on a logarithmic number scale. Numerosity tuning functions showing these characteristics were also indirectly derived through functional imaging in humans (Piazza et al., 2004; Nieder, 2004; Jacob and Nieder, 2009; Kersey and Cantlon, 2017).

This argues that the way in which numerosity-selective neurons encode numerical quantity gives rise to the psychophysical characteristics captured by the Weber–Fechner law. Moreover, the quantitative parameters derived from the neuronal tuning functions, such as the widths of the tuning functions, are comparable between monkeys and crows (Nieder and Miller, 2003; Ditz and Nieder, 2015). All these findings argue that primates and crows engage the same ANS when representing absolute numerosity.

In the human literature, it is hotly debated whether the brain represents numerosity separately for simultaneous versus sequential presentation formats, or abstractly and format-independently. The neuronal data from monkeys and crows both argue for a neuronal two-stage process when

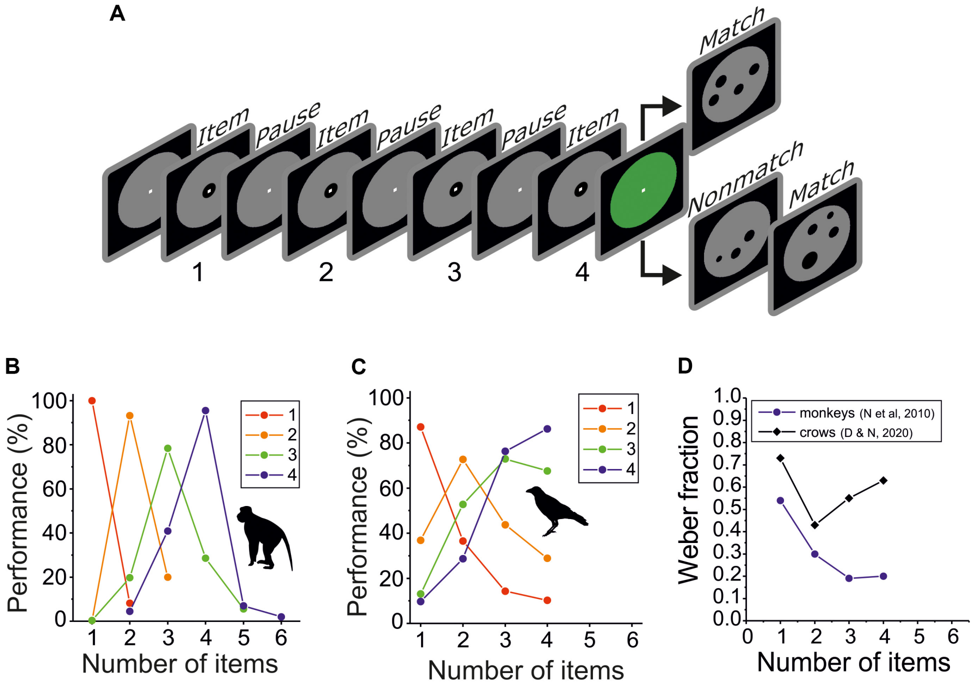


FIGURE 4 | Discrimination performance for sequentially presented numerosity. **(A)** Layout of the delayed match-to-numerosity task (DMNT) for four sequentially presented single dot in the sample period. **(B)** Average numerosity performance functions of two rhesus macaques in the sequential DMNT for target numerosity 1–4 (data from Nieder, 2012). **(C)** Average numerosity performance functions of two carrion crows in the sequential DMNT for target numerosity 1–4 (data from Ditz and Nieder, 2020). **(D)** Weber fractions for small sequential-numerosity discriminations of two macaques and two crows. Weber fractions derived from the functions shown in **(B,C)**, respectively.

these two fundamentally different number formats need to be represented. During the sensory presentation stage, the number of sequentially presented items is extracted by one population of numerosity-tuned neurons, whereas the numerosity in dot arrays is represented by another population of numerosity-tuned neurons (Nieder et al., 2006; Ditz and Nieder, 2020). At this sensory stage of number processing, neurons therefore responded format-dependently. However, once the sensory presentation phase had ended, yet another neuronal population represents numerosity format-independently. This third, format-independent population of neurons maintains numerical information in working memory and also predicts performance success (Nieder et al., 2006; Ditz and Nieder, 2020). In summary, sequential and simultaneous number formats engage different and temporally succeeding populations of format-dependent and format-independent numerosity-selective neurons.

Combining the DMNT with electrophysiological recordings not only provided insights into the behavioral relevance of sensory number representations (Viswanathan and Nieder, 2015), but also enables insights into how numerical information is maintained in working memory and further processed according to behavioral principles (rules) (Cantlon and Brannon, 2007b; Bongard and Nieder, 2010; Vallentin et al., 2012; Eiselt and Nieder, 2013; Cantlon et al., 2016). An in-depth treatment of the neuronal correlates of number representations is beyond the scope of this article concerned with psychophysical results but can be found in recent reviews (Nieder, 2016, 2020b).

CONCLUSION

In his Null Hypotheses, Macphail (1985) suggests that "*neither quantitative nor qualitative differences among the intellects of non-human vertebrates*" existed. The current analyses show that both the quantitative and qualitative aspect of this hypothesis are violated.

The first, quantitative aspect of Macphail (1985) Null Hypotheses proves to be an untenable assertion. As shown in the current review, the three vertebrate species that master elaborate absolute numerosity judgments systematically differ in their precision. The two primate species (humans and monkeys) consistently showed higher (and surprisingly similar) accuracy when discriminating numerosities in a non-symbolic manner. If quantitative differences emerge already for only three investigated vertebrate species, even more pronounced differences can be expected for a broader range of vertebrate species.

In addition, also the second, qualitative aspect of Macphail (1985) Null Hypotheses proves to be an untenable assertion. This is because abstract and flexible judgments of absolute numerosity have so far only been mastered by humans, simian primates and selected bird species, mammalian and avian species that belong to the most cognitively advanced vertebrate classes. This suggests that species from other vertebrate classes (fish, amphibians, and non-avian reptiles) are not capable of flexible absolute numerosity representations. Of course, one may argue that the blank spots of numeracy in the vertebrate phylogenetic tree will be filled with time and more investigations. After all, fish (DeLong et al., 2017) amphibians (Uller et al., 2003), and non-avian reptiles (Gazzola et al., 2018) show relative numerical competence. In fact, some species of teleost fish show unexpected numerical (Miletto et al., 2020) and cognitive skills (Bloch et al., 2019) suggesting that they may also grasp absolute numerosity judgments. However, I predict that amphibians and non-avian reptiles will never master absolute numerosity tasks because they seem to lack the necessary behavioral flexibility (or intelligence) to solve such abstract tasks.

In sum, and in contrast to Macphail's (1985) Null Hypotheses, clear quantitative as well as qualitative differences among the numerical intellects of non-human vertebrates exist. In the field of numerical competence, and most likely also across other cognitive competence, Macphail's Null Hypotheses is untenable.

AUTHOR CONTRIBUTIONS

AN conceptualized and wrote the manuscript.

FUNDING

This work was supported by Deutsche Forschungsgemeinschaft (DFG) grants NI 618/2-1, NI 618/3-1, NI 618/4-1, and NI 618/10-1.

ACKNOWLEDGMENTS

The author thanks Helen Ditz for help with re-analyses of data, and Diana Liao for reading an earlier version of this manuscript.

REFERENCES

Agrillo, C., Dadda, M., Serena, G., and Bisazza, A. (2008). Do fish count? spontaneous discrimination of quantity in female mosquitofish. *Anim. Cogn.* 11, 495–503. doi: 10.1007/s10071-008-0140-9

Bloch, S., Froc, C., Pontiggia, A., and Yamamoto, K. (2019). Existence of working memory in teleosts: establishment of the delayed matching-to-sample task in adult zebrafish. *Behav. Brain Res.* 370:111924. doi: 10.1016/j.bbr.2019.111924

Bongard, S., and Nieder, A. (2010). Basic mathematical rules are encoded by primate prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2277–2282. doi: 10.1073/pnas.0909180107

Brannon, E. M., and Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science* 282, 746–749. doi: 10.1126/science.282.5389.746

Cantlon, J. F., and Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychol. Sci.* 17, 401–406. doi: 10.1111/j.1467-9280.2006.01719.x

Cantlon, J. F., and Brannon, E. M. (2007a). Basic math in monkeys and college students. *PLoS Biol.* 5:e328. doi: 10.1371/journal.pbio.0050328

Cantlon, J. F., and Brannon, E. M. (2007b). How much does number matter to a monkey (*Macaca mulatta*)? *J. Exper. Psychol.* 33, 32–41. doi: 10.1037/0097-7403.33.1.32

Cantlon, J. F., Merritt, D. J., and Brannon, E. M. (2016). Monkeys display classic signatures of human symbolic arithmetic. *Anim. Cogn.* 19, 405–415. doi: 10.1007/s10071-015-0942-5

Çavdaroğlu, B., and Balcı, F. (2016). Mice can count and optimize count-based decisions. *Psychon. Bull. Rev.* 23, 871–876. doi: 10.3758/s13423-015-0957-6

Cordes, S., Gelman, R., Gallistel, C. R., and Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychon. Bull. Rev.* 8, 698–707. doi: 10.3758/bf03196206

Dacke, M., and Srinivasan, M. V. (2008). Evidence for counting in insects. *Anim. Cogn.* 11, 683–689. doi: 10.1007/s10071-008-0159-y

Davis, H., and Albert, M. (1986). Numerical discrimination by rats using sequential auditory stimuli. *Anim. Learn. Behav.* 14, 57–59. doi: 10.3758/bf03200037

DeLong, C. M., Barbato, S., O'Leary, T., and Wilcox, K. T. (2017). Small and large number discrimination in goldfish (*Carassius auratus*) with extensive training. *Behav. Process.* 141(Pt 2), 172–183. doi: 10.1016/j.beproc.2016.11.011

Ditz, H. M., and Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7827–7832. doi: 10.1073/pnas.1504245112

Ditz, H. M., and Nieder, A. (2016). Numerosity representations in crows obey the Weber-Fechner law. *Proc. R. Soc. B* 283:20160083. doi: 10.1098/rspb.2016.0083

Ditz, H. M., and Nieder, A. (2020). Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. *Nat. Commun.* 11:686.

Eiselt, A. K., and Nieder, A. (2013). Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *J. Neurosci.* 33, 7526–7534. doi: 10.1523/jneurosci.5827-12.2013

Fechner, G. T. (1860). *Elemente der Psychophysik*, Vol. 2. Leipzig: Breitkopf & Härtel.

Fernandes, D. M., and Church, R. M. (1982). Discrimination of the number of sequential events. *Anim. Learn. Behav.* 10, 171–176. doi: 10.3758/bf03212266

Frank, M. C., Everett, D. L., Fedorenko, E., and Gibson, E. (2008). Number as a cognitive technology: evidence from Pirahã language and cognition. *Cognition* 108, 819–824. doi: 10.1016/j.cognition.2008.04.007

Gazzola, A., Vallortigara, G., and Pellitteri-Rosa, D. (2018). Continuous and discrete quantity discrimination in tortoises. *Biol. Lett.* 14:20180649. doi: 10.1098/rsbl.2018.0649

Gordon, P. (2004). Numerical cognition without words: evidence from amazonia. *Science* 306, 496–499. doi: 10.1126/science.1094492

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., and Dyer, A. G. (2018). Numerical ordering of zero in honey bees. *Science* 360, 1124–1126. doi: 10.1126/science.aar4975

Jacob, S. N., and Nieder, A. (2009). Tuning to non-symbolic proportions in the human frontoparietal cortex. *Eur. J. Neurosci.* 30, 1432–1442. doi: 10.1111/j.1460-9568.2009.06932.x

Kersey, A. J., and Cantlon, J. F. (2017). Neural tuning to numerosity relates to perceptual tuning in 3-6-year-old children. *J. Neurosci.* 37, 512–522. doi: 10.1523/jneurosci.0065-16.2016

Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F., and Nieder, A. (2018). Single neurons in the human brain encode numbers. *Neuron* 100, 753–761.e4. doi: 10.1016/j.neuron.2018.08.036

Macphail, E. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 308, 37–51. doi: 10.1098/rstb.1985.0008

Macphail, E. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695.

Mandler, G., and Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen.* 111, 1–22. doi: 10.1037/0096-3445.111.1.1

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature* 315, 57–59. doi: 10.1038/315057a0

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *J. Exp. Anal. Behav.* 1, 109–121. doi: 10.1901/jeab.1958.1-109

Meck, W. H., and Church, R. M. (1983). A mode control model of counting, and timing processes. *J. Exp. Psychol. Anim. Behav. Process.* 9, 320–334. doi: 10.1037/0097-7403.9.3.320

Merten, K., and Nieder, A. (2009). Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J. Cogn. Neurosci.* 21, 333–346. doi: 10.1162/jocn.2008.21032

Miletto, P. M. E., Pecunioso, A., Dadda, M., and Agrillo, C. (2020). Searching for the critical p of macphail's null hypothesis: the contribution of numerical abilities of fish. *Front. Psychol.* 11:55. doi: 10.3389/fpsyg.2020.00055

Miller, E. K., Nieder, A., Freedman, D. J., and Wallis, J. D. (2003). Neural correlates of categories and concepts. *Curr. Opin. Neurobiol.* 13, 198–203. doi: 10.1016/s0959-4388(03)00037-0

Moll, F. W., and Nieder, A. (2014). The long and the short of it: rule-based relative length discrimination in carrion crows, *Corvus corone*. *Behav. Process.* 107, 142–149. doi: 10.1016/j.beproc.2014.08.009

Murofushi, K. (1997). Numerical matching behavior by a chimpanzee (Pan troglodytes): subitizing and analogue magnitude estimation. *Jpn. Psychol. Res.* 39, 140–153. doi: 10.1111/1468-5884.00050

Nieder, A. (2004). The number domain- can we count on parietal cortex? *Neuron* 44, 407–409. doi: 10.1016/j.neuron.2004.10.020

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11860–11865. doi: 10.1073/pnas.1204580109

Nieder, A. (2016). The neuronal code for number. *Nat. Rev. Neurosci.* 17, 366–382. doi: 10.1038/nrn.2016.40

Nieder, A. (2019). *A Brain for Numbers: The Biology of the Number Instinct*. Cambridge, MA: MIT Press.

Nieder, A. (2020a). Neural constraints on human number concepts. *Curr. Opin. Neurobiol.* 60, 28–36. doi: 10.1016/j.conb.2019.10.003

Nieder, A. (2020b). The adaptive value of numerical competence. *Trends Ecol. Evol.* 35, 605–617. doi: 10.1016/j.tree.2020.02.009

Nieder, A., Diester, I., and Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science* 313, 1431–1435.

Nieder, A., Freedman, D. J., and Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297, 1708–1711. doi: 10.1126/science.1072493

Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157.

Nieder, A., and Miller, E. K. (2004a). A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7457–7462. doi: 10.1073/pnas.0402239101

Nieder, A., and Miller, E. K. (2004b). Analog numerical representations in rhesus monkeys: evidence for parallel processing. *J. Cogn. Neurosci.* 16, 889–901. doi: 10.1162/089892904970807

Pepperberg, I. M. (1994). Numerical competence in an African grey parrot (*Psittacus erithacus*). *J. Comp. Psych.* 108, 36–44. doi: 10.1037/0735-7036.108.1.36

Piantadosi, S. T., and Cantlon, J. F. (2017). True numerical cognition in the wild. *Psychol. Sci.* 28, 462–469. doi: 10.1177/0956797616686862

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44, 547–555. doi: 10.1016/j.neuron.2004.10.014

Pica, P., Lemer, C., Izard, V., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi: 10.1126/science.1102085

Scarf, D., and Colombo, M. (2020). Columban simulation project 2.0: numerical competence and orthographic processing in pigeons and primates. *Front. Psychol.* 10:3017. doi: 10.3389/fpsyg.2019.03017

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334:1664. doi: 10.1126/science.1213357

Smirnova, A. A., Lazareva, O. F., and Zorina, Z. A. (2000). Use of number by crows: investigation by matching and oddity learning. *J. Exp. Anal. Behav.* 73, 163–176. doi: 10.1901/jeab.2000.73-163

Stancher, G., Rugani, R., Regolin, L., and Vallortigara, G. (2015). Numerical discrimination by frogs (*Bombina orientalis*). *Anim. Cogn.* 18, 219–229. doi: 10.1007/s10071-014-0791-7

Tudusciuc, O., and Nieder, A. (2010). Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. *Exp. Brain Res.* 207, 221–231. doi: 10.1007/s00221-010-2452-7

Uller, C., Jaeger, R., Guidry, G., and Martin, C. (2003). Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Anim. Cogn.* 6, 105–112. doi: 10.1007/s10071-003-0167-x

Vallentin, D., Bongard, S., and Nieder, A. (2012). Numerical rule coding in the prefrontal, premotor, and posterior parietal cortices of macaques. *J. Neurosci.* 32, 6621–6630. doi: 10.1523/jneurosci.5071-11.2012

van Oeffelen, M. P., and Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Percept. Psychophys.* 32, 163–170. doi: 10.3758/bf03204275

Viswanathan, P., and Nieder, A. (2015). Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. *Curr. Biol.* 25, 1259–1269. doi: 10.1016/j.cub.2015.03.025

Weber, E. H. (1850). "Der tastsinn und das gemeingefühl," in *Handwörterbuch der Physiologie Part 2*, Vol. 3, ed. R. Wagner (Nicosia: TP Verone Publishing), 481–588.

Whalen, J., Gallistel, C. R., and Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. *Psychol. Sci.* 10, 130–137. doi: 10.1111/1467-9280.00120

Xia, L., Emmerton, J., Siemann, M., and Delius, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *J. Compar. Psychol.* 115, 83–91. doi: 10.1037/0735-7036.115.1.83

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 21 August 2020
doi: 10.3389/fpsyg.2020.02072

Are There Differences in "Intelligence" Between Nonhuman Species? The Role of Contextual Variables

Michael Colombo* and Damian Scarf

Department of Psychology, University of Otago, Dunedin, New Zealand

We review evidence for Macphail's (1982, 1985, 1987) *Null Hypothesis*, that nonhumans animals do not differ either qualitatively or quantitatively in their cognitive capacities. Our review supports the *Null Hypothesis* in so much as there are no qualitative differences among nonhuman vertebrate animals, and any observed differences along the qualitative dimension can be attributed to failures to account for contextual variables. We argue species do differ quantitatively, however, and that the main difference in "intelligence" among animals lies in the degree to which one must account for contextual variables.

Keywords: ***null hypothesis***, **contextual variables, Macphail, intelligence, species differences**

OPEN ACCESS

Edited by:
Thomas Bugnyar,
University of Vienna, Austria

Reviewed by:
Zhanna Reznikova,
Institute of Systematics and Ecology of Animals (RAS), Russia
Colin Allen,
University of Pittsburgh, United States

***Correspondence:**
Michael Colombo
colombo@psy.otago.ac.nz

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *20 March 2020*
Accepted: *27 July 2020*
Published: *21 August 2020*

Citation:
Colombo M and Scarf D (2020) Are There Differences in "Intelligence" Between Nonhuman Species? The Role of Contextual Variables. Front. Psychol. 11:2072. doi: 10.3389/fpsyg.2020.02072

MACPHAIL'S CLAIM

> *In the present case we should, then, conclude that there are no differences, either qualitative or quantitative, among vertebrates (excluding man; Macphail, 1985, p. 46).*

In 1985, Macphail advocated the *Null Hypothesis* for animal intelligence which stated that there are no differences, either qualitative or quantitative, in intelligence across nonhuman species. Macphail later published his *Null Hypothesis* as a target article in *Behavioral and Brain Sciences* (Macphail, 1987). It is fair to say that the peer commentaries were generally negative. For example, Sternberg (1987, p. 680) stated that "Macphail has made a valiant but not wholly successful effort" while Elepfandt (1987, p. 662) commented with respect to the newly emerging study of vertebrate intelligence that "This new growth should not be stunted by narrow views or precipitate conclusions." Perhaps the most scathing comment was lodged by Goldman-Rakic and Preuss (1987, p. 667) who stated that "Macphail's 'null hypothesis' is merely the epitaph on the head stone of comparative cognition."

Rather than stunting the growth of comparative cognition or becoming the epitaph on its headstone, in the more than three decades since the publication of *Vertebrate intelligence: the null hypothesis* (Macphail, 1985) there has been an explosion of research into the cognitive capacities of animals. Topics, such as *episodic memory*, *theory of mind*, and *planning for the future* were little investigated in 1985, whereas now they form the mainstay of animal cognition studies. And other topics such as the *representation of equivalence relations* (reflexivity, symmetry, and transitivity) have enjoyed a long research history and have continued to generate considerable insight into the mental abilities of nonhuman animals. In light of the wealth of data that has accumulated since Macphail (1985) published his *Null Hypothesis*, the aim of this article is to see whether it has stood the test of time: are there really no differences, qualitative or quantitative, in the cognitive abilities of vertebrate animals?

SOME BACKGROUND ISSUES

In the present article, we review the current status of Macphail's claim that there are no differences, either qualitative or quantitative, in intelligence across nonhuman vertebrate species. Many of the criticisms directed at Macphail (1987) concerned his use of the term "intelligence." For example, Barlow (1987, p. 657) put it perfectly when he stated "because there is not yet any generally agreed upon definition of intelligence that enables a quantitative scale to be defined for it…it cannot justifiably be said that quantitative differences either do, or do not, exist." We believe this is a fair criticism, and furthermore agree with Hodos (1987, p. 668) when he stated that "we should not become bogged down with a general intelligence concept for animals because its measurement is well beyond our grasp."

Limitations in the definition of "intelligence" aside, the field of comparative cognition is about comparing the abilities of different animals in order to understand not only their capacities but also the evolution of the mental abilities of humans. Evaluating how animals differ in "intelligence," however, may not be the best approach. Rather, we think a better approach is to concentrate on specific, definable, and measurable capacities that allow direct comparisons to be made between species. D'Amato and Salmon (1984, p. 149) put forward such a view with respect to comparing the cognitive abilities of different species when they said "how much simpler the task would be if we could identify a relatively small number of kernel cognitive capabilities that would allow us, through their measurement, to make reasonable statements about the cognitive potentials and capacities of various species." To this extent, we focus on a set of such kernel cognitive abilities that have been the subject of extensive investigations across species: reflexivity, symmetry, transitivity, and serial-order behavior, as well as touch upon some more contemporary kernel cognitive abilities such as episodic memory and ToM.

Besides issues surrounding the use of the term "intelligence," another caveat concerns our use of the term "cognition." We use the term more for ease of exposition than necessarily to indicate that our animals are solving tasks using processes that go beyond, or are unexplained by, behavioral principles encompassed by operant and classical conditioning or associative processes. Effectively, our use of the term "cognition" is synonymous with Shettleworth's (1998, p. 5) definition that it encompasses "the mechanisms by which animals acquire, process, store and act on information from the environment."

A third caveat is that it is not our intention to compare the cognitive abilities of nonhuman animals with those of humans. Such comparisons have been dealt with extensively in a recent review (Penn et al., 2008). Rather, our aim is to compare the cognitive abilities of nonhuman animals and specifically, address the value of the *Null Hypothesis*. Similarly, because most research has been conducted on either apes, monkeys, rats, or birds, our comparisons are limited to these species. That said, these species offer a sufficient range of evolutionary independence, as well as differences in neuroanatomy and niches, to forestall any criticism that we failed to sample widely.

THE ROLE OF CONTEXTUAL VARIABLES

In 1965, Bitterman advanced the idea of a contextual variable, a noncognitive factor that accounts for the differences in observed behavior between species. Speaking with respect of the inability of fish to display improvements in reversal learning on both spatial and visual task, Bitterman (1965, p. 95) stated that:

> *"Another possibility is that the conditions under which the fish has been tested are to blame for its poor showing, that the difference in performance is to be traced not only to a difference in capability but also to an inequality in some contextual variable, such as sensory demand, motor demand, degree of hunger, or attractiveness of reward."*

Bitterman (1965, p. 95) also foresaw the problem with the notion of contextual variables when he stated "Can we ever, then, rule out the possibility that a difference in performance of two different animals in such an experiment stems from a difference in some confounded contextual variable?" Macphail (1985, p. 39) revisited the notion of contextual variables in his paper and echoed the same concern when he stated that "There is no finite catalogue of potentially relevant contextual variables: how, therefore, could their effects be conclusively ruled out?"

While the concerns around the issue of contextual variables are reasonable, we believe contextual variables do lie on a continuum of importance and relevance. Although one might be justified in doing so, few would be tempted to argue that a difference in ability between species A and B was because the stimuli used in the experiment with species A were different in size to those used in the experiment with species B. On the other hand, an apparatus that prevents an animal from properly processing a stimulus would indeed be a valid appeal to a contextual variable. Indeed Macphail (1985) went on to conclude that the importance of contextual variables cannot be overlooked, and we fully subscribe to that view. As we will show in the current review, contextual variables often play a role in the outcome of whether an animal can display a certain ability.

We first focus on a set of cognitive capacities referred to as equivalence relations (reflexivity, symmetry, and transitivity). Although the idea of equivalence relations may not spark the notion of cognitive prowess, equivalence relations underlie a number of complex behaviors. According to Sidman (2018, p. 33), for example, equivalence relations play a central role "in making language such a powerful factor in our everyday social intercourse with each other."

Reflexivity

The first equivalence relation we explore is reflexivity known better in the animal cognition literature as the "same-different" or "matching" concept. The task most frequently used to explore whether animals can form a matching concept is the simultaneous matching-to-sample (SMS) task. Although there are many variants, the basic procedure is very simple. An animal is shown a sample stimulus, for example, either a circle or vertical line geometric form. After responding to the sample stimulus two comparison stimuli appear on either side of the sample stimulus, one the same as the sample and the other different. The animal must respond to the comparison stimulus that is the same as the sample stimulus. In this example, from trial to trial, the sample alternates between the circle and vertical line stimuli.

An animal can solve a SMS task in one of three main ways (Skinner, 1950; Farthing and Opuda, 1974; Carter and Werner, 1978). One way is by learning each of the possible configurations of the sample and comparison stimuli. With two stimuli (A and B), and the stimuli arranged so that the sample stimulus appears in the center and the comparison stimuli appear on either side of the sample stimulus, there are four possible sample-comparison configurations, AAB, BAA, BBA, and ABB. According to the *configuration* view, the animal learns that the configurations AAB and BBA mean peck the left stimulus to obtain a reward, and the configurations ABB and BAA mean peck the right stimulus to obtain a reward. A second way to solve a SMS task is by learning a series of *stimulus-response associations* such as "if circle was the sample then press the circle comparison stimulus" and "if vertical line was the sample then press the vertical line comparison stimulus." Finally, a third way to solve a SMS task is by learning a *generalized matching concept* such as "peck the comparison stimulus that matches the sample stimulus." Solving the task by implementing a generalized matching concept is "a necessary consequence of reflexivity, which therefore conveys the notion of sameness" (Sidman et al., 1982, p. 24).

To untangle which of the three possible ways an animal may be solving a SMS task, a transfer test is conducted in which the subjects are presented with novel stimuli, such as red and green. There are a multitude of issues about the conditions that must prevail during the transfer test in order to infer solution by a matching concept. First, the "novel" stimuli must be truly novel in the sense that one should not be able to invoke the notion of stimulus generalization to account for the good transfer performance. In other words, if we train an animal with a circle and vertical line as the stimuli, and tested them with oval and tilted line, the good performance on the transfer test is more likely attributable to stimulus generalization than the application of a matching concept. To avoid the pitfall of stimulus generalization, the stimuli on the transfer test should be completely different (i.e., orthogonal) to the training stimuli. In our example, the transfer stimuli of red and green are orthogonal to the training stimuli of a circle and a vertical line.

Another critical issue is how good does transfer performance have to be with the novel stimuli to infer solution by a matching concept? The basic idea is that if the animals had learned the original task by adopting a matching concept they ought to transfer rapidly to novel stimuli because a *matching concept* tends to be independent of the stimuli. On the other hand, if the animals had learned the original task using either the *configuration* rule or the *stimulus-response association* rule then performance with the novel stimuli should be poor because both of these processes are dependent on the original stimuli, and indeed it might take the animal as many trials to learn the task with the novel transfer stimuli as it did to learn the task with the training stimuli. Naturally, it is rare that either of these extreme situations prevail, and often we are left with measures of savings from which one must use their best judgment as to what process the animal had employed. For example, if it took an animal 500 trials to learn the original task with circle and vertical line stimuli, and they took 50 trials to learn the task with red and green stimuli, is that sufficiently good performance from which to infer that the original task had been learned using a matching concept? Most would probably say yes. But then what about 100 trials?

There is ample evidence across a wide range of species that animals learn to solve a SMS task by applying a matching concept. Chimpanzees, both adult (Nissen et al., 1948; Robinson, 1955) and infant (Oden et al., 1988), readily transfer to novel stimuli to the point that one could almost talk about near-perfect levels of performance on the first few trials. For example, as a group, the infant chimpanzees in the Oden et al. (1988) study took 816 trials to learn the matching task with the training stimuli to a level of about 85% correct, and continued to score at that level across the first 24 trials with a variety of different novel stimuli. Although not quite to the level of competence of the chimpanzees, monkeys also are capable of showing high levels of transfer with novel stimuli (Mello, 1971; Milner, 1973; D'Amato et al., 1985a).

Outside of non-human primates, studies have focused largely on the abilities of pigeons. Early studies either failed to find evidence of a matching concept (Cumming and Berryman, 1961; Farthing and Opuda, 1974; Holmes, 1979), provided at best weak evidence for a matching concept (Wilson et al., 1985a,b), or the evidence for a matching concept was open to alternative explanations (Zentall and Hogan, 1974, 1978; Urcuioli and Nevin, 1975; Edwards et al., 1983). One such alternative explanation was common coding of stimuli. For example, Zentall and Hogan (1974) trained pigeons with red and green stimuli and then tested with yellow and blue, and the birds showed reasonably good levels of transfer to the "novel" stimuli. Unfortunately, pigeons tend to code yellow and red as similar, and blue and green as similar (Wright and Cummings, 1971), so the transfer seen was nothing more than an instance of stimulus generalization, that is, a violation of the principle or orthogonality. Even a further study (Zentall and Hogan, 1976) in which pigeons trained with a circle and cross geometric forms and then transferred to (clearly novel) red and green stimuli showed high levels of transfer, but failed to recognize that pigeons learn a SMS task with red and green stimuli very quickly (Zentall and Hogan, 1974), thus casting doubt that the rapid transfer to red and green was due to the application of a matching concept.

Early pigeon matching concept studies tended to support the view that, rather than learning a matching concept, the behavior of the pigeons could be best described as learning a series of stimulus-response or configuration associations. The evidence for pigeons forming a matching concept, however, took a big step forward when Wright (1997) showed that the number of responses emitted to the sample stimulus is a critical determinant of whether pigeons will form a matching concept. Different groups of birds were trained to emit either an FR0, FR1, FR10, or FR20 to the sample stimulus, and then tested with novel stimuli under the same response conditions. Wright (1997) found that birds trained with either and FR0 or FR1 failed to transfer to novel stimuli, whereas those trained with FR10 or FR20 showed levels of performance with the novel stimuli similar (or equivalent in the case of the FR20 condition) to their terminal performance with the training stimuli. The number of responses emitted to the sample stimulus was a contextual variable that had been overlooked in many early pigeon studies, where few responses were required to the sample stimulus.

According to Wright (1997), configurational learning is the dominant and preferred learning strategy for pigeons, and in order to display evidence of a matching concept, one must first break the predisposition to process the sample and comparison stimuli as a configuration. Effectively, the larger the FR requirement, the more likely it is that the animal divorces itself from configural learning, and the more likely it will then adopt a matching concept. Take the case of the FR0 condition. The sample and comparison stimuli are presented at the same time, and so it is unlikely that the birds even appreciate that there is a "sample" stimulus that needs to be matched to one of the "comparison" stimuli. And why would they? In effect, the only solution under an FR0 condition is to treat the entire display of "sample" and "comparison" stimuli as a unitary whole, that is, a configuration, and direct your responses accordingly. On the other hand, in the FR20 condition, the sample appears and then only after 20 responses do the comparison stimuli appear. The structure of this task encourages the animals to perceive the sample as something they have to match to the comparison, and as a result, pigeons are more inclined to adopt a matching concept, and transfer to novel stimuli.

A subsequent study by Colombo et al. (2003) uncovered yet another contextual variable that must be adjusted before pigeons will display a matching concept. These authors were surprised when their FR20 pigeons failed to transfer to novel stimuli. They noted, however, that another difference between the Wright (1997) study and their study was that Wright (1997) had initially trained their birds with three stimuli, whereas Colombo et al. (2003) trained theirs with just two. Although training with two versus three stimuli may not seem like an impactful contextual variable, two training stimuli yield four possible sample-comparison configurations, whereas three training stimuli yield 12 possible sample-comparison configurations. Indeed when Colombo et al. (2003) trained another group of birds with three stimuli and an FR20 requirement, they transferred to novel stimuli at a very high level. Thus *number of training stimuli* is also a contextual variable. They reasoned that while it might be possible to learn the right/left responses associated with four configurations, learning the right/left responses associated with 12 configurations might pose difficulty for the animals, and encourage the use of a matching concept to solve the task.

In summary, if one designs the experiment properly, one can show levels of transfer in pigeons virtually identical to levels of transfer in monkeys (Colombo et al., 2003). It is true that, in the case of the pigeon, one must impose an FR20 to the sample stimulus and train them with three stimuli, compared to monkeys that show transfer with an FR1 to the sample stimulus and training with just two stimuli. Once these contextual variables are accounted for, however, the performance of pigeons becomes indistinguishable from that of monkeys. This is true not only for the conditions that results in successful transfer, but also the conditions that result in unsuccessful transfer (see **Figure 1**). Both the D'Amato et al. (1985a) and Colombo et al. (2003) studies employed the same training and testing format, in that the animals were trained with a number of stimuli and then tested over four sessions with novel stimuli as well as the training stimuli. It is clear from **Figure 1** that when the contextual variables of FR and number of training stimuli are adjusted, the transfer performance of the birds is indistinguishable from that of the monkeys, both in terms of the successful transfer to a novel color and form stimulus (left panel), as well as unsuccessful transfer to two novel form stimuli (right panel). More on the difference between transfer to color/form and form/form stimuli later.

One final point in the matching concept literature deserves some attention. Premack (1983) has made the claim that animals can be distinguished on the basis of the type of matching procedure that is employed. According to Premack (1983), the procedures discussed in all the above studies are what he calls "successive" matching tasks, where the response of same or different are directed to the physical stimuli themselves (e.g., press the red comparison stimulus if the sample was red).

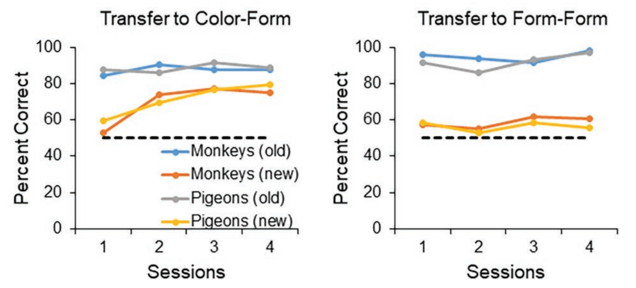


FIGURE 1 | Transfer performance of monkeys and pigeons. The monkey data are based on D'Amato et al. (1985a) and the pigeon data are based on Colombo et al. (2003). The animals were tested over four 48-trial sessions, with half of the trials dedicated to the training (old) stimuli, and half dedicated to the novel (new) stimuli. The **left panel** shows the transfer performance to a novel color and form stimulus (for training the monkeys were trained with two form stimuli and the pigeons trained with three form stimuli). The **right panel** shows the transfer performance to two novel form stimulus (for training the monkeys were trained with a color and a form stimulus, whereas the pigeons were trained with two form stimuli and a color stimulus). When the contextual variables are set appropriately for the pigeons (training with three stimuli and an FR20 on the sample), both monkeys and pigeons transfer readily to a novel color and form stimulus, but not to two novel form stimuli.

Premack (1983) believes that the ability to solve such "successive" matching tasks is ubiquitous among animals. On the other hand, a "simultaneous" matching task can only be solved, not just by any chimpanzee, but only language-trained chimpanzees. In the "simultaneous" procedure, the discriminanda to which the judgments of same and different must be made are separate from the actual stimuli being judged as same and different. For example, if the animal was presented with stimuli A and B it would have to choose the cue that signifies "different," say a red rectangle, or if presented with stimuli A and A it would have to choose the cue that signifies "same," a yellow rectangle (Premack et al., 1978). There seems little doubt that chimpanzees can solve such "simultaneous" tasks (Premack et al., 1978), and despite Premack's (1983) claim of a language-training prerequisite, so too can non-language-trained monkeys (Sands and Wright, 1980; D'Amato and Colombo, 1989).

Whether pigeons can solve simultaneous matching tasks has, as is often the case for pigeons, taken longer to show. Early positive reports were marred by alternative interpretations, such as the animals potentially learning the fixed order of the left/right responses associated with the "same" and "different" outcomes (Santiago and Wright, 1984), or a failure to fully balance the design thereby allowing the birds to solve the task using item-specific associations (Edwards et al., 1983). Far better transfer performance has been obtained on simultaneous matching tasks when the discriminanda consisted of arrays of multiple same and multiple different stimuli (Santiago and Wright, 1984; Wasserman et al., 1995; Cook et al., 1997). If, in fact, the birds are processing the specific items in the arrays then these studies would provide evidence for pigeons being able to solve simultaneous matching tasks. The criticism with these studies, however, is that the novel "same" and "different" arrays are really not novel. If instead of looking at the individual items that compose an array the animals are processing a global feature, perhaps a measure of the "entropy" of the stimulus array, then the "novel" arrays are really not novel after all (Young et al., 1997). More recently, however, Blaisdell and Cook (2005) have shown that pigeons can perform a simultaneous matching task when only two stimuli are presented at a time, and they transfer to novel stimuli at a level that would suggest evidence of a matching concept.

When Macphail (1985) made his claim that there were no qualitative or quantitative differences among species, he was referring to only vertebrate species. To drive the point home concerning the absence of differences among vertebrates in the ability to form a matching concept, it is worth finishing this section with a matching concept study using invertebrates. Giurfa et al. (2001) showed that honeybees also solve a SMS task using a matching concept. They used a Y-maze with the bees encountering the sample stimulus on the stem of the Y-maze and the comparison stimuli on the arms of the Y-maze. The bees easily learned the task and showed perfect transfer to novel stimuli. So exceptional was the performance of the bees that not only did they transfer to novel visual stimuli but they also transferred the matching concept across modalities, an ability that has never been shown even in non-human primates (see D'Amato et al., 1985a). In summary, when the contextual variables are adjusted for each species, a number of animals display transfer to novel stimuli at a level that would suggest the employment of a matching concept. Whether, in fact, it is necessary to formulate the performance in terms of the cognitive construct of a *matching concept*, as opposed to the operation of associative processes, is an issue to which we will return at the end of this review.

Symmetry

The second equivalence relation we explore is known as symmetry. When you learn the name of an object, say "door," from then on, the word "door" brings to mind an image of a door. Likewise an image of a door brings to mind the word "door." This is an example of *symmetry*, a bidirectional association between two stimuli. Symmetry in the context of the animal literature is usually trained using a version of the matching-to-sample task called the symbolic or conditional matching-to-sample, in which different sample stimuli are mapped onto different comparison stimuli. The aim of the task is therefore not to match, in terms of sameness, a comparison stimulus to a sample stimulus, but to choose the comparison stimulus that is associated with the sample stimulus. For example, if A1 and A2 are the sample stimuli, and B1 and B2 are the comparison stimuli, then when A1 appears as the sample the correct choice is B1, whereas when A2 appears as the sample the correct choice is B2. To test for symmetry, B1 and B2 now become the sample stimuli, and A1 and A2 become the comparison stimuli. If the learned relationships, A1→B1 and A2→B2 are symmetrical, then when presented with B1 or B2 as the sample stimuli the subject should chose A1 and A2, respectively. Although animals readily learn symbolic matching-to-sample tasks, demonstrating symmetry in a number of species has proven difficult.

It is worth mentioning at this point that the term symmetry typically implies that the backward association is learned to the same degree as the forward association. By this harsh definition, it would appear that there is little or no evidence for such symmetry in nonhuman animals. As in most cases, in the nonhuman animal literature we accept a significant backward association (albeit less pronounced than the forward one) as evidence of symmetry. With this in mind, Tomonaga et al. (1991) trained three chimpanzees to match one of two sample colors to one of two comparison shapes to a criterion of at least 80%, then overtrained the animals for hundreds of trials, and then tested for symmetry over 12 trials. Keep in mind that testing for the emergence of an ability over a mere 12 trials is a tall order, as animals are often impaired by any change in testing conditions. Nevertheless, one of three trained chimpanzees performed above chance on the symmetry test, providing evidence that chimpanzees are capable of forming symmetrical relations. The evidence for symmetry in chimpanzees, however, is by no means uniformly positive. Yamamoto and Asano (1995), for example, found their one chimpanzee displayed no evidence for symmetry after training with one stimulus set, but after specific training and testing with six stimulus sets, a procedure called exemplar training, symmetry did emerge.

Demonstrating symmetry in monkeys has also been met with great difficulty. Sidman et al. (1982) failed to show any evidence of symmetry in monkeys trained with geometric (vertical and horizontal line) samples and color comparison stimuli. McIntire et al. (1987) purported to show evidence of symmetry in macaque monkeys; however, their conclusions were met with considerable criticisms on the basis that the tested-for relations were already trained (see Hayes, 1989). Surprisingly, the study by D'Amato et al. (1985b), one that is commonly cited as a negative finding (Hayes, 1989; Sidman, 1994; Lionello-DeNolf and Urcuioli, 2002; Frank and Wasserman, 2005), provides some favorable evidence for symmetry in monkeys. D'Amato et al. (1985b) argued that the use of vertical and horizontal line comparison stimuli in the Sidman et al. (1982) study could have been the contextual variable that put the monkeys at a disadvantage. Employing far more discriminable stimuli as sample and comparisons, and also assessing performance over the first 12 trials, D'Amato et al. (1985b) showed evidence for significant backward associations in two of the six monkeys tested.

Numerous studies have explored the extent to which pigeons display symmetry, and positive findings have been difficult to obtain. Early studies either failed to find any evidence for even backward associations (Lipkens et al., 1988), were criticized for alternative interpretations when they did (Vaughan, 1988; Hayes, 1989) or much like for chimpanzees and monkeys, found at best only weak evidence for backward associations (Hogan and Zentall, 1977; Richards, 1988). Interestingly in the Hogan and Zentall (1977) study, some of the positive evidence for symmetry was seen early in the test for symmetry but then dissipated, an outcome also observed by D'Amato et al. (1985b) with monkeys. Given the context of this article it is perhaps fitting to include one possibility raised by Hogan and Zentall (1977, p. 14) as to why the pigeons fare poorly on symmetry tasks: "it is also possible that the development of backward associations depends upon the species-specific functional value of such associations (i.e., humans may need to be able to develop backward associations whereas pigeons may not)."

Lionello-DeNolf and Urcuioli (2002) also failed to find evidence for symmetry in pigeons, but their study is worth mentioning because it represents one of the earliest attempts to address possible contextual variables that may be preventing pigeons (and possibly other animals) from displaying symmetry. Drawing from McIlvane et al.'s (2000) notion of *stimulus response topography* that pigeons may process aspects of a stimulus that interfere with the aspects of interest in tests of symmetry, Lionello-DeNolf and Urcuioli (2002) reasoned that during the symmetry test not only do the sample and comparison exchange roles but they also exchange positions. Thus, pigeons seem to code not only the features of the stimulus but also the positions of the stimuli as part of the *stimulus response topography*. Take the situation in matching tasks where the sample stimulus typically appears in a central position and the comparison stimuli appear to either side of the sample position. For the test of symmetry, the comparison stimuli now appear in the central position. To a human it might be irrelevant that the comparison stimulus now appears in a position that it has never appeared in before, but to a nonhuman, position may be part of the *stimulus response topography*, and hence, nonhuman animals may fail the symmetry test because it is unclear how they should behave when stimuli appear in positions that they have never appeared in before. Lionello-DeNolf and Urcuioli (2002) therefore trained their animals so that the sample and comparison stimuli could appear in any of a number of positions, thus effectively training "position" out as a component of the *stimulus response topography*. Despite this training the pigeons still failed to show any evidence for symmetry, a finding that, marginal as the evidence for symmetry is in non-human primates, further seems to distance pigeons from nonhumans in their ability to form symmetrical relations.

Frank and Wasserman (2005) noted, however, that in addition to the stimuli being associated with their spatial location, they are also associated with their temporal location. In other words, if the relation A1→B1 is trained and then the relation B1→A1 is tested, item B has never appeared first. Similar to the case for position mentioned earlier, if item B now appears first, we as humans may quickly assume that because it appears first it must be serving in the role of a sample stimulus, but again there is no reason why other animals should make that assumption. To account for the potentially controlling influence of the contextual variable of temporal location, Frank and Wasserman (2005) used a successive go/no-go matching tasks, where the sample and comparison stimuli appear successively in the same position, and the subject required to make a go response to the second stimulus if it is paired with the first (e.g., A1→B1), and a no-go response (i.e., withhold responding) to the second stimulus if it is not paired with the first (e.g., A1→B2). To control for the potentially disruptive effects of the contextual variable of temporal order and the fact that, for example, stimulus B had never appeared first, the pigeons were trained not only with symbolic relations (A1→B1, A2→B2) but also with identity relations (A→A and B→B), thus training the animals that both stimuli A and B can occur in any temporal position. With these contextual variables in mind, the pigeons displayed robust symmetry. Frank and Wasserman (2005, p. 157) concluded that "symmetry can be obtained with nonhuman animals under proper conditions of training and testing." Interestingly, the one successful chimpanzee in the Tomonaga et al. (1991) study was also trained with both symbolic and identity relations.

In summary, Frank and Wasserman's (2005) elegant study shows that once contextual variables are taken into account, pigeons can display symmetry, and do so to a level not that dissimilar to chimpanzees. Furthermore, there is little if any evidence that would distinguish the performance of non-human primates and birds with respect to the formation of symmetry. Much like the matching concept literature, researchers are now investigating whether symmetry can be demonstrated by invertebrates. Given a recent attempt by Moreno et al. (2012) with honeybees, it seems only a matter of time before an invertebrate species can be shown to display symmetry.

Transitivity

The third equivalence relation we explore is known as transitivity. There is little need to appeal to the notion of contextual variables because most species have been shown capable of solving transitivity tasks. Rather, we include a brief mention of this topic to complete our discussion of equivalence relations, and more importantly to highlight another issue we wish to briefly address in this review, namely the desire to interpret the behavior of nonhuman animals in overtly cognitively-rich terms.

Transitivity is an operation whereby given the information that A is smarter than B, and B is smarter than C, one makes the logical conclusion that A is smarter than C, even though no direct information about the relationship between A and C was ever given. According to Piaget (1928), the ability to solve such a three-term transitive inference task does not develop until approximately 7 years of age, a conclusion that was challenged by Bryant and Trabasso (1971), who demonstrated robust transitive inference abilities in 4-, 5-, and 6-year olds. Although the main purpose of this review is to compare nonhuman animals, the procedure used by Bryant and Trabasso (1971) is worth mentioning because very similar training procedures have been used to explore transitivity in nonhuman animals. In their study, children were trained to discriminate between colored rods of different lengths. The rods were presented in pairs, and training consisted of repeated exposures to four training pairs, A+B−, B+C−, C+D−, and D+E−, with the letters representing the different lengths of rods (e.g., A was the longest and E the shortest) and the "+" and "−" indicating the correct and incorrect stimulus, respectively, to select. For example, when presented with pair CD and prompted with the question "*which rod is longer?*", the subject should select item C.

Of course, in learning a five-item transitive series, A is always correct, E is always incorrect, and B, C, and D are both correct and incorrect depending on the pair in which they appear. With a five-item series there are 10 possible pairs with which to test the subject (AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE). Of these, we expect the subject to perform well with any pair that contains item A (AB, AC, AD, and AE) because in training item A was always correct. We also expect the subjects to perform well with any pair that contains item E (AE, BE, CE, and DE) because in training item E was always incorrect and hence should always be avoided in favor of the other stimulus. Finally, we clearly expect them to perform well with a pair that was one of the training pairs (AB, BC, CD, and DE), leaving as the critical test for transitivity pair BD. Bryant and Trabasso (1971) found that 4-, 5-, and 6-year olds performed at high levels on pair BD.

Studies using nonhuman animal subjects tend to follow the same general procedure adopted by Bryant and Trabasso (1971) of initially training the animals on the four premise pairs AB, BC, CD, and DE, and then testing them on the critical BD pair. Using this procedure chimpanzees (Gillan, 1981), monkeys (McGonigle and Chalmers, 1977), rats (Davis, 1992), and pigeons (von Fersen et al., 1991; Paz-y-Miño C et al., 2004) have all been found to perform at high levels on the critical BD test pair, and indeed achieve levels of performance not too different from that reported by Bryant and Trabasso (1971) for young children. To be sure there have been failures by pigeons to show transitivity (D'Amato et al., 1985b), but there have also been failures by primates to display transitivity (Sidman et al., 1982). Despite the occasional failure, there is no need to appeal to contextual variables, because in general pigeons solved transitivity tasks as well as other animals.

A key feature surrounding many of these studies is the extent to which the high level of performance on the BD test pair reflects a cognitive/logical operation or a behavioral/associative operation. In the cognitive/logical camp is the view that while learning how to respond to each of the five premise pairs (e.g., A+B−, B+C−, C+D−, and D+E−) animals form a hierarchical linear mental representation of how the five stimuli are related to one another (e.g., A > B > C > D > E), and use that representation to guide them as to how to respond to the critical BD test pair. In the behavioral/associative camp, no linear mental representation of the five items is formed. Rather, solution of the critical BD pair is based on conducting an associative computation based on reward values assigned to each of the items (Value Transfer Theory), or by relying on previously learned premise pairs to solve the BD problem (Binary Sampling Model). According to the Value Transfer Theory (von Fersen et al., 1990, 1991), different strengths are assigned to each of the five stimuli as a function of which pairs they have appeared in during training and whether they were associated with the always rewarded stimulus A or never-rewarded stimulus E. As a result of such associations, item B is ranked higher than item D, and so the animal will choose B when presented with pair BD. Indeed, the resulting rankings can be used to very accurately predict which stimulus an animal will select when any two stimuli are paired.

The Binary Sampling Model (McGonigle and Chalmers, 1977) is also a simple yet effective noncognitive account of why an animal selects item B during the critical BD test. According to the model (see **Figure 2**), upon seeing pair BD the animal attempts to solve the task as if it were either pair BC, CD, or BD. A test session typically consists of numerous presentations of pair BD, and according to the model, there is a 1 in 3 chance that either of the three pairs is selected on any given trial. Given that each pair is selected 33% of the time, we can think of each pair as having 0.33 units to contribute to the solution of the BD problem. If the animal attempts to solve the BD pair as if it were pair BC, it will select B because B+C− is one of the training pairs where the animal is taught to select B. Item B therefore accumulates 0.33 units. If the animal attempts to solve the BD pair as if it were pair CD it will select C because C+D− is one of the training pairs where the animal is taught to select C. But keep in mind that there is no item C to select because, remember, the animal is presented with pair BD not CD. However, if the animal was attempting to solve pair BD as if it were pair CD the animal is also trained on that pair to avoid item D, and so the animal avoids item D in pair BD and selects item B. Item B again gets all 0.33 units, bringing its current tally to 0.66 units. Finally, the animal may attempt to solve pair BD as if it was pair BD. Unfortunately, pair BD is not a training pair and so no associations have been established between items B and D, and the animal will randomly select B half the time and D half the time, and the 0.33 available

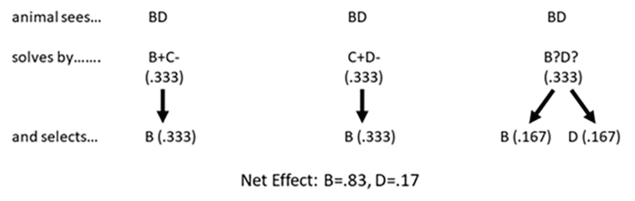


FIGURE 2 | The Binary Sampling Model. According to McGonigle and Chalmers (1977), an animal attempts to solve the BD pair as if it were either pair BC, CD, or BD. B+C− is a training pair and so it will select item B. C+D− is also a training pair where the animal is taught to select C. Unfortunately, because the animal has been presented with pair BD, there is no item C being displayed and hence no item C to select. But if the animal were trying to solve the BD pair as if it was item C+D− being displayed, it also learned to avoid D, which is what the animal does, and again selects item B. Finally, if the animal attempts to solve the BD pair as if it were BD, it has received no training with these two stimuli presented together, and randomly chooses between them. The net effect is the animal will select B 83% of the time and D 17% of the time, which happens to be very close the performance levels many animals achieved with the BD test.

units gets split between the two items. The final tally is that item B gets 0.83 units and item D gets 0.16 units, which when expressed in terms of percent correct is remarkably close to the performance of animals with the BD test pair across a wide range of studies.

Both the Value Transfer Theory and Binary Sampling Model, as well as other noncognitive accounts of transitivity (see Delius and Siemann, 1998), very nicely account for the high levels of BD test pair performance without the need to appeal to cognitive accounts such as hierarchical mental representations. Of course, there have been challenges to these simpler accounts of transitive inference (Steirn et al., 1995; Lazareva and Wasserman, 2012) but it is hard to overlook the power of associative strength *via* reinforcement history (Siemann et al., 1996). It is difficult to do justice in this review to the complex transitive inference literature, but whether one believes in cognitive/logical accounts or behavioral/associative accounts, one thing is certain, there is no support for the view that monkeys perform such tasks any differently to pigeons. The recent demonstration of this ability in invertebrates means this question can now be extended beyond vertebrates (Tibbetts et al., 2019). Perhaps more importantly, the fact that a species with just 0.001% of the neurons in a human brain (Azevedo et al., 2009; Menzel, 2012) can pass the task should call into question the cognitively-rich terms with which researchers describe transitive inference.

Serial-Order Behavior

Conceptually related to studies of transitivity are studies that explore the serial-order abilities of animals. The serial-order task, also known as the simultaneous chaining procedure, has provided a wealth of information on the structure of the representations believed to underlie transitive judgments. The task is straightforward, and like the transitivity procedure, often uses five stimuli. Rather than presenting the five stimuli as four training premise pairs, however, in the serial-order task the animals are trained to respond to five simultaneously presented stimuli in a specific order, namely, A→B→C→D→E. Both monkeys and pigeons can learn to perform the five-item serial-order tasks to the same high levels (D'Amato and Colombo, 1988; Terrace 1993; Scarf and Colombo, 2010). To determine what the animals have learned, much like in the transitivity test, subjects are given a pairwise test consisting of all 10 possible pairs of stimuli that can be generated from the five-item list (AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE). A correct response on the pairwise test requires that the animals respond to the two displayed items in a manner consistent with their order in the five-item series. When presented with pair BC, for example, to obtain a reward the animal must first respond to item B and then to item C.

The pairwise test has provided considerable insight into the processes that different animals use in learning the original five-item serial-order task. In fact, until recently, the performance on the pairwise tests, as well as latency measures that can generated from the correct responses, seemed to provide some of the best evidence that monkeys and pigeons process serial-order information in fundamentally different ways (Terrace, 1993; Scarf and Colombo, 2008). For example, in terms of performance across the 10 pairs, monkeys perform at very high levels on all test pairs, whereas pigeons perform at high levels only on pairs that contain either item A or item E. Importantly, pigeons perform at chance levels on the internal pairs BC, BD, and CD (see **Figure 3**). Such an outcome is consistent with the view that in the course of learning a serial-order task monkeys form a mental representation of the list and used that representation to guide their behavior (D'Amato and Colombo, 1988). Pigeons, on the other hand, seem unable to form such a representation, and rather learn a simple set of behavioral rules such as "(1) Respond first to item A. (2) Respond last to item D. (3) Respond to any other item by default" (Terrace, 1993, p. 164).

Further evidence that monkeys form a mental representation of the series comes from two types of analysis of the latency data to respond to the first and second items of the displayed pair. In the case of the *first-item effect*, the latency to the first item of the pair is averaged across all pairs that share the same first item. In other words, the latency to item A is averaged across pairs AB, AC, AD, and AE, the latency to item B averaged across pairs BC, BD, and BE, the latency to item C averaged across pairs CD and CE, and the latency to item D is based on the only pair that has item D as a first item, pair DE (see **Figure 4**, left panel). Monkeys clearly display a *first-item effect*, in that the latency to respond to the first item of a pair is longer the further along the list that the first item lies. For example, the latency to respond to item C in pair CD takes longer than the latency to respond to item B in pair BD. Such a latency function suggests that the monkeys are accessing the list at item A and progressing through the list in a linear fashion trying to match the item in memory to a displayed item. In contrast to monkeys, pigeons show a flat *first-item effect*.

Monkeys also display what is known as a *missing-item effect* (**Figure 4**, right panel). The *missing-item effect* refers to the

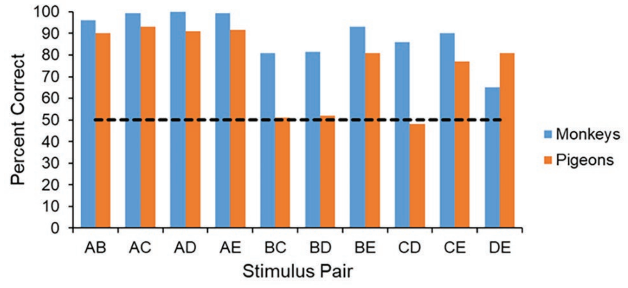


FIGURE 3 | Performance across the 10 pairs during the pairwise test. The monkeys perform well on all pairs, whereas the pigeons only perform well on pairs that have either an item A or an item E, and perform at chance on the internal pairs that are missing these items.

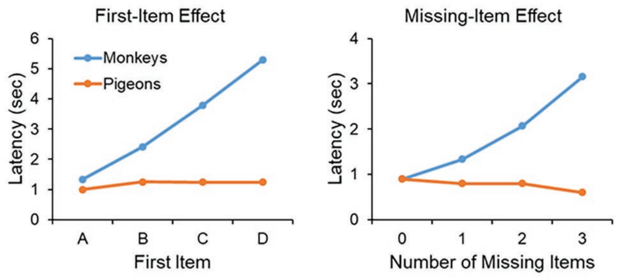


FIGURE 4 | Left panel: The *first-item effect*. The latency to respond to the first item of a pair for monkeys and pigeons as a function of whether the first item was A (averaged across pairs AB, AC, AD, and AE), B (averaged across pairs BC, BD, and BE), C (averaged across pairs CD and CE), or D (based on pair DE only). Monkeys show a linear increase across first item whereas pigeons do not. The data on based on correct trials. **Right panel**: The *missing-item effect*. The latency to respond to the second item of a pair as a function of whether the second item was separated from the first by 0 missing items (averaged across pairs AB, BC, CD, and DE), 1 missing item (averaged across pairs AC, BD, and CE), 2 missing items (averaged across pairs AD and BE), or 3 missing items (based on pair AE only). Monkeys show a linear increase across the number of missing items whereas pigeons do not. The data on based on correct trials.

latency to respond to the second item of a pair as a function of the distance from the first item to the second item. In pairs AB, BC, CD, and DE, there are no missing items in that the second stimulus of a pair occurs directly after the first stimulus. Pairs AC, BD, and CE have one missing item, pairs AD and BE have two missing items, and pair AE has three missing items. Monkeys display a very clear *missing-item effect* in that the latency to respond to the second item of a pair is a function of the number of missing items between the first and second item. For example, monkeys are faster to respond to item D in pair CD than item D in pair BD. The reason is because in pair CD there are no missing items to access, whereas in pair BD the monkey must access one missing item, item C. In contrast to the monkeys, pigeons do not display a *missing-item effect*.

The performance across the 10 pairs, as well as the presence of a *first-item effect* and a *missing-item effect*, supports the view that in the course of learning a serial-order task monkeys form a linear mental representation of the items and use that representation to guide their behavior, for example, during the pairwise test. In contrast, the absence of these effects in pigeons suggests that they solve the serial-order task in a fundamentally different way to monkeys. These views fit well with the notion that the success of the monkeys may very well be related to their ability to respond appropriately to dominance hierarchies (Cheney et al., 1986), something that is not necessary for the pigeon, whose social structure has a far less hierarchical organization (Masure and Allee, 1934).

Is it really the case, however, that pigeons have no knowledge of the ordering of the stimuli in a serial-order task, or has revealing that ability been masked by some contextual variable? Recall that the pairwise test occurs once the animals have reached a certain level of proficiency on the five-item serial-order task, and consists of presenting the subjects with all 10 pairs of stimuli that can be generated from the five items (AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE). Furthermore, each of the 10 pairs is shown a number of times within a session (typically four times within a 40-trial session). We wondered whether the structure of the pairwise test, and the surprise at being shifted from a five-item task to a pairwise test with all 10 pairs intermixed within a session, was perhaps causing the pigeons difficulty. Was the dramatic change in context the contextual variable that accounted for the pigeons' poor performance on the pairwise test? We explored this possibility across two experiments (Scarf and Colombo, 2010).

In one experiment, we trained four pigeons on a four-item serial-order task and another four pigeons on a five-item serial-order task. Instead of then delivering a pairwise test of six pairs (the number that can be generated from a four-item list) for the birds trained on the four-item task, or 10 pairs for the birds trained on the five-item task, we attempted to mitigate the effects of the context change by presenting the four-item-trained birds with just pair BC (the critical internal pair after training on the four-item list) or the five-item-trained birds with just pair BD (a critical internal pair after training on a five-item list). The BC or BD pairs were presented 40 times per session. We reasoned that if the pigeons learned nothing about the order of items B and C, or the order of items B and D, then those tested on the positive pair condition (BC+ or BD+) and rewarded for pressing B→C or B→D should fare no better than those tested on the negative pair condition (BC− and BD−) and rewarded from responding to the items in the opposite direction, that is C→B or D→B. The results are shown in **Figure 5**. Clearly animals trained on the positive pairs acquired the task significantly faster than those trained on the negative pairs, suggesting that if the conditions are set up properly, pigeons display evidence that they understand the order of the internal items on four-item and five-item serial-order tasks.

When tested with just one pair, the birds were able to indicate that they did understand that item B comes before item C, or item B comes before item D, and thus provide us with evidence that they did understand, at least at some rudimentary level, the organization of the internal items in a series. That said, the positive pair birds did experience far

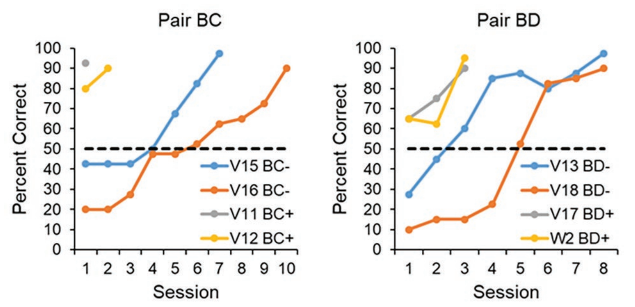


FIGURE 5 | Performance on pair BC after training on a four-item list **(left panel)** and BD after training on a five-item list **(right panel)**. Two animals each were trained on the BC+, BC−, BD+, and BD− conditions, where the "+" indicated that a reward could be obtained by responding to the items in the order in which they appeared in the original sequence (B→C or B→D), and the "−" indicated that a reward could be obtained by responding to the items in the order opposite to what they appeared in the original sequence (C→B or D→B). If the birds learned nothing about the order of the internal items, then when presented with these pairs they ought to take as long to learn the positive pair condition as the negative pair condition. Rather, it is clear that the birds tested with the positive pairs fared far better than those tested with the negative pairs.

more trials (40–80 for the BC+ birds, 120 for the BD+ birds) on their respective pairs than that typically experience by the monkeys on those same pairs during a regular pairwise test (usually around 8–12 trials). We wondered, then, if pigeons could ever display high levels of performance on a critical pair, as did the monkeys, following exposure to a limited number of trials. To test this notion we again modified the pairwise testing procedure. For the second experiment, pigeons were trained on a four-item serial-order task and presented with the critical BC pair as a probe of four trials embedded against a baseline of 36 trials dedicated to the standard (A→B→C→D) four-item serial-order task. The test was run for four sessions giving a total number of 16 BC trials, a number very similar to that experienced by the monkeys. The results are shown in **Figure 6**. All four birds performed at very high levels across the 16 BC probe trials. For comparison, also shown in the figure is the performance of pigeons on the BC pair when it was delivered in a standard pairwise test format in which the six pairs that can be generated following training with a four-item list (AB, AC, AD, BC, BD, and CD) are presented intermixed within a session (Straub and Terrace, 1981) with no baseline A→B→C→D trials. Clearly, pigeons can perform well on a critical pair after limited exposure to that pair, but only when the context of the overall test is not dramatically changed from the training situation.

We have shown that by mitigating the effects of a dramatic change in context, pigeons can perform well on a critical internal test pair, thus supporting the view that they do understand the order of the internal items in a list. It would seem that for the pigeons, for whatever reason, displaying all the pairs at once as in a standard pairwise test is a contextual variable that prevents them from displaying their understanding of the organization of the items of a four-item and five-item lists.

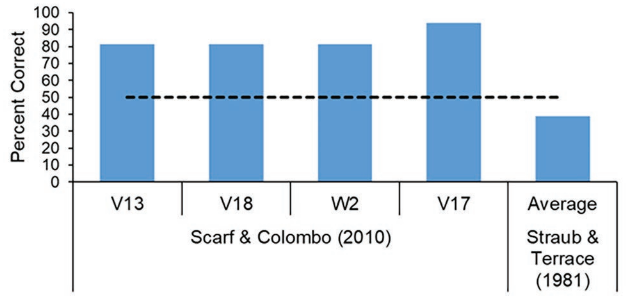


FIGURE 6 | Performance on pair BC when delivered as probes embedded against a baseline four-item serial-order task. All four pigeons performed at high levels after exposure to only 16 BC probes trials. Also shown is the BC performance of Straub and Terrace (1981) pigeons where the BC trials were delivered in a standard pairwise tests fashion along with all the other pairs in a session.

Episodic Memory and Theory of Mind

There are many tasks that have been used to probe the abilities of nonhuman animals, for which there are not only no differences in performance across species, but also for which some of the most compelling evidence for a particular ability actually comes from birds, rather monkeys or chimpanzees. Such experiments speak very clearly to the *Null Hypothesis*. A case in point is episodic memory. Episodic memory refers to the recollections of personal experiences of one's life. Tulving (1972) originally envisioned episodic memory as consisting of memory for *what* the event was, *where* the event occurred, and *when* in one's life the event happened, colloquially referred to as WWW memory. Later, Tulving (1985) refined his definition to include the concept of autonoetic consciousness (autonoesis), the phenomenological experience that the memory one retrieves is indeed something that has happened to you in the past. If episodic memory is defined as requiring autonoesis, which can only be accessed by a verbal report, then it is unlikely that any nonhuman animal can satisfy the criterion for possessing episodic memory. However, if we revert to Tulving's (1972) original definition of episodic memory as memory for *what*, *when*, and *where*, then there is accumulating evidence that a variety of animals possess episodic memory, or at least what some have cautiously referred to as *episodic-like* memory.

The Clayton and Dickinson (1998) study still ranks as the most compelling evidence to date that nonhumans, in their case scrub jays, can use *what*, *where*, and *when* information to guide their behavior. Since the publication of the Clayton and Dickinson (1998) study, there have been many other attempts at showing WWW-memory in a number of species such as rats (Bird et al., 2003; Babb and Crystal, 2005; Ergorul and Eichenbaum, 2007), pigeons (Skov-Rackette et al., 2006), monkeys (Hampton et al., 2005), and apes (Schwartz et al., 2002, 2004, 2005; Mulcahy and Call, 2006). In many cases, these experiments have alternative explanations that do not necessitate the attribution of episodic memory (see Colombo and Hayne, 2010). In others, the evidence can be tantalizingly close to that of the Clayton and Dickinson (1998) study with jays (Mulcahy and Call, 2006), but always seems to fall just short of the performance of the jays, although some of the more recent work by Crystal and his colleagues raises rats' abilities on par with those of the jays (for a review see Crystal, 2011). The same is true for the ability to plan for a future need, which was very elegantly shown in jays (Raby et al., 2007) and then chimpanzees (Inoue and Matsuzawa, 2007), and more recently in rats (Crystal, 2013). It is also important to bear in mind that even the Clayton and Dickinson (1998) study is not without its critics who oppose the view that the jays are displaying episodic memory (Suddendorf and Busby, 2003; Suddendorf and Corballis, 2007). Nevertheless, with proper experimental designs in place, it is simply a matter of time before all animals show high levels of proficiency on WWW tasks.

Studies exploring the capacity of animals to display Theory of Mind (ToM) is another example where birds display remarkable abilities. Premack and Woodruff (1978) posed the question: "*Does a chimpanzee have a theory of mind?*" On the basis of the ability of chimpanzees to select the proper picture depicting a solution to a previously seen 30-s video clip of a person facing a dilemma, the authors concluded that chimpanzees do have a ToM. Similarly, Povinelli and colleagues compared two forms of mental state attribution, role reversal and the concept of a knower versus a guesser (Povinelli et al., 1990; Povinelli, 1993). In the case of the role reversal experiment, the chimpanzees were able to appreciate not only their role in securing food, but also that of the human they were paired with, so that if switched to the other's role, they still succeeded in obtaining food. Likewise, in the knower-guesser experiment, the chimpanzees provided evidence that they understood that the person who remained in the room (the knower) had knowledge of the whereabouts of the hidden food, whereas the person that left the room (the guesser) did not, so that when given the choice they chose the location indicated by the knower rather than the guesser.

Although neither the video-clip, role-reversal, nor knower-guesser experiments have been conducted with birds, Emery and Clayton (2001) did examine the effects of experience and social context on the ability of scrub jays to cache food. Jays were given the opportunity to cache food either in the presence of an observer jay or in private. The authors found that jays were far more likely to recache their food if they had previously cached while being observed, suggesting that they understood the intentions of the observing jay. Indeed, only those jays that themselves had experienced pilfering caches displayed such an ability, whereas naïve jays did not recache any more in the observed condition than the in-private condition. These results support the age-old adage that "It takes a thief to know a thief," and highlights the remarkable ability of these birds with respect to mental state attribution.

To be sure there are critics of all these studies, indeed Povinelli (1994) has since conceded the chimpanzees may have learnt to respond to a behavioral cue rather than infer each of the experimenters' knowledge state, a far simpler take on ToM than mental state attribution. In a critique of the ToM literature Heyes (1998, p. 101) evaluated the empirical evidence that chimpanzees possess a ToM and concluded that "in every case where non-human primate behavior has been interpreted as a sign of ToM, it could instead have occurred by chance or as a product of nonmentalistic processes such as associative learning or inference based on nonmental categories." And similarly, the findings of Emery and Clayton (2001) can also be attributed to simple learning processes and associations. Although we subscribe to these simpler interpretations, the main point we wish to make now, however, is that there is no evidence to suggest that a particular capacity such as episodic memory, or ToM (or any of the previous abilities we have discussed) is present in one species and not another.

MACPHAIL REVISITED

Our review is not exhaustive in the sense that we have not examined every task on which species have been compared. For example, how different species perform on habituation, classical conditioning, and instrumental conditioning tasks, what Macphail labeled as "simple" tasks, have been extensively reviewed by Macphail (1982, 1985, 1987), and it was not our intention to go over those again, mainly because there is probably little disagreement that vertebrates perform similarly on such "simple" tasks. Rather, our goal was to evaluate Macphail's (1985) *Null Hypothesis* in light of the recent explosion of interest in the mental abilities of nonhuman animals, and the tasks that have been used to infer these abilities. These tasks are those referred to by Macphail as "complex" tasks, and Macphail recognized that disagreement over his *Null Hypothesis* would focus on these "complex" tasks.

We have reviewed a large number of such "complex" tasks such as reflexivity (matching concept), symmetry, and serial-order behavior, and have shown that differences in performance between species can be traced to a contextual variable, be it the FR requirement to the sample stimulus or the number of training stimuli in the case of reflexivity, aspects of the stimulus response topography in the case of symmetry, or the testing situation in the case of serial-order behavior. For other tasks, such as transitivity, episodic memory, and ToM, the performance of birds rivals, and at times exceeds that of non-human primates. Our review of the literature indicates that there is very little difference in the performance on these "complex" tasks across a range of vertebrate species. On the basis of the above review, and notwithstanding the potential pitfalls inherent in all such comparisons, we agree with Macphail (1985, p. 39) when he stated that "there is currently no phenomenon of learning demonstrable in one (non-human) vertebrate species that has not been found in all other vertebrates in which it has been sought systematically."

Qualitative Differences Versus Quantitative Differences

> *By a qualitative difference between species is meant the possession by one species of a mechanism that is absent in another.... A quantitative difference between two*

> *species would mean that one species used a mechanism or mechanisms common to both species more efficiently than the other," (*Macphail, 1985*, p. 38).*

We do not mean to imply that there are no instances of a particular task in which the performance of one species exceeds that of another. Indeed, there are many such cases. It is hard to escape the fact, therefore, that species do differ quantitatively. The ease with which chimpanzees and monkeys can learn tasks is all too apparent, and although speed of learning is not the best proxy for cognitive abilities, it does speak to some difference in processing capacity, even once issues such as contextual variables are account for. And the mere fact that a pigeon needs a testing situation set up in a specific way, whereas a monkey may not, further speaks to a quantitative difference at the phenomenological level, and possibly also at the process level (see further discussion below). These quantitative differences also surely extend to the range of transfer situations with a more restricted range in pigeons than that seen in monkeys, and indeed a more restricted range in monkeys than that seen in chimpanzees or humans (Weinstein, 1941). And after all, it is the ability to transfer to novel situations, which is really the hallmark of what we call intelligence, and in this respect the abilities of humans exceeds that of monkeys, just as the abilities of monkeys more than likely exceeds that of pigeons. Indeed, we would argue that the main difference in "intelligence" among animals lies in the degree to which one must account for contextual variables, which in turn reflects the level of flexibility of an animal's behavioral repertoire. Surely, the extra cortical tissue of a primate brain, even once one accounts for body size, is what allows it to express behaviors in less restricted manners, and surely that is what lies at the heart of "intelligence."

Associative Processes or Cognitive Processes?

An important point to bear in mind when comparing the performance of species on a particular task is that similar looking graphs do not imply similar underlying processes. Just because a pigeon shows levels of transfer on a matching task similar to that of monkeys, or performs similarly on tests of transitivity, does not mean that it is invoking the same processes to solve the task as a monkey. A similar point was trenchantly put forth by Gallup in his reply to Epstein et al.'s (1981) demonstration of self recognition abilities in pigeons when he stated that "Simply because you can mimic the behavior of one species by reinforcing a series of successive approximations to what looks like the same routine in another, it does not follow that the behavior of the former species necessarily arose in the same way" (Gallup, 1985, p. 633). Although a fair criticism, the simple fact is that there is virtually no evidence to suggest that pigeons are solving complex tasks differently from monkeys, or monkeys differently from chimpanzees, once, of course, contextual variables are taken into consideration. The fact that pigeons, monkeys, and chimpanzees are solving tasks similarly is supported not only by the success-testing metric, but also more importantly by the signature-testing metric, which explores the various signatures of performance on a task (Taylor, 2014; Scarf and Colombo, 2020).

The important question for comparative cognition is not whether an animal can solve a task or not, but rather *how do they solve tasks*? We invoke constructs, such as a *matching concept*, *symmetry*, *transitivity*, and *orthographic processing* as if these constructs are explanations of behavior. They are not, they are just labels for a behavior. Epstein et al. (1981, p. 696) put it beautifully when they said that "such constructs impede the search for the controlling variables of the behavior they are said to procedure." The temptation to richly interpret an animal's behavior is pervasive (Haith, 1998; Shettleworth, 2010). Speaking for our own research, we may argue that pigeons have a matching concept (Colombo et al., 2003), abstract numerical abilities (Scarf et al., 2011), and orthographic processing (Scarf et al., 2016), but we do not believe that pigeons (or monkeys) succeed on such tasks because they have advanced cognitive skills. Rather we use these constructs, much in the same way that Skinner, Epstein, and their colleagues used them in the Columban simulation studies (self-awareness: Epstein et al., 1981; symbolic communication: Epstein et al., 1980; insight: Epstein et al., 1984), to mimic the constructs that have been used with primates, for whom we feel much more comfortable adopting such labels.

If not "cognitive" processes, then what processes underlies these impressive abilities? We surely underestimate the power of simpler (but not simple) accounts such as associative learning or reinforcement-learning processes (Dickinson, 2012; Hanus, 2016; Haselgrove, 2016). We doubt that our pigeons (or the monkeys) are truly engaging in "orthographic processing" and breaking down each four-letter word they see into its constituent pairs, and evaluating the frequency with which each pair is likely to occur in words or nonwords (Grainger et al., 2012). Rather, we agree with Vokey and Jamieson (2015, see also Linke et al., 2017) that the monkeys and the birds are likely mapping novel words onto prototypic "word" and "nonword" templates, an impressive and certainly not a simple ability to be sure, but one that differs from an "orthographic" account. Similarly, we might invoke "mental representations" as processes governing the behavior of pigeons and monkeys on a transitivity task, but simpler accounts such as Value Transfer Theory and Binary Sampling Model go a long way to explain the behavior. True these simpler accounts may not explain every nuance of a behavior that has been observed (and they should), but how much of that might reflect our lack of understanding of these simpler accounts, as opposed to a shortcoming of these simpler accounts?

The issue we touch on above is a critical issue for comparative cognition, and it is impossible to do it justice as a side note of a few paragraphs. We agree with Allen (2014, p. 76) that there is too much "trophy hunting," and that those theories that are available are not formalized to a sufficient degree to truly untangle the difference between associative and cognitive models of behavior. But models are critically important if we are to advance the field, especially process-based models (Luce, 1995; Buckner, 2011). That said, models themselves are not without their limitations. For example, Smith et al. (2016) note that associative models based mainly on reinforcement principles, and cognitive models based mainly on uncertainty responses, are mathematically the same, and that unless one wishes to invoke Morgan's canon, there is little reason to accept one over the other. We take a different view that perhaps the reason these models are mathematically identical is because the processes underlying them are not as different as we think; surely uncertainty monitoring is intimately tied to not only our recent but also our remote reinforcement history. As Crystal (2011, p. 417) states "if an uncertainty response was never reinforced, it seems unlikely that it would be produced by the subject, and it seems virtually impossible that it would be used functionally to express uncertainty or escape a difficult trial."

CONCLUDING COMMENTS

We have reviewed a number of studies, and we hopefully have convinced the reader that in situations where one species outperforms another the reason can often be traced to contextual variables. Macphail (1985) concluded that he did not overestimate the importance of contextual variables, and more than three decades later we would agree that contextual variables do underlie many of the differences in performance seen across species. In a companion paper (Scarf and Colombo, 2020), we have also shown that the similarities extend not only to performance on a task, but also the signatures that underlie successful performance on a task. Taken together, we fully support Macphail's view that there are at least no qualitative differences across vertebrate species, and certainly none between birds and monkeys. On the other hand, we think there is ample support for the view that there are quantitative differences across species. Perhaps by perceiving the world through a quantitative lens of differences of degree, we can better tackle the divide between associative processes and cognitive processes.

AUTHOR CONTRIBUTIONS

MC and DS conceptualized and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Research was funded by a Royal Society of New Zealand Marsden grant 19-UOO-162 to MC.

REFERENCES

Allen, C. (2014). Models, mechanisms, and animal minds. *South. J. Philos.* 52, 75–97. doi: 10.1111/sjp.12072

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., et al. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* 513, 532–541. doi: 10.1002/cne.21974

Babb, S. J., and Crystal, J. D. (2005). Discrimination of what, when, and where: implications for episodic-like memory in rats. *Learn. Motiv.* 36, 177–189. doi: 10.1016/j.lmot.2005.02.009

Barlow, H. B. (1987). Efficiency, versatility, cognitive maps, and language. *Behav. Brain Sci.* 10, 657–658. doi: 10.1017/S0140525X0005500X

Bird, L. R., Roberts, W. A., Abroms, B., Kit, K. K., and Crupi, C. (2003). Spatial memory for food hidden by rats (*Rattus norvegicus*) on the radial maze: studies of memory for where, what, and when. *J. Comp. Psychol.* 117, 176–187. doi: 10.1037/0735-7036.117.2.176

Bitterman, M. E. (1965). The evolution of intelligence. *Sci. Am.* 212, 92–100. doi: 10.1038/scientificamerican0165-92

Blaisdell, A. P., and Cook, R. G. (2005). Two-item same-different concept learning in pigeons. *Learn. Behav.* 33, 67–77. doi: 10.3758/BF03196051

Bryant, P., and Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature* 232, 456–458. doi: 10.1038/232456a0

Buckner, C. (2011). Two approaches to the distinction between cognition and 'mere association'. *Int. J. Comp. Psychol.* 24, 314–348.

Carter, D. E., and Werner, T. J. (1978). Complex learning and information processing by pigeons: a critical analysis. *J. Exp. Anal. Behav.* 29, 565–601. doi: 10.1901/jeab.1978.29-565

Cheney, D., Seyfarth, R., and Smuts, B. (1986). Social relationships and social cognition in nonhuman primates. *Science* 234, 1361–1366. doi: 10.1126/science.3538419

Clayton, N. S., and Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* 395, 272–274. doi: 10.1038/26216

Colombo, M., Cottle, A., and Frost, N. (2003). Degree of representation of the matching concept in pigeons (*Columba livia*). *J. Comp. Psychol.* 117, 246–256. doi: 10.1037/0735-7036.117.3.246

Colombo, M., and Hayne, H. (2010). "Episodic memory: comparative and developmental issues" in *Oxford handbook of developmental and comparative neuroscience*. eds. M. S. Blumberg, J. H. Freeman and S. R. Robinson (New York, NY: Oxford University Press), 617–636.

Cook, R. G., Katz, J. S., and Cavoto, B. R. (1997). Pigeon same-different concept learning with multiple stimulus classes. *J. Exp. Psychol. Anim. Behav. Process.* 23, 417–433. doi: 10.1037/0097-7403.23.4.417

Crystal, J. D. (2011). Navigating the interface between learning and cognition. *Int. J. Comp. Psychol.* 24, 412–436.

Crystal, J. D. (2013). Remembering the past and planning for the future in rats. *Behav. Processes* 93, 39–49. doi: 10.1016/j.beproc.2012.11.014

Cumming, W. W., and Berryman, R. (1961). Some data on matching behavior in the pigeon. *J. Exp. Anal. Behav.* 4, 281–284. doi: 10.1901/jeab.1961.4-281

D'Amato, M. R., and Colombo, M. (1988). Representation of serial order in monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. Behav. Process.* 14, 131–139. doi: 10.1037/0097-7403.14.2.131

D'Amato, M. R., and Colombo, M. (1989). On the limits of the matching concept in monkeys (*Cebus apella*). *J. Exp. Anal. Behav.* 52, 225–236. doi: 10.1901/jeab.1989.52-225

D'Amato, M. R., and Salmon, D. P. (1984). "Cognitive processes in Cebus monkeys" in *Animal cognition*. eds. H. L. Roitblat, T. G. Bever and H. S. Terrace (Hillsdale, New Jersey: Erlbaum).

D'Amato, M. R., Salmon, D. P., and Colombo, M. (1985a). Extent and limits of the matching concept in monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. Behav. Process.* 11, 35–51. doi: 10.1037//0097-7403.11.1.35

D'Amato, M. R., Salmon, D. P., Loukas, E., and Tomie, A. (1985b). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *J. Exp. Anal. Behav.* 44, 35–47. doi: 10.1901/jeab.1985.44-35

Davis, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *J. Comp. Psychol.* 106, 342–349. doi: 10.1037/0735-7036.106.4.342

Delius, J. D., and Siemann, M. (1998). Transitive responding in animals and humans: exaptation rather than adaptation. *Behav. Processes* 42, 107–137. doi: 10.1016/S0376-6357(97)00072-7

Dickinson, A. (2012). Associative learning and animal cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2733–2742. doi: 10.1098/rstb.2012.0220

Edwards, C. A., Jagielo, J. A., and Zentall, T. R. (1983). "Same/different" symbol use by pigeons. *Anim. Learn. Behav.* 11, 349–355. doi: 10.3758/BF03199787

Elepfandt, A. (1987). Comparative cognition: inadequate approach, precipitate conclusions. *Behav. Brain Sci.* 10, 661–662. doi: 10.1017/S0140525X00055059

Emery, N. J., and Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature* 414, 443–446. doi: 10.1038/35106560

Epstein, R., Kirshnit, C. E., Lanza, R. P., and Rubin, L. C. (1984). 'Insight' in the pigeon: antecedents and determinants of an intelligent performance. *Nature* 308, 61–62. doi: 10.1038/308061a0

Epstein, R., Lanza, R. P., and Skinner, B. F. (1980). Symbolic communication between two pigeons (*Columbia livia domestica*). *Science* 207, 543–545. doi: 10.1126/science.207.4430.543

Epstein, R., Lanza, R. P., and Skinner, B. F. (1981). "Self-awareness" in the pigeon. *Science* 212, 695–696. doi: 10.1126/science.212.4495.695

Ergorul, C., and Eichenbaum, H. (2007). The hippocampus and memory for "what," "where," and "when". *Learn. Mem.* 11, 397–405. doi: 10.1101/lm.73304

Farthing, G. W., and Opuda, M. J. (1974). Transfer of matching-to-sample in pigeons. *J. Exp. Anal. Behav.* 21, 199–213. doi: 10.1901/jeab.1974.21-199

Frank, A. J., and Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *J. Exp. Anal. Behav.* 84, 147–165. doi: 10.1901/jeab.2005.115-04

Gallup, G. G. (1985). Do minds exist in species other than our own? *Neurosci. Biobehav. Rev.* 9, 631–641. doi: 10.1016/0149-7634(85)90010-7

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *J. Exp. Psychol. Anim. Behav. Process.* 7, 150–164. doi: 10.1037/0097-7403.7.2.150

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., and Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature* 410, 930–933. doi: 10.1038/35073582

Goldman-Rakic, P. S., and Preuss, T. M. (1987). Wither comparative psychology? *Behav. Brain Sci.* 10, 666–667. doi: 10.1017/S0140525X00055102

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., and Fagot, J. (2012). Orthographic processing in baboons (*Papio papio*). *Science* 336, 245–248. doi: 10.1126/science.1218152

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behav. Dev.* 21, 167–179. doi: 10.1016/S0163-6383(98)90001-7

Hampton, R. R., Hampstead, B. M., and Murray, E. A. (2005). Rhesus monkeys (*Macaca mulatta*) demonstrate robust memory for what and where, but not when, in an open-field test of memory. *Learn. Motiv.* 36, 245–259. doi: 10.1016/j.lmot.2005.02.004

Hanus, D. (2016). Casual reasoning versus associative learning: a useful dichotomy or a strawman battle in comparative psychology. *J. Comp. Psychol.* 130, 241–248. doi: 10.1037/a0040235

Haselgrove, M. (2016). Overcoming associative learning. *J. Comp. Psychol.* 130, 226–240. doi: 10.1037/a0040180

Hayes, S. C. (1989). Nonhumans have not yet shown stimulus equivalence. *J. Exp. Anal. Behav.* 51, 385–392. doi: 10.1901/jeab.1989.51-385

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behav. Brain Sci.* 21, 101–148. doi: 10.1017/S0140525X98000703

Hodos, W. (1987). Animal general intelligence: an idea ahead of its time. *Behav. Brain Sci.* 10:668. doi: 10.1017/S0140525X00055138

Hogan, D. E., and Zentall, T. R. (1977). Backward associations in the pigeon. *Am. J. Psychol.* 90, 3–15. doi: 10.2307/1421635

Holmes, P. W. (1979). Transfer of matching performance in pigeons. *J. Exp. Anal. Behav.* 31, 103–114. doi: 10.1901/jeab.1979.31-103

Inoue, S., and Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Curr. Biol.* 17, R1004–R1005. doi: 10.1016/j.cub.2007.10.027

Lazareva, O. F., and Wasserman, E. A. (2012). Transitive inference in pigeons: measuring the associative values of stimuli B and D. *Behav. Processes* 89, 244–255. doi: 10.1016/j.beproc.2011.12.001

Linke, M., Bröker, F., Ramscar, M., and Baayen, H. (2017). Are baboons learning "orthographic" representations? Probably not. *PLoS One* 12:e0183876. doi: 10.1371/journal.pone.0183876

Lionello-DeNolf, K. M., and Urcuioli, P. J. (2002). Stimulus control topographies and test of symmetry in pigeons. *J. Exp. Anal. Behav.* 78, 467–495. doi: 10.1901/jeab.2002.78-467

Lipkens, R., Kop, P. F., and Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances of pigeons. *J. Exp. Anal. Behav.* 49, 395–409. doi: 10.1901/jeab.1988.49-395

Luce, R. D. (1995). Four tensions concerning mathematical modelling in psychology. *Annu. Rev. Psychol.* 46, 1–27. doi: 10.1146/annurev.ps.46.020195.000245

Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Oxford, England: Clarendon Press.

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 37–51. doi: 10.1098/rstb.1985.0008

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695. doi: 10.1017/S0140525X00054984

Masure, R. H., and Allee, W. C. (1934). The social order in flocks of the common chicken and pigeons. *Auk* 51, 306–327. doi: 10.2307/4077659

McGonigle, B. O., and Chalmers, M. (1977). Are monkeys logical? *Nature* 267, 694–696. doi: 10.1038/267694a0

McIlvane, W. J., Serna, R. W., Dube, W. V., and Stromer, R. (2000). "Stimulus control topography coherence and stimulus equivalence: reconciling test outcomes with theory" in *Experimental and applied analysis of human behaviour*. eds. J. Leslie and D. E. Blackman (Reno, NV: Context Press), 85–110.

McIntire, R. D., Cleary, J., and Thompson, T. (1987). Conditional relations by monkeys: reflexivity, symmetry, and transitivity. *J. Exp. Anal. Behav.* 47, 279–285. doi: 10.1901/jeab.1987.47-279

Mello, N. K. (1971). Alcohol effects on delayed matching to sample performance by rhesus monkeys. *Physiol. Behav.* 7, 77–101. doi: 10.1016/0031-9384(71)90239-3

Menzel, R. (2012). The honeybee as a model for understanding the basis of cognition. *Nat. Rev. Neurosci.* 13, 758–768. doi: 10.1038/nrn3357

Milner, A. D. (1973). Matching within and between sense modalities in the monkey (*Macaca mulatta*). *J. Comp. Physiol. Psychol.* 83, 278–284. doi: 10.1037/h0034427

Moreno, A. M., de Souza, D. D. G., and Reinhard, J. (2012). A comparative study of relational learning capacity in honeybees (*Apis mellifera*) and stingless bees (*Melipona rufiventris*). *PLoS One* 7:e51467. doi: 10.1371/journal.pone.0051467

Mulcahy, N. J., and Call, J. (2006). Apes save tools for future use. *Science* 312, 1038–1040. doi: 10.1126/science.1125456

Nissen, H. W., Blum, J. S., and Blum, R. A. (1948). Analysis of matching behavior in chimpanzee. *J. Comp. Physiol. Psychol.* 41, 62–74. doi: 10.1037/h0061040

Oden, D. L., Thompson, R. K. R., and Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes*). *J. Exp. Psychol. Anim. Behav. Process.* 14, 140–145. doi: 10.1037/0097-7403.14.2.140

Paz-y-Miño C, G., Bond, A. B., Kamil, A. C., and Balda, R. P. (2004). Pinyon jays use transitive inference to predict social dominance. *Nature* 430, 778–781. doi: 10.1038/nature02723

Penn, D. C., Holyaok, K. J., and Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 109–130. doi: 10.1017/S0140525X08003543

Piaget, J. (1928). *Judgement and reasoning in the child*. London, England: Routledge Chapman & Hall.

Povinelli, D. J. (1993). Reconstructing the evolution of mind. *Am. Psychol.* 48, 493–509. doi: 10.1037/0003-066X.48.5.493

Povinelli, D. J. (1994). Comparative studies of animal mental state attribution: a reply to Heyes. *Anim. Behav.* 48, 239–241. doi: 10.1006/anbe.1994.1233

Povinelli, D. J., Nelson, K. E., and Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 104, 203–210. doi: 10.1037/0735-7036.104.3.203

Premack, D. (1983). The codes of man and beasts. *Behav. Brain Sci.* 6, 125–136. doi: 10.1017/S0140525X00015077

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/S0140525X00076512

Premack, D., Woodruff, G., and Kennel, K. (1978). Paper-marking test for chimpanzees: simple control for social cues. *Science* 202, 903–905. doi: 10.1126/science.202.4370.903

Raby, C. R., Alexis, D. M., Dickinson, A., and Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature* 445, 919–921. doi: 10.1038/nature05575

Richards, R. W. (1988). The question of bidirectional associations in pigeons' learning of conditional discrimination tasks. *Bull. Psychon. Soc.* 26, 577–579. doi: 10.3758/BF03330126

Robinson, J. S. (1955). The sameness-difference discrimination problem in chimpanzee. *J. Comp. Physiol. Psychol.* 48, 195–197. doi: 10.1037/h0042463

Sands, S. F., and Wright, A. A. (1980). Serial probe recognition performance by a rhesus money and a human with 10- and 20-item lists. *J. Exp. Psychol. Anim. Behav. Process.* 6, 386–396. doi: 10.1037/0097-7403.6.4.386

Santiago, H. C., and Wright, A. A. (1984). Pigeon memory: same/different concept learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. *J. Exp. Psychol. Anim. Behav. Process.* 10, 498–512. doi: 10.1037/0097-7403.10.4.498

Scarf, D., Boy, K., Uber Reinert, A., Devine, J., Güntürkün, O., and Colombo, M. (2016). Orthographic processing in pigeons (*Columba livia*). *Proc. Natl. Acad. Sci. U. S. A.* 113, 11272–11276. doi: 10.1073/pnas.1607870113

Scarf, D., and Colombo, M. (2008). Representation of serial order: a comparative analysis of humans, monkeys, and pigeons. *Brain Res. Bull.* 76, 307–312. doi: 10.1016/j.brainresbull.2008.02.022

Scarf, D., and Colombo, M. (2010). Representation of serial order in pigeons (*Columba livia*). *J. Exp. Psychol. Anim. Behav. Process.* 36, 423–429. doi: 10.1037/a0020926

Scarf, D., and Colombo, M. (2020). Columban simulation project 2.0: numerical competence and orthographic processing in pigeons and primates. *Front. Psychol.* 10:3017. doi: 10.3389/fpsyg.2019.03017

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334:1664. doi: 10.1126/science.1213357

Schwartz, B. L., Colon, M. R., Sanchez, I. C., Rodriguez, I. A., and Evans, S. (2002). Single-trial learning of "what" and "who" information in a gorilla (*Gorilla gorilla gorilla*): implications for episodic memory. *Anim. Cogn.* 5, 85–90. doi: 10.1007/s10071-002-0132-0

Schwartz, B. L., Hoffman, M. L., and Evans, S. (2005). Episodic-like memory in a gorilla: a review and new findings. *Learn. Motiv.* 36, 226–244. doi: 10.1016/j.lmot.2005.02.012

Schwartz, B. L., Meissner, C. M., Hoffman, M., Evans, S., and Frazier, L. D. (2004). Event memory and misinformation effects in a gorilla (*Gorilla gorilla gorilla*). *Anim. Cogn.* 7, 93–100. doi: 10.1007/s10071-003-0194-7

Shettleworth, S. J. (1998). *Cognition, evolution and behavior*. New York, NY: Oxford University Press.

Shettleworth, S. J. (2010). Clever animals and killjoy explanations in comparative psychology. *Trends Cogn. Sci.* 14, 477–481. doi: 10.1016/j.tics.2010.07.002

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperative.

Sidman, M. (2018). What is interesting about equivalence relations and behavior? *Perspect. Behav. Sci.* 41, 33–43. doi: 10.1007/s40614-018-0147-8

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., and Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *J. Exp. Anal. Behav.* 37, 23–44. doi: 10.1901/jeab.1982.37-23

Siemann, M., Delius, J. D., and Wright, A. A. (1996). Transitive responding in pigeons: influences of stimulus frequency and reinforcement history. *Behav. Processes* 37, 185–195. doi: 10.1016/0376-6357(96)00020-4

Skinner, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* 57, 193–216. doi: 10.1037/h0054367

Skov-Rackette, S. I., Miller, N. Y., and Shettleworth, S. J. (2006). What-where-when memory in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 32, 345–358. doi: 10.1037/0097-7403.32.4.345

Smith, J. D., Zakrzewski, A. C., and Church, B. A. (2016). Formal models in animal-metacognition research: the problem of interpreting animals' behavior. *Psychon. Bull. Rev.* 23, 1341–1353. doi: 10.3758/s13423-015-0985-2

Steirn, J. N., Weaver, J. E., and Zentall, T. R. (1995). Transitive inference in pigeons: simplified procedures and a test of value transfer theory. *Anim. Learn. Behav.* 23, 76–82. doi: 10.3758/BF03198018

Sternberg, R. J. (1987). Difficulties in comparing intelligence across species. *Behav. Brain Sci.* 10, 679–680. doi: 10.1017/S0140525X00055278

Straub, R. O., and Terrace, H. S. (1981). Generalization of serial learning in the pigeon. *Anim. Learn. Behav.* 9, 454–468. doi: 10.3758/BF03209775

Suddendorf, T., and Busby, J. (2003). Mental time travel in animals? *Trends Cogn. Sci.* 7, 391–396. doi: 10.1016/s1364-6613(03)00187-6

Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: what is mental time travel and is it unique to humans? *Behav. Brain Sci.* 30, 299–313. doi: 10.1017/S0140525X07001975

Taylor, A. H. (2014). Corvid cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 361–372. doi: 10.1002/wcs.1286

Terrace, H. S. (1993). The phylogeny and ontogeny of serial memory: list learning by pigeons and monkeys. *Psychol. Sci.* 4, 162–169. doi: 10.1111/j.1467-9280.1993.tb00481.x

Tibbetts, E. A., Agudelo, J., Pandit, S., and Riojas, J. (2019). Transitive inference in *Polistes* paper wasps. *Biol. Lett.* 15:20190015. doi: 10.1098/rsbl.2019.0015

Tomonaga, M., Matsuzawa, T., Fujita, K., and Yamamoto, J. (1991). Emergence of symmetry in visual conditional discrimination by chimpanzees (*Pan troglodytes*). *Psychol. Rep.* 68, 51–60. doi: 10.2466/pr0.1991.68.1.51

Tulving, E. (1972). "Episodic and semantic memory" in *Organization of memory*. eds. E. Tulving and W. Donaldson (New York: Academic Press), 382–403.

Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1–12. doi: 10.1037/h0080017

Urcuioli, P. J., and Nevin, J. A. (1975). Transfer of hue matching in pigeons. *J. Exp. Anal. Behav.* 24, 149–155. doi: 10.1901/jeab.1975.24-149

Vaughan, W. (1988). Formation of equivalence sets in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 14, 36–42. doi: 10.1037/0097-7403.14.1.36

Vokey, J. R., and Jamieson, R. K. (2015). A visual-familiarity account of evidence for orthographic processing in baboons (*Papio papio*). *Psychol. Sci.* 25, 991–996. doi: 10.1177/0956797613516634

von Fersen, L., Wynne, C. D. L., Delius, J. D., and Staddon, J. E. R. (1990). Deductive reasoning in pigeons. *Naturwissenschaften* 77, 548–549. doi: 10.1007/BF01139271

von Fersen, L., Wynne, C. D. L., Delius, J. D., and Staddon, J. E. R. (1991). Transitive inference formation in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 17, 334–341. doi: 10.1037/0097-7403.17.3.334

Wasserman, E. A., Hugart, J. A., and Kirkpatrick-Steger, K. (1995). Pigeons show same-different conceptualization after training with complex visual stimuli. *J. Exp. Psychol. Anim. Behav. Process.* 21, 248–252. doi: 10.1037//0097-7403.21.3.248

Weinstein, B. (1941). Matching-from-sample by rhesus monkeys and by children. *J. Comp. Psychol.* 31, 195–213. doi: 10.1037/h0063449

Wilson, B., Mackintosh, N. J., and Boakes, R. A. (1985a). Matching and oddity learning in the pigeon: transfer effects and the absence of relational learning. *Q. J. Exp. Psychol. B* 37, 295–311. doi: 10.1080/14640748508401172

Wilson, B., Mackintosh, N. J., and Boakes, R. A. (1985b). Transfer of relational rules in matching and oddity learning by pigeons and corvids. *Q. J. Exp. Psychol. B* 37, 313–332. doi: 10.1080/14640748508401173

Wright, A. A. (1997). Concept learning and learning strategies. *Psychol. Sci.* 8, 119–123. doi: 10.1111/j.1467-9280.1997.tb00693.x

Wright, A. A., and Cummings, W. W. (1971). Color-naming functions for the pigeon. *J. Exp. Anal. Behav.* 15, 7–17. doi: 10.1901/jeab.1971.15-7

Yamamoto, J., and Asano, T. (1995). Stimulus equivalence in a chimpanzee (*Pan troglodytes*). *Psychol. Rec.* 45, 3–21.

Young, M. E., Wasserman, E. A., and Garner, K. L. (1997). Effects of number of items on the pigeon's discrimination of same from different visual displays. *J. Exp. Psychol. Anim. Behav. Process.* 23, 491–501. doi: 10.1037/0097-7403.23.4.491

Zentall, T. R., and Hogan, D. E. (1974). Abstract concept learning in the pigeon. *J. Exp. Psychol.* 102, 393–398. doi: 10.1037/h0035970

Zentall, T. R., and Hogan, D. E. (1976). Pigeons can learn identity or difference, or both. *Science* 191, 408–409. doi: 10.1126/science.191.4225.408

Zentall, T. R., and Hogan, D. E. (1978). Same/different concept learning in the pigeon: the effect of negative instances and prior adaptation to transfer stimuli. *J. Exp. Anal. Behav.* 30, 177–186. doi: 10.1901/jeab.1978.30-177

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 31 August 2020
doi: 10.3389/fpsyg.2020.01835

Validity of Cognitive Tests for Non-human Animals: Pitfalls and Prospects

Michèle N. Schubiger[1,2]*, Claudia Fichtel[3,4] and Judith M. Burkart[1]

[1] Evolutionary Cognition Group, Department of Anthropology, University of Zurich, Zurich, Switzerland, [2] World Ape Fund, London, United Kingdom, [3] Behavioural Ecology and Sociobiology Unit, German Primate Center, Göttingen, Germany, [4] Leibniz ScienceCampus "Primate Cognition", Göttingen, Germany

Comparative psychology assesses cognitive abilities and capacities of non-human animals and humans. Based on performance differences and similarities in various species in cognitive tests, it is inferred how their minds work and reconstructed how cognition might have evolved. Critically, such species comparisons are only valid and meaningful if the tasks truly capture individual and inter-specific variation in cognitive abilities rather than contextual variables that might affect task performance. Unlike in human test psychology, however, cognitive tasks for non-human primates (and most other animals) have been rarely evaluated regarding their measurement validity. We review recent studies that address how non-cognitive factors affect performance in a set of commonly used cognitive tasks, and if cognitive tests truly measure individual variation in cognitive abilities. We find that individual differences in emotional and motivational factors primarily affect performance via attention. Hence, it is crucial to systematically control for attention during cognitive tasks to obtain valid and reliable results. Aspects of test design, however, can also have a substantial effect on cognitive performance. We conclude that non-cognitive factors are a minor source of measurement error if acknowledged and properly controlled for. It is essential, however, to validate and eventually re-design several primate cognition tasks in order to ascertain that they capture the cognitive abilities they were designed to measure. This will provide a more solid base for future cognitive comparisons within primates but also across a wider range of non-human animal species.

Keywords: comparative psychology, cognitive testing, test design, task validity, contextual variables, non-cognitive factors, individual differences, primates

Edited by:
Michael Colombo,
University of Otago, New Zealand

Reviewed by:
Brent Alsop,
University of Otago, New Zealand
Jennifer Vonk,
Oakland University, United States
Michael Beran,
Georgia State University,
United States

****Correspondence:***
Michèle N. Schubiger
mnschubiger@protonmail.com

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *30 April 2020*
Accepted: *03 July 2020*
Published: *31 August 2020*

Citation:
Schubiger MN, Fichtel C and Burkart JM (2020) Validity of Cognitive Tests for Non-human Animals: Pitfalls and Prospects.
Front. Psychol. 11:1835.
doi: 10.3389/fpsyg.2020.01835

INTRODUCTION

Comparative psychologists design and use cognitive tests to investigate and compare cognitive performance and capacities of extant non-human animal species and humans (mainly children). The ultimate goal is to better understand how animal minds are organized and to reconstruct the evolution of mind, including the human one. To date, a variety of animal species has been compared in cognitive studies, ranging from mammals (terrestrial and aquatic ones) to birds (e.g., Lambert et al., 2018), fish (e.g., Bhsary and Brown, 2014), reptiles (e.g., Wilkinson and Huber, 2012) and invertebrates such as cephalopods (mainly octopuses: Mather and Kuba, 2013;

Amodio et al., 2019) or insects (mainly bees: Chittka, 2017; Solvi et al., 2020). The rationale behind selecting certain species for cognitive studies is typically a low or high degree of variation in brain size or socio-ecological factors such as breeding systems, social structures or feeding ecologies, to better understand the selective pressures driving cognitive abilities. Of particular interest for comparative psychologists are cognitive comparisons between larger-brained species (e.g., non-human primates, elephants, dolphins, and birds from the corvid and parrot families; Tomasello and Call, 1997; Pepperberg, 2002; Emery and Clayton, 2004; Plotnik et al., 2006; Byrne et al., 2009; Maestripieri, 2012; Manger, 2013; Güntürkün, 2014) and smaller-brained species (e.g., rodents, pigeons: Scarf et al., 2011; Matzel and Sauce, 2017) in order to examine the cognitive potential of large brains. More recently, an increasing number of mammalian carnivore taxa are also being studied to better understand the cognitive abilities of this large and in many aspects heterogenous order (e.g., domestic and wild dogs, hyeanas, bears, and meerkats: Townsend et al., 2012; Bensky et al., 2013; Holekamp and Benson-Amram, 2017; Dale et al., 2019; for a review see: Vonk and Leete, 2017).

Valid species comparisons pose an immense challenge for comparative psychology, because obviously, applying a physically identical task is not sufficient to warrant fair species comparisons, which led some researchers to argue this may render meaningful species comparisons impossible (e.g., MacPhail, 1987). A central challenge of contemporary comparative psychology is that both the nature of cognitive abilities and their potential evolution need to be inferred and reconstructed from individual performance scores in human-devised cognitive tests. In comparative psychology, such cognitive tasks often address cognitive abilities from broad domains such as physical cognition to deal with the non-social world and social cognition to deal with the social world (Tomasello and Call, 1997). However, when dealing with more cognitively demanding problems, such as completely novel ones, an individual is required to draw on cognitive resources that can be applied across a wide range of problems from diverse domains. In primates, brain size predicts species-difference in such a global domain-general cognitive ability, which is reflected in an IQ-like performance score of a species (*G*) across a set of diverse tasks and this performance score increases from humans' evolutionary most distant primate relatives (lowest scores in strepsirrhines) to their closest primate relatives (highest scores in haplorrhines and particularly in great apes; Deaner et al., 2006; Reader et al., 2011). At least in humans, such a domain-general intellectual ability is unequivocal at the individual level (*g*) and increasing evidence in non-human animals suggests considerable evolutionary continuity of *g* (as assessed via reasoning ability and behavioral flexibility), at least within the mammalian lineage (mainly rodents and dogs: as reviewed in Burkart et al., 2017a) including some primates (e.g., tamarins: Banerjee et al., 2009; chimpanzees: Hopkins et al., 2014; orangutans: Damerius et al., 2019). However, among primates, the current evidence is still mixed and controversial with a number of studies finding no support for a *g* (e.g., Herrmann et al., 2007; Amici et al., 2012). This could be consistent with Macphail's null hypothesis that no major domain-general species differences are expected in primates other than humans (MacPhail, 1985), but could also be an artifact of differences in tests and procedures. In order to empirically assess such potential intra-species differences in domain-general cognition, the focus of comparative cognition studies has recently been shifted from using single tasks toward cognitive test batteries (sets of at least five tasks from various cognitive "domains"). Gaining a better understanding of individual differences will help to achieve the ultimate goal of comparative psychology: meaningful comparisons across a wide range of species.

Cognitive testing of non-human primates and other animals poses several issues for comparative psychology, which become particularly apparent when using test batteries. In this paper, we will therefore use established cognitive test batteries to illustrate some of the most prominent issues researchers generally encounter when testing non-human primates but also other animals. Two of these issues are that the outcomes of cognitive studies should be replicable within individuals or single species and that the cognitive tasks should reliably capture variation in the cognitive abilities they were designed to measure.

Are the Outcomes of Cognitive Studies With Non-human Primates Replicable?

It is conceivable that at least some comparative cognition tasks do not measure the cognitive abilities they were intended to measure. Besides aspects of the test-design including task sensitivity (i.e., the level of task difficulty should be appropriate to detect individual variation in cognitive performance without producing ceiling or floor effects), several non-cognitive biases may affect the outcomes of both original studies and their replications. Human cognitive test batteries are regularly evaluated regarding psychometric criteria such as their validity (e.g., construct validity: whether the cognitive tasks actually do measure the supposed cognitive abilities their human developers attempted to measure), and reliability (e.g., test-retest-reliability: whether repeated administration of the same tasks to the same participants leads to reproducible outcomes). In comparative psychology with non-human primates and other animals, however, such evaluations are currently largely missing. This might be problematic, especially in light of the recent replication crisis in human psychology highlighting that many original research findings, even from studies using established experimental paradigms, cannot be replicated by fellow researchers (Open Science Collaboration, 2012; Pashler and Wagenmaekers, 2012; Earp and Trafimow, 2015). One consequence of this crisis was a lack of confidence in scientific studies including the used research practices. Comparative psychology has so far largely been spared by this crisis of confidence, but it may be even more susceptible than human psychology owing to its lack of replication studies (e.g., see Farrar et al., 2020 for specific challenges in this field).

The few primate studies that assessed the re-test reliability of primate cognition tests include one that used different test sessions within single cognitive tasks of the Tamarin Test Battery (TTB) as a proxy for repeated testing and found a relatively low correlation between performance scores (Banerjee et al., 2009).

In an adapted version of the original memory task from the TTB, marmosets were tested with a no-delay session before and after the six different delay conditions (Schubiger et al., 2016). Although they performed in both above chance level, their performance declined from the former to the latter. This suggested that although the marmosets were still able to solve this very simple task, they might not have been fully motivated anymore to make correct choices after having completed the full memory task with five delay sessions. However, in another study, Hopkins et al. (2014) re-tested a subset of their chimpanzee sample with 13 tasks of the Primate Cognition Test Battery (PCTB, Herrmann et al., 2007) 2 years later and found that overall, performance remained relatively stable (besides improvements in four physical and decreased performance in one social PCTB task). In addition, a recent meta-analysis on the repeatability of cognitive performance from 25 species of six animal classes revealed that cognitive performance could be replicated in both temporal repeatability by comparing performance on several exposures of the same task and contextual repeatability by comparing individual performance on different tasks that measure the same putative cognitive ability (Cauchoix et al., 2018). Hence, to assess the validity of cognitive tests it is important to include, if possible, assessments of repeatability.

In comparative psychology, large samples, particularly of non-human primates, are often not available (ManyPrimates et al., 2019a,b), which limits the statistical power of an empirical study (but see Smith and Little, 2018 for advantages of small sample sizes). The next best option to validate obtained test results in such cases are replications in independent samples of the same species. Some findings were replicated within species and were also found in evolutionarily closely related species, which also establishes the external validity of an optimized task (e.g., a memory task that is often part of test batteries to establish *g*: Schubiger et al., 2016, see also **Table 1**).

In comparative psychology, the construct validity of cognitive tasks appears particularly important when assessing individual and species differences in one cognitive ability or domain (e.g., inhibitory control: MacLean et al., 2014). However, this is arguably not equally the case when individuals and species are tested with test batteries, i.e., sets of cognitive tasks that assess performance in various cognitive abilities that are not isolated from each other (but may overlap to some degree; Huber, 2017; Ramus, 2017), provided these tasks do measure aspects of cognition. Therefore, the central issue in comparative psychology is to establish internal task validity more generally, that is whether a task truly measures individual and species differences in the cognitive abilities it is supposed to measure rather than variation in factors that are not primarily of cognitive nature and might bias the outcomes of a study. Such biases may be especially problematic when the same individuals are tested with test batteries consisting of several tasks that might not be controlled for confounding factors. In test batteries that are used to evaluate general intelligence (*g*), the main question is whether individual performance is correlated across tasks, and such correlations can be the result of confounds rather than a true positive manifold (i.e., a positive correlation of unbiased individual cognitive performance scores; Burkart et al., 2017a). While some confounding factors are overt and obvious and can be relatively easily dealt with, such as sensory-motor differences, other factors may not be as straightforward.

Overtly Necessary Preconditions for Valid Species Comparisons

Evolutionarily distantly related vertebrate taxa such as mammals and birds whose cognitive performance is often compared using the same tasks and test-setups, can vary greatly in sensory-physiological and morphological variables such as vision, olfaction and dexterity. Such differences are even more pronounced when comparing cognition of vertebrate taxa to invertebrate taxa such as cephalopods or bees. Essential for valid cognitive comparisons is that three basic preconditions are met. Every tested individual should possess (i) the sensory (e.g., visual or auditory) abilities to easily perceive the test apparatus and distinguish the test stimuli, (ii) the motor skills to easily handle the test apparatus and (iii) sufficient motivation to participate in and attend to the cognitive task at hand (Schubiger, 2019). The first two preconditions are related to test fairness (i.e., comparable conditions for all individuals to understand a cognitive task and perform well in it) and are arguably best met in conservative comparisons between evolutionary closely related taxa such as within primates or birds, whereas motivational and attentional aspects can probably not be equally well controlled by restricting comparisons to closely related taxa.

Reducing Sensory-Motor Influences on Cognitive Performance by Testing Primate Species

Even when conducting cognitive species comparisons within the primate order, some differences in sensory-motor skills remain that might affect individual test performance if not considered when constructing test apparatuses. For instance, many strepsirrhines (lemurs and lorises) are nocturnal and possess limited color vision (particularly along the red-green spectrum) and dexterity, while they appear to rely more on olfaction than haplorrhines (monkeys and apes), most of whom are diurnal and have excellent stereoscopic color vision. Yet different haplorrhine primate species, and even different individuals within a single species, also vary to some degree in perception and dexterity (King, 2016; Heldstab et al., 2016). Importantly, such differences have to be considered when planning and conducting cognitive tests. For instance, nocturnal strepsirrhines such as mouse lemurs were tested under infrared light to adapt testing conditions to their activity period (Kittler et al., 2018; Fichtel et al., 2020). In marmosets, most males are dichromats (red-green color blind) whereas most females have trichromatic vision owing to a cone receptor polymorphism (Pessoa et al., 2005; Freitag and Almeida Pessoa, 2012). In order to ensure these individuals were equally able to perceive the test stimuli, researchers either used yellow and blue colored test stimuli or refrained from using colored stimuli altogether by using black and white ones instead (Schmon, 2011; Strasser and Burkart, 2012; Schubiger et al., 2015, 2016, 2019). Furthermore, one central characteristic of the callitrichids (marmosets and tamarins) is that they have claws rather than typical primate hands with fingernails, which needs to be

TABLE 1 | Effect of internal (subject-related) and external (test design-related) non-cognitive factors on an individual's motivation to participate, attend, and use an appropriate response strategy, and on its cognitive performance in commonly used comparative cognition tasks.

Non-cognitive factors	Participation/attention/ response motivation	Performance	Cognitive task(s)/skill(s)	Species	Reference(s)
1. Internal (subject-related) factors					
a) Intrinsic factors					
Personality-related motivational differences					
• High levels of trait anxiety	↓	↓	Reversal learning	Long-tailed macaques	Toxopeus et al., 2005
• High openness	↑	↑	Training	Brown capuchins	Morton et al., 2013
• High assertiveness	↓	↓			
• High levels of boldness vs. shyness	↑ bold	↑ bold	Physical PCTB tasks	Chimpanzees and orangutans	Herrmann et al., 2007
	=	=	Social PCTB tasks		
	↑ bold (males)	=	Physical and social PCTB tasks	Olive baboons	Schmitt et al., 2012
	↓ shy (females)	=		Long-tailed macaques	
	=	=	Physical PCTB tasks	Ring-tailed and ruffed lemurs	Fichtel et al., 2020
	↓ Shy	↑ Shy	Social PCTB tasks	ruffed lemurs	
High emotional reactivity	↓ (males)	=	Object permanence	Common marmosets	Schubiger et al., 2015
b) Social factors					
Rearing conditions, housing conditions, and previous contact with humans					
• Impoverished rearing conditions	↓	↓	Reversal learning (Transfer Index)	Chimpanzees	Davenport et al., 1973
• Enriched rearing conditions	↑	↑	Mirror self-recognition	Chimpanzees Gorillas	Gallup et al., 1971 Patterson and Cohn, 1994; Posada and Colell, 2007
• Enriched vs. standard nursery-reared	↑	↑	Joint attention (30 BSID tasks) and cooperation (1 IBR task)	Infant chimpanzees	Bard et al., 2014
• Mother-reared vs. nursery-reared	(=)	=	13of the 16 PCTB tasks	Chimpanzees	Hopkins et al., 2014
• Social housing and high levels of human care	(↑)	↑	OTB tasks: inhibitory control, reversal learning, problem solving, causal reasoning	Sumatran and Bornean orangutans	Damerius et al., 2017b
• Degree of previous contact with humans (Human Orientation Index, HOI)	↑ High HOI	↑ High HOI	Problem-solving		Damerius et al., 2017a
Social tolerance and organization					
• High social tolerance	(↑)	↑	Cooperation	Tonkean macaques and rhesus macaques Bonobos and chimpanzees	Petit et al., 1992 Hare et al., 2007

(Continued)

TABLE 1 | Continued

Non-cognitive factors	Participation/attention/ response motivation	Performance	Cognitive task(s)/skill(s)	Species	Reference(s)
	(↑)	↑	Causality scale of the PCTB	Chimpanzees	Herrmann et al., 2010
	(↑)	↑	Theory of Mind scale of the PCTB	Bonobos	
	(↑)	↑	Pointing cups (social task 13 of PCTB), inhibitory control	Barbary macaques, long-tailed macaques, rhesus macaques, and Tonkean macaques,	Joly et al., 2017
		=	All other PCTB tasks		
• High levels of allomaternal care	(=)	=	Physical cognition tasks	Many different non-human primate species	Reviewed in Burkart and van Schaik, 2010; Burkart and van Schaik, 2016
	(↑)	↑	Social cognition tasks		
• High degree of fission-fusion	(↑)	↑	Five inhibitory control tasks	Bonobos, chimpanzees, gorillas, orangutans, long-tailed macaques, spider monkeys, and capuchin monkeys	Amici et al., 2008
	(↑)	↑	Spatial memory (similar to task 1 of the physical PCTB scale), Support (similar to task 8 from the physical PCTB scale)		Amici et al., 2010
c) Demographic factors					
Social rank	=	=	Physical and social PCTB tasks	Olive baboons and long-tailed macaques	Schmitt et al., 2012
				Ring-tailed and ruffed lemurs	Fichtel et al., 2020
Sex (gender)	(↑) males	↑ males	Physical PCTB tasks of the Space scale	Chimpanzees and orangutans	Herrmann et al., 2007
	(↑) females	↑ females	Physical PCTB tasks of the Quantity Scale	Human children	
	=	=	Physical and social tasks of the PCTB	Olive baboons and long-tailed macaques	Schmitt et al., 2012
				Chimpanzees	Hopkins et al., 2014
				Barbary macaques, long-tailed macaques, rhesus macaques, and Tonkean macaques	Joly et al., 2017
				Ring-tailed lemurs, ruffed lemurs, and mouse lemurs	Fichtel et al., 2020
	↓ males	=	Object permanence	Common marmosets	Schubiger et al., 2015
Age					
• Older subjects		↑	Causality scale of the PCTB	Bonobos and chimpanzees	Herrmann et al., 2010
		↓	Theory of Mind scale of the PCTB		
		↑	Most physical PCTB tasks	Chimpanzees (only females tested)	Lacreuse et al., 2014
		↓	Spatial memory (physical PCTB task); Attentional state and gaze following (social PCTB tasks)		
	(↓)	↓	Inhibitory control	Common marmosets	Gokcekus, 2020
	(=) curious subjects	= curious subjects			

(Continued)

TABLE 1 | Continued

Non-cognitive factors	Participation/attention/ response motivation	Performance	Cognitive task(s)/skill(s)	Species	Reference(s)
2. External (test design-related) factors					
a) Task format					
Physical cognition tasks					
• Possibility to rake-in food reward (rather than just push) with tool	↑	↑	Trap-tube	Bonobos, chimpanzees, gorillas, and orangutans	Mulcahy and Call, 2006
• Food retrievable with finger rather than tool		↑	Two-trap box	Chimpanzees	Seed et al., 2009
• Inedible test stimuli (tokens) • Food type difference between test stimuli and rewards	↑	↑	Quantity discrimination	Olive baboons and long-tailed macaques	Schmitt and Fischer, 2011
	=	=		Brown capuchins	Gazes et al., 2018
• Edible test stimuli	↑	↑		Brown capuchins	Addessi et al., 2008; Gazes et al., 2018
• High quality rewards	↑	↑		Brown capuchins and squirrel monkeys	Gazes et al., 2018
• Low probability of success by chance (large number of test stimuli)	(↑)	↑	Memory (modified after TTB)	Common marmosets, common squirrel monkeys	Schubiger et al., 2016
	(↑)	↑	Uncertainty monitoring (computerized metacognition task)	Rhesus monkeys and capuchin monkeys	Beran et al., 2014
				Capuchin monkeys	Beran et al., 2016
• Multimodal exploration of test stimuli	↑	↑	Visual discrimination	Capuchin monkeys	Carducci et al., 2018
Social cognition tasks					
• Competitive rather than cooperative experimenter cues	↑	↑	Object-choice	Chimpanzees	Hare and Tomasello, 2004
• Experimenter's cue already in place when subject enters test area	(↑)	↑			Barth et al., 2005
				Common marmosets	Burkart and Heschl, 2006
• Low probability of random success	(↑)	↑			
• Eye contact at time of experimenter's communicative cue	(↑)	↑		Bonobos, chimpanzees, and orangutans	Mulcahy and Call, 2009; Mulcahy and Suddendorf, 2011
• Larger distance between test stimuli	↑	↑			
	↑	↑	Perspective taking	Chimpanzees	Hare et al., 2000; Karin-D'Arcy and Povinelli, 2002
b) Opportunistic testing					
Excluding subjects with motivational issues					
• Excluding subjects who need more testing time	↑	=	Inhibitory control (A-not-B and detour-reaching, adapted from the TTB) and memory	Common marmosets and squirrel monkeys	Schubiger et al., 2019
			Quantity discrimination and reversal learning (adapted from the TTB)	Common marmosets	

No effect: =, negative effect: ↓, positive effect: ↑. Note: Symbols in parentheses () indicate that participation/attention and response motivation were only indirectly assessed.

taken into account when designing tasks in which subjects need to manipulate objects (as e.g., Schubiger et al., 2016; Schubiger et al., 2019).

While test designers and experimenters can largely control sensory-motor confounding factors by using appropriate test apparatuses, they may have limited control over several other non-cognitive factors during testing. For instance, motivational aspects remain a potential source of bias on cognitive performance. Examples are inner states and predispositions that affect how individuals approach and attend to their non-social (including cognitive tasks) and social environment (including the human experimenter and the cognitive test situation). However, deliberate test designs and analytical methods might alleviate some of these issues (e.g., Schubiger et al., 2015).

Starting from primate cognition studies, we here review recent studies that exemplary address (1) how individual differences in several non-cognitive factors affect participation and performance in cognitive tasks commonly used for within and between-species comparisons and (2) how aspects of test design and human-induced biases directly or indirectly affect cognitive performance. Finally, we evaluate (3) how recent studies that used cognitive test batteries may be affected by such effects.

OVERVIEW OF STUDIES ON NON-COGNITIVE FACTORS THAT POTENTIALLY CONFOUND COGNITIVE PERFORMANCE

Individuals can differ considerably in terms of non-cognitive factors (i.e., intrinsic ones such as individual differences in personality, emotion and motivation) and different species differ in social factors (such as levels of social tolerance or social organization or structure) that might affect their cognitive task performance. Recent comparative cognition studies have started to assess a number of such internal (subject-related) and external (test design-related) non-cognitive factors and their potential effects on cognitive performance. Here, we review several relevant studies and their findings regarding whether they affected cognitive task performance or not (see also **Table 1**).

Internal (Subject-Related) Factors

Intrinsic Factors

Personality-related motivational differences

Personality. At least some personality traits have been shown to affect cognitive performance in non-human primates. For instance, trait anxiety as assessed by the monkeys' sustained reaction to a loud noise, was negatively correlated with the cognitive performance of long-tailed macaques in a reversal learning task (Toxopeus et al., 2005). Personality traits of non-human primates have also been assessed in a comparable manner to humans using the Hominoid Personality Questionnaire (HPQ; King and Figuereo, 1997; Weiss, 2017). Traits such as openness (behaviorally associated with the time an individual devotes to playing with conspecifics) and assertiveness (behaviorally associated with an individual's aggressive behavior toward conspecifics) were associated with the participation and performance of capuchin monkeys *(Sapajus apella)* in two training tasks that preceded cognitive testing. Subjects with more open or less assertive personalities and particularly those exhibiting a combination of both trait expressions were more motivated to participate and also performed better in the training tasks than less open and highly assertive subjects (Morton et al., 2013).

Temperament and neophobia. In non-human primates, temperament or neophobia is generally measured as the latency to approach novel objects, humans, or food. Bolder chimpanzees *(Pan troglodytes)* and orang-utans *(Pongo pygmaeus)* approached such novel situations more quickly and performed better in some physical subtests of the PCTB than their shyer conspecifics. In the social subtests, however, individual differences in temperament were not associated with the apes' cognitive performance in either domain (Herrmann et al., 2007). The opposite pattern was found for ruffed lemurs in that shyer subjects who took longer to approach and spent less time in the testing area performed better in the social PCTB tasks than their bolder conspecifics. This was not the case for ring-tailed lemurs whose temperament did not correlate with performance in the physical or social tasks of the PCTB (Fichtel et al., 2020). Olive baboons, particularly males, spent more time next to new objects than long-tailed macaques and showed a shorter approach latency toward new stimuli than long-tailed macaques, particularly females. Their performance in the PCTB, however, was not associated with these species and sex differences in temperament (Schmitt et al., 2012). Importantly thus, species can interact with influences of personality, temperament and neophobia in predicting cognitive outcomes.

Emotional reactivity. While an individual's temperament and personality traits are fairly stable over time, its emotional reactivity may differ depending on the context and be particularly strongly expressed in the test situation. Common marmoset (*Callithrix jacchus*) subjects who showed a strong spontaneous emotional reaction to the experimenter and the test situation participated in fewer trials of an object permanence task than their less emotionally reactive conspecifics (Schubiger et al., 2015). Elevated emotional arousal, which the marmosets visibly and auditorily expressed (via piloerection of the tail as well as increased vigilance, mobbing vocalizations and avoidance behavior), was particularly apparent in the majority of male individuals and affected their attention in trials in which they participated. However, when strict pre-defined stop criteria were applied to abort a test session when a subject's state of elevated emotional arousal persisted, their cognitive performance was not affected.

Social Factors

Rearing conditions, housing conditions, and previous contact with humans

Individual differences in rearing conditions, housing conditions and previous contact with humans may affect cognitive performance in primates, which is of particular importance when testing and comparing primates from more heterogenous populations such as in different zoos, sanctuaries and in the wild. One example is that a young individual who has abundant opportunity to learn socially from its adult conspecifics (ideally by being mother-raised), is able to acquire a larger set of cognitive skills than an individual who is deprived of this opportunity (such as an orphan growing up with peers; reviewed in van Schaik and Burkart, 2011).

Adult chimpanzees, for instance, who had experienced *impoverished rearing conditions* in the first two years of their lives, performed poorly in a reversal learning task (as reflected in their lower transfer indices, a measure of cognitive flexibility) than their mother-raised conspecifics (Davenport et al., 1973), even though all chimpanzees had been living at the same facility during the last six years prior to cognitive testing.

Enriched rearing conditions, on the other hand, favored whether chimpanzees (Gallup et al., 1971) and gorillas (Patterson and Cohn, 1994; Posada and Colell, 2007) showed evidence of mirror self-recognition. Moreover, infant chimpanzees growing up in enriched nursery-care conditions developed better socio-cognitive abilities than their conspecifics raised in standard nursery care, particularly those abilities related to joint attention (as measured by 30 tasks of the Bayley Scales of Infant Development for human infants: BSID, Bayley, 1969) and cooperation (as rated by the experimenter using the Infant Behavior Record: IBR, Bayley, 1969; Bard et al., 2014). A more recent study with zoo-housed chimpanzees, however, found that being *mother-reared vs. nursery-reared* was not associated with how well adult chimpanzees performed in the subset of 13 physical and social PCTB tasks they were tested with (Hopkins et al., 2014).

Being *housed with conspecifics* in zoos and sanctuaries and *being cared for by humans* allows individuals to be more curious and explorative toward their surroundings than their single-housed conspecifics. Such favorable rearing and housing conditions, most likely in combination with higher exposure to human artifacts, facilitated performance in the Orangutan Test Battery (OTB, Damerius et al., 2017b).

Individual differences in orangutans' *previous contact with humans* have recently been quantified by a new composite measure that assesses individual differences in the subjects' behavioral response to unfamiliar humans, the Human Orientation Index (HOI, Damerius et al., 2017a). Individuals who had been more exposed to humans exhibited higher HOI-scores than those with limited human exposure, were more explorative and also more successful than less human-oriented orangutans at solving the honey-box task of the OTB, in which they had to use tools to extract honey from a wooden apparatus (also see Damerius et al., 2019).

Social organization and social tolerance

Primate species differ regarding their social organization such as the spatiotemporal cohesion of the societies they live in (Kappeler and van Schaik, 2002). Fission-fusion societies, for instance, are characterized by dynamic group compositions with changing associations between individuals both in time and in strength (Dunbar, 1988). In haplorrhine primates more generally, ape and monkey species living in social systems characterized by a *high degree of fission-fusion dynamics*, such as great apes and spider monkeys, performed better in two physical cognition tasks similar to the ones of the PCTB ("spatial memory" and "shape") and in several inhibitory control tasks from other sources (Amici et al., 2008, 2010).

Primate species also differ considerably in terms of *social tolerance levels*. Tolerant primate societies are characterized by less steep dominance hierarchies, low levels of conflicts without clear directionality, and feeding in close proximity (Jaeggi et al., 2010a; Fichtel et al., 2018). All these factors can facilitate highly social behavior such as cooperating in solving problems and prosocial acts such as proactively sharing food with conspecifics (Jaeggi and van Schaik, 2011; Burkart et al., 2014). For instance, in cooperative tasks (that required two individuals to simultaneously lift heavy stones to uncover food items or pull-in a feeding platform they would both be able to access) more tolerant macaque and ape species performed better than less tolerant ones. In particular, Tonkean macaques performed better than rhesus macaques (Petit et al., 1992) and bonobos better than chimpanzees (Hare et al., 2007). Likewise, socially more tolerant macaque species performed better than less socially tolerant macaque species in an inhibitory control task and in one PCTB task from the social scale (pointing cups) that tests the subject's ability to communicate to the experimenter in which location food has earlier been placed by a second person, but not in any other PCTB tasks (Joly et al., 2017). Finally, the cooperative callitrichid monkeys (marmosets and tamarins with higher levels of *allomaternal care*) exhibit generally high levels of social tolerance, which appears to facilitate performance in socio-cognitive tasks such as social learning or cooperative problem solving compared to their less tolerant sister taxa (Burkart and van Schaik, 2010, 2016). Importantly, at least in primates, social tolerance can also differ considerably between different groups of the same species (Jaeggi et al., 2010b; Cronin et al., 2014; Burkart, 2015) and thus mask or exaggerate potential species differences in cognitive performance.

Demographic Factors

The potential effects of individual differences in demographic factors on cognitive performance are well known and often taken into account by comparing balanced, unbiased samples to avoid confounding effects or statistically controlled in studies using cognitive test batteries.

Social rank

An individual's rank in its social group is typically inferred via the number of decided conflicts between individuals gathered via focal observations. An individual's social rank did not affect its cognitive performance in any of the PTCB tasks in olive baboons, long-tailed macaques (Schmitt et al., 2012) or ring-tailed and ruffed lemurs (Fichtel et al., 2020).

Sex

Male chimpanzees in Herrmann et al.'s (2007) PCTB study performed better than females in the space scale and male children performed better than female children in the quantities scale. However, in Hopkins et al.'s (2014) chimpanzee study, a subject's sex did not affect its performance in the PCTB and neither did it in baboons and long-tailed macaques, although male and female individuals differed in terms of temperament in these two Old World primate species (Schmitt et al., 2012). Similarly, sex did not influence the performance of three lemur species in the PCTB (Fichtel et al., 2020). In marmosets, males were generally more easily emotionally aroused in the test situation than females and less food motivated to participate in the cognitive tasks (Schubiger et al., 2015). In addition, males were often more attentive to their surroundings than the test apparatus in front of them, which is in line with males showing more vigilance behavior in the wild (Koenig, 1998). Importantly, however, because they were given the opportunity to leave the test situation as soon as they became unmotivated and inattentive, and because inattentive trials were excluded from the analysis, they performed comparably to their female conspecifics (Schubiger et al., 2015).

Age

Generally, as in humans, cognitive abilities recruiting fast and flexible (fluid) mental processes and maintaining information such as executive functions (e.g., inhibitory control) are also predicted to decrease with increasing age in non-human primates (e.g., Deary et al., 2009). Cognitive abilities that improve with experience, on the other hand, such as many social ones, are predicted to increase over an individual's lifetime (for a review see Burkart et al., 2017a). Interestingly, the opposite pattern was found in chimpanzees (Lacreuse et al., 2014) in that age had a positive effect on individual performance in the physical cognition tasks of the PCTB (with the exception of the spatial memory task) but a negative effect on their performance in two of the socio-cognitive tasks (attentional state, gaze following). Similarly, older chimpanzees and bonobos performed better in some of the physical tasks (causality scale) but not as well as younger individuals in some social tasks (Theory of Mind scale, Herrmann et al., 2010). In the other studies applying the PCTB, no such age-effects on cognitive performance were reported (Schmitt and Fischer, 2011; Fichtel et al., 2020). Another interesting finding was that high levels of curiosity appeared to alleviate cognitive decline in marmosets (Gokcekus, 2020). While marmosets' ability to inhibit directly reaching into a transparent barrier (detour-reaching tasks) declined with age in individuals with low curiosity scores, this ability remained stable in particularly curious individuals and in some cases even increased with age. Hence, the influence of age on performance in cognitive tests is not straightforward.

External (Test Design-Related) Effects on Performance

Task Format

Task format, i.e., the way in which a cognitive task is designed and how (many) test stimuli are presented to the subject, not only generally affected cognitive performance in several empirical comparative cognition studies but it sometimes did so in different ways in different species.

Physical cognition tasks

One physical task that has been used to test the causal understanding of non-human primates is the *trap-tube task*, originally designed by Visalberghi and Limongelli (1994) testing capuchin monkeys. In the original trap-tube task, a food reward was placed in a transparent tube that contained a trap in its middle. In order to retrieve the reward, the subject had to use a stick tool and push the reward out of the tube while preventing it from falling into the trap. To ensure that chimpanzees were not using simple distance rules rather than causal reasoning, the trap was later moved to the side and the reward placed in the tube's middle instead (Limongelli et al., 1995), but this could not rule out other simple rules such as always pushing the food away from the trap, which chimpanzees tended to do even when the trap was inverted and non-functional (Povinelli, 2000). The strongest evidence for great apes' causal understanding of the trap-tube problem comes from studies using modified task versions. For instance, great apes performed better in a modified trap-tube task, in which the tube was wider so that the apes could also use the stick tool to *rake-in the reward* (thereby pulling it away from the trap) rather than having to push the reward away from the trap (Mulcahy and Call, 2006). This suggests that improving the ecological validity of the task, which allowed the apes to use the tool in a more natural way, helped reveal their causal understanding. In a further modified version, the *Two Trap Box task*, the reward was placed on a shelf that had a trap on each end (one of which was blocked depending on the trial). Through the transparent front of the box, the subject had visual access to the test apparatus and could use its finger to move the reward away from the trap *without requiring a tool*, which made the task easier for chimpanzees (Seed et al., 2009). A similar task version consisting of a box with six channels each containing a trap was used in the OTB (Damerius et al., 2019), and a considerable number of individuals succeeded, suggesting that the causal problem was easier to solve when they could use their fingers instead of a tool.

In a *quantity discrimination task* (requiring subjects to choose the larger of two amounts of items), two Old World monkey species, olive baboons (*Papio anubis*) and long-tailed macaques (*Macaca fascicularis*) performed better when the test stimuli were *inedible* (i.e., tokens) as opposed to edible items (i.e., raisins). Interestingly, the same monkeys performed equally well with edible test stimuli when the food type of the edible items they were rewarded with differed from the *food type* of the test stimuli (Schmitt and Fischer, 2011). Contrarily, capuchins monkeys (*Cebus sapajus apella*) performed better with *edible* stimuli than tokens (Addessi et al., 2008; independent of food types: Gazes et al., 2018) and generally better than squirrel monkeys (*Saimiri sciureus*). When the quality of food rewards was modulated, both New World monkey species performed better with *higher quality rewards* independent of how long they had to wait to be rewarded (Gazes et al., 2018). This heterogeneous influence most likely emerged because highly attractive rewards on the one hand increase an individual's motivation, but on the other hand elicit prepotent reactions and thus can increase demands on inhibitory control. Depending on which influence prevails, high quality rewards can both increase and decrease task performance.

Task format can also include how stimuli are presented, e.g., only visually or visually and haptically. This turned out to influence performance in a *visual discrimination task*, in capuchin monkeys (*Sapajus* spp.). They were more successful at distinguishing between two objects when they had access to *haptic* in addition to purely visual information by being allowed to touch and manipulate the objects suggesting they benefited from this multimodal exploration (Carducci et al., 2018).

Finally, task formats can vary with regard to *chance probabilities of success*. Modifications of these probabilities in physical cognition tasks revealed that common marmosets and common squirrel monkeys performed better in a *memory task* in which they had a choice between nine containers, only one of which was baited with a food item, than in the original two-choice version of the task. Lowering the chance probability of success from 50 to 11% made wrong choices in the nine-choice memory task more costly and is likely to have indirectly enhanced the monkeys' motivation to favor an appropriate rather than a random choice strategy and thus more reliably assessed their memory decline over increasingly longer time delays (Schubiger et al., 2016). Similarly, capuchin monkeys showed better evidence of uncertainty monitoring in a computerized *metacognition task* by more often selecting the escape option when chance levels of success were lower than when they were higher, whereas rhesus macaques appeared less sensitive to higher costs of incorrect choices (Beran et al., 2014, 2016).

Social cognition tasks

One of the most extensively used social cognition test paradigms in comparative psychology is the *object choice task* (originally developed by Anderson et al., 1995) in which the experimenter sits or stands opposite the subject, hides a food item in one of at least two containers, and then provides the subject with at least one (visual and/or auditory) communicative cue to indicate the food's location before the subject is allowed to choose one of the containers. Primates, especially great apes, have been shown to perform poorly in comparison to many other animal species, including distantly related mammals such as canids as well as birds (e.g., Bräuer et al., 2006; Giret et al., 2009). However, although this has been interpreted as the apes' inability to understand human-given communicative cues, the original test setup used for primates differed from the one used for other animal species and several modifications to the primate version substantially improved the apes' performance (for a detailed review see Mulcahy and Hedge, 2012), the most relevant of which we list here.

Competitive experimenter cues. In one object-choice study, chimpanzees performed better in an object-choice task if the experimenter's cue was competitive in that he extended his arm in an attempt to grab the baited container rather than pointing at it in a cooperative manner (Hare and Tomasello, 2004). However, as Mulcahy and Hedge (2012) pointed out, the competitive task version also differed from the cooperative one in that the former included a potential inhibition component and higher costs of an incorrect choice. Before being able to choose one of the two containers, the apes had to open a corresponding transparent panel in the testing window. Having to do so might not only have prevented them from making ambiguous choices but also enhanced their motivation to attend to the experimenter's cue(s). This in turn might have helped them to perform better in the competitive task version.

Experimenter's cue already in place when subject enters the test area. In Barth et al.'s (2005) object choice study, chimpanzees performed poorly in the original version of the task in which the experimenter provided the communicative cue (head and eyes directed toward the baited container) only once the subject was directly in front of the experimenter. In contrast, the same chimpanzees located the reward much more successfully in a modified task version, in which the experimenter initialized the gaze cue before the subject entered the test area. This indicates that when entering the test area, the chimpanzees immediately looked into the direction of the experimenter's cue and as they were approaching, they veered in this direction which would result in them arriving at and choosing the correct container. Similarly, in a marmoset study, the subjects' access to choosing containers was restricted until the experimenter had provided the pointing gesture toward the baited container which might have facilitated basic inhibitory and attentional processes required to make correct choices (Burkart and Heschl, 2006).

Lowering the chance probability of random success. Burkart and Heschl's (2006) task version also differed from other object-choice studies in that marmosets were presented with nine containers, only one of which contained a reward. This version was directly compared to a two container version, and the marmosets performed much better in the first one, which probably enhanced their motivation to attend to the experimenter's cues because incorrect choices involved a higher cost than in the traditional object choice task with two choice options. Together with similar findings on chance probabilities from two physical cognition tasks (i.e., memory and metacognition) mentioned earlier, this suggests that lowering the chance probability of making correct choices at random by increasing the number of available choice options may positively affect performance in tasks across cognitive domains.

Increasing the distance between test stimuli. Other modifications to the original object choice task, such as increasing the distance between the containers in which the food item is hidden, have also been shown to positively affect the performance of bonobos, chimpanzees, and orangutans in object choice tasks (e.g., Mulcahy and Call, 2009). The distance between test stimuli also turned out to be responsible for discrepant results in *perspective-taking tasks* that assess whether the subject knows if a conspecific individual present in the test situation can see a reward (i.e., because it is visible from both the subject's and the conspecific's point of view) or not (i.e., because a barrier obstructs the conspecific's view). Initial findings by Hare et al. (2000) suggesting that chimpanzees knew what their conspecific could see could not be readily replicated (Karin-D'Arcy and Povinelli, 2002), but it turned out that this was owing to variation in spatial factors of the set-up. In the meantime, this paradigm has been applied to a variety of species with varying results, and it is not entirely clear which differences represent true species differences and which ones may be affected by spatial factors too (e.g., capuchin monkeys: Hare et al., 2003 vs. common marmosets: Burkart and Heschl, 2007).

Establishing eye contact when giving the cue. Although it is effortful to establish eye contact with some non-human primate subjects and this is not possible with all species (e.g., owing to gaze aversion or being perceived as a threatening gesture), ensuring in this way that the subject is attentive to the experimenter's cues has been shown to improve the performance of bonobos, chimpanzees, and orangutans in the object choice task (Mulcahy and Call, 2009; Mulcahy and Suddendorf, 2011).

Opportunistic Testing

Excluding Subjects With Motivational Issues

One specific issue of comparative psychology is that not all species and not all individuals within a given species are equally motivated to continually participate in cognitive tasks. An individual's lack of motivation to do so can critically affect the course of a study because the individual will require substantially more than the allocated or available testing time to complete the cognitive tasks. This can be particularly problematic in cognitive studies with non-human primates because access to respective testing facilities is often temporally limited, which constrains the time available for a study. Researchers often deal with this constraint by following an opportunistic approach of only testing individuals who readily participate and are most likely to complete the tasks in the available testing time and excluding those who are not. However, such opportunistic testing might bias a study's outcome if the excluded individuals not only differ from the selected ones in motivational factors but also in terms of cognitive abilities.

In a recent study with common marmosets and common squirrel monkeys, this issue has been addressed by including less consistently motivated individuals and allowing them additional time to complete the six tasks of a cognitive test battery at their own pace (Schubiger et al., 2019). A direct comparison of individuals who needed additional *testing time* to those who were consistently motivated showed that both groups performed equally well in all tasks. This suggests that opportunistic testing and the selection bias that results from it does not necessarily affect a study's outcome. Whether this also applies to other species still needs to be established.

HOW ARE INDIVIDUAL AND SPECIES COMPARISONS USING COGNITIVE TEST BATTERIES AFFECTED BY NON-COGNITIVE FACTORS?

When using test batteries to assess individual differences within a species, priority is given to the same individuals completing all tasks. This is because obtaining a complete data set enables researchers to conduct factor-analytical performance analysis whereas dropouts would complicate this approach. A second goal is to obtain large enough sample sizes to reach sufficient statistical power. As current data suggest, using an opportunistic (as opposed to a randomized) approach by only selecting those individuals as subjects who are most likely to stay consistently motivated and complete all tasks is not a major issue. At least not as long as researchers report that and why this approach was followed and as long as dropouts and their performance in the few tasks they completed are also reported in detail.

However, some individuals (or species) might also require more time to get used to a new task because they are neophobic and more cautious when approaching the test apparatus for the first time. It is therefore advisable to allow every subject to familiarize itself with the basic test apparatus and to only start testing when the subject appears comfortable with all components. Highly neophilic individuals, on the other hand, tend to approach and get used to the test apparatus much more quickly with the risk that some of these individuals might also more quickly lose interest once the task is not novel anymore. In order to enable later replications of a study, it is therefore important to describe in detail how subjects were familiarized with the tasks prior to testing, how their motivation was regained if necessary, and which criteria were used to objectively decide when a test session started and when it had to be aborted.

Besides opportunistic testing, using several experimenters rather than just one is another way to test as many subjects as possible with all tasks in a limited testing period. Although training different experimenters to use the same standardized methods helps reducing experimenter effects that might bias the subjects' cognitive performance, a certain risk of such unintended biases remains. Herrmann et al. (2007), for instance, used five different experimenters in the original PCTB study, including two experimenters with the rule that every subject was tested by the same experimenter with all tasks. Since one group of chimpanzees performed better than the other, Herrmann and her colleagues could not tease apart in how far these differences were purely cognitive in nature or also experimenter-induced.

Another issue is that not all subjects are equally motivated to participate in food-reward tasks and different food types do not have the same value for all individuals. One way of limiting such individual differences is to use tokens as test stimuli rather than food items (e.g., Addessi et al., 2008; Schmitt and Fischer, 2011). However, subjects need to be trained to use tokens, which limits the usefulness of this approach, particularly for large-scale species comparisons because test battery tasks should not require any previous experience. Quantity discrimination tasks in which the subject has to choose the larger of two quantities (of food or token items) to pass a trial have shown to be particularly susceptible to the type of test stimuli and rewards (Schmitt and Fischer, 2011; Gazes et al., 2018). Regardless of whether subjects chooses between two amounts of tokens or food items, the number of rewards usually corresponds to the chosen amount of test stimuli. However, this procedure differs from the one used in all other test battery tasks in which the subject usually only receives one reward in case of a correct choice, which could be one possible explanation why the task appeared to be difficult for squirrel monkeys in Gazes et al.'s (2018) study. Therefore, in the quantity discrimination task of their adapted test battery, Schubiger et al. (2019) used two amounts of edible "tokens" of low food quality as test stimuli, which made them interesting enough to attend to the task (and more interesting than non-food tokens) but not desirable as rewards. If the subject correctly chose the larger amount, it was, as in all other tasks, rewarded with one highly desirable food item. They found that the dropout rate in the quantity discrimination task was particularly low in comparison to most other tasks and the marmosets performed better in this task than in most others. Whether this was a consequence of the setup and reward contingency remains to be determined in future studies.

The most comprehensive test battery currently available for non-human primates is the PCTB consisting of a physical and social cognition scale that each comprise several subtests amounting to a total of 16 cognitive tasks (Herrmann et al., 2007). Initially, the PCTB was applied to the largest sample of great apes (chimpanzees and orangutans) that had ever been tested in comparative psychology and to 2.5 years-old human children who outperformed both ape species in most social but not in the physical cognition tasks. In the last decade, the full PCTB or parts of it (ranging from six to 13 tasks) have been used to assess and compare the cognitive abilities of ten other non-human primate species (with some minor adaptations). Besides an independent chimpanzee sample (tested with 13 tasks: Hopkins et al., 2014) the tested species included bonobos *(Pan paniscus,* Herrmann et al., 2010) one small ape species (lar gibbons, *Hylobathes lar*, *Yocom, 2010,* tested with six tasks), five Old World monkey species (olive baboons*,* *Papio anubis*, Schmitt et al., 2012; longtailed-macaques, *Macaca fascicularis*, Schmitt et al., 2012; Joly et al., 2017; Barbary macaques, *Macaca sylvanus*, rhesus macaques, *Macaca mulatta*, and Tonkean macaques, *Macaca tonkeana*, Joly et al., 2017) and three lemur species (tested with all tasks; black-and-white ruffed lemurs, *Varecia variegata*, ring-tailed lemurs, *Lemur catta*, and mouse lemurs, *Microcebus murinus*, Fichtel et al., 2020). In addition, four bird species (parrots) have recently also been tested with the full PCTB (African grey parrots, *Psittacus erithacus,* blue-headed macaws, *Primolius couloni*, blue-throated macaws, *Ara glaucogularis,* and great green macaws, *Ara ambiguous*, Krasheninnikova et al., 2019). Unanticipatedly and in contrast to previous meta-analytic studies (Deaner et al., 2006; Reader et al., 2011), the primate studies found that overall Old World monkeys and lemurs (who as strepsirrhines represent the evolutionarily most distant primate relatives of great apes) performed largely comparable to great apes, particularly in the social scale. Contrarily, all four parrot species performed inferior to great apes in both the physical and social scale of the PCTB (Krasheninnikova et al., 2019). This was unanticipated because parrots (besides corvids and owls) belong to the birds with the largest brain size and parts of their brains have been described as homologous to the mammalian neocortex (Jarvis et al., 2005; Güntürkün and Bugnyar, 2016). Based on their powerful brains and their remarkable cognitive abilities that have been demonstrated in several tasks and sometimes been considered to match or even exceed those of non-human primates (e.g., Pepperberg, 2006), the parrots were expected to perform relatively well in the PCTB.

Three explanations appear most plausible for this arguably unexpected pattern of results: (i) the tested species do not differ in terms of cognitive abilities (which would be in line with Macphail's null hypothesis), (ii) small differences in task designs rather than cognitive ability masked species differences in cognitive performance, or (iii) the levels of task sensitivity were not appropriate to identify between-species variation and instead led to ceiling (i.e., mainly very high performance scores) or floor effects (i.e., mainly very low performance scores). While the first possibility appears unlikely to explain the primate and parrot findings based on what is known about their cognitive abilities, the two other two possibilities or a combination of the two (depending on the tasks and species) appear more plausible.

The lack of clear-cut performance differences between the different primate taxa points to ceiling effects in most tasks of the PCTB with relatively good performance levels in all species most likely owing to the relatively low task sensitivity (Fichtel et al., 2020). Moreover, all PCTB studies with monkeys and lemurs also found floor effects for at least one physical (tool use) and one social cognition task (social learning). In fact, only great apes passed the tool use task that required the ability to use a stick tool to rake a food reward into reach whereas no other primate species did. Doing so might have been too challenging for species exhibiting either a medium (baboons, macaques) or low (lemurs) level of precision grip (Torigoe, 1985; Kittler et al., 2018). This is not surprising because, even in captivity, great apes use stick tools more often and more skillfully than other primate species.

Comparably to the primates, all tested macaw species also performed relatively poorly in the tool-use task despite other parrot species (such as Goffin's cockatoos, *Cacatua goffiniana*) having been shown to be skillful at using a stick to retrieve food in previous experiments. According to Krasheninnikova et al. (2019), this indicates that morphological rather than cognitive constraints such as their longer maxilla and a less muscular tongue made it difficult for the macaws to maneuver the stick and pull the food reward into reach.

The social learning task of the PCTB, for which another floor effect was found, required subjects to solve a problem using the same solution that a human experimenter had demonstrated, i.e., retrieving a food item out of a transparent or opaque tube using the same behavioral actions. It is not surprising that human children performed better than great apes in this task because children have been shown to over-imitate actions of adults by even copying unnecessary or unsuccessful steps or methods of a human demonstrator whereas chimpanzees did not (Horner and Whiten, 2005). In addition, however, children could learn from a conspecific demonstrator whereas all other species had to learn from a hetero-specific demonstrator, the human experimenter. Among non-human primates, great apes possess the most similar preconditions to children in that their hands and manipulation skills resemble those of humans the most (Heldstab et al., 2016). Consequently, a social learning task in which subjects could learn from a conspecific demonstrator and that is adapted to the manipulative skills of monkeys and lemurs might have been more informative (Fichtel et al., 2020).

A striking result was, however, that the parrots performed at chance-levels in most of the tasks of the PCTB, indicating that non-cognitive factors as well as aspects of task design may have played a role. Particularly in the space scale that largely consisted of object permanence tasks, all primates outperformed the macaws despite parrots having been shown to pass such invisible displacement tasks in previous studies, and even before reaching adulthood. As Krasheninnikova et al. (2019) suggested, having to choose containers by touching them with their beaks might have made it more demanding for the parrots to inhibit prepotent impulses to touch containers. Based on earlier findings on parrots' numerical cognitive abilities, the African grey parrots and macaws would also have been expected to perform much better in the quantity tasks. The authors' finding that many individuals seemed to choose in a random manner in many PCTB tasks, particularly in those that involved only two choice options, is in line with earlier findings on primates that subjects may not always be motivated to attend to the task and use an appropriate choice strategy when they have a 50% probability of making a correct choice by chance and being rewarded (e.g., Burkart and Heschl, 2007; Schubiger et al., 2016, 2019; Fichtel et al., 2020).

It is important to mention that while great apes only received one to three trials per task of the PCTB, other primates (Old World monkeys and lemurs) and parrots received up to six trials per subtest. In the object-permanence tasks of the physical scale, for instance, the monkeys and lemurs received six trials so that all spatial positions and combinations of the baited cups were evenly distributed. In principal, participating in the double amount of trials might have given them an advantage over the great apes in that they had the opportunity to learn and perform better across trials. However, the monkey and lemur results were stable and there were no learning effects from the first to the second half of trials, which makes it unlikely that they had a substantial learning advantage over the great apes (Schmitt and Fischer, 2011; Fichtel et al., 2020).

Therefore, the most plausible explanation for the apparent lack of cognitive differences between various primate species based on individual performance scores in the PCTB, is that some of the tasks are not valid or sensitive enough to reveal differences between species (Fichtel et al., 2020). For instance, all primate species from lemurs to great apes performed very well in the spatial memory task of the PCTB's space subscale, which was basically an object permanence task for most species as it involved keeping track of two food items placed in two of three cups without a delay between the baiting of the cup and the subject's choice. Only the transposition task, in which keeping track of the food item became more demanding, revealed some species differences, which indicates that this task's level of difficulty was appropriate to distinguish between species while the other tasks of the space subscale led to ceiling effects.

The floor effects in the primates' and parrots' performance demonstrate that researchers are facing a trade-off when constructing cognitive test batteries. While the test set-ups need to be sufficiently abstract to also identify a putative domain-general cognitive ability rather than only capturing narrow domain-specific adaptations, the task apparatuses also need to be ecologically valid enough to be easily perceived and handled by all individuals. While closely related species can be largely tested with the same test apparatuses and setups with only minor adjustments, large-scale species comparisons might require more changes. This might particularly be the case for species with low dexterity levels or those who have to use their beaks (birds) or noses (e.g., canids and elephants) to handle the test apparatus and make responses. As the parrot results suggest, validated tasks should be used that can be adjusted for as many species as possible while keeping the cognitive task itself as similar as possible. This is an immense challenge for comparative psychology that yet needs to be accomplished.

One task that may be promising for meaningful species comparisons is the reversal learning task (Rumbaugh, 1971) in which all individuals (of every species) first have to master an initial discrimination by reaching the same predefined criterion. In the actual test, the discrimination is then reversed and it is assessed how quickly the subjects switch to the new discrimination in relation to their pre-reversal performance (Transfer Index). Despite their timeliness, reversal learning tasks have rarely been part of cognitive test batteries. Moreover, the pivotal measure, the Transfer index, has only been determined in one study that used a modified version of the TTB task to test marmosets (Schubiger et al., 2019). The latter task version was optimized in that two differently patterned black-and-white plates were used as test stimuli under which the reward could be hidden, rather than presenting and hiding the reward in the experimenter's hands (two different colors of gloves, including a green one) as in the original TTB. This optimization was applied in order to minimize potential effects of individual differences in the subjects' color-vision as well as potential experimenter effects. Including reversal learning tasks in future test batteries will allow to compare the cognitive flexibility of different individuals and species regardless of how long they needed to reach the initial criterion. Doing so would also help to establish whether new experimental data support findings from metanalytical research in which reversal learning performance was the best predictor of general intelligence across species (Deaner et al., 2006).

Most importantly, however, our findings illustrate the importance of conducting basic validations of cognitive tasks and test batteries in comparative psychology before applying them to a broader range of different species. This can be achieved by establishing that the cognitive tasks truly measure differences and similarities in aspects of cognition rather than other aspects that are not primarily cognitive in nature. One possible way of doing so is to first establish that each cognitive task (of a future test battery) identifies intra- and inter-individual cognitive variation in one species by assessing which potentially confounding factors can be ruled out or have to be controlled as far as possible. In a second step this species could then be compared to its closest evolutionary relatives and then to more distant ones. Another way is to use one or a small number of established and validated cognitive tasks to compare the cognitive abilities of a wide range of different species (e.g., as in the Many Primates project, ManyPrimates et al., 2019a,b). Importantly, in the latter type of studies, relatively overt differences in sensory-motor aptitudes need to be considered, particularly when comparing evolutionarily distantly related taxa (e.g., such as primate to bird species). While minor adaptations to the basic test setup and apparatus might be sufficient in some tasks, other tasks might require more substantial modifications to be suitable for a wide range of species. This is a challenge, which might demand (repeated) re-designing of tasks that turn out to be unsuitable but a challenge worth pursuing in order to establish that cognitive tests truly capture cognitive differences and similarities between individuals and species. Optimized test batteries consisting of tasks that are largely controlled for these factors will more accurately measure if and how species compare or differ in terms of cognitive abilities without non-cognitive factors playing a substantial role. This will provide a more solid base for meaningful inferences and conclusions regarding how these similarities or differences may have evolved.

CONCLUSION

In this review, we provided an overview of recent studies that assessed (subject- and test design-related) non-cognitive factors that may confound the outcomes of primate cognition tasks in general and primate test batteries in particular. In order to take into account sensory-motor species-differences, we have largely focused on studies with primates rather than other mammals or birds (with the exception of the PCTB that was applied to several parrot species). Our findings from these typical comparative cognition tasks suggest that individual differences in non-cognitive internal (subject-related) factors (such as personality-mediated intrinsic motivational factors) affected cognitive performance primarily via attention, which in principle can be controlled or at least quantified. Depending on the individuals and species tested, differences in social and demographic factors may positively or negatively affect performance. Unless cognitive comparisons specifically account for the influence of such factors on cognitive performance, it is therefore essential to report these potential sources of variation and control them if possible.

We conclude that non-cognitive factors are a minor issue if experimenters ensure they only test attentive individuals who are motivated to use appropriate response strategies. This is best achieved by either presenting more than two choice options to the subjects whenever possible or by using modified two-choice task versions that prevent motivational issues. Since relatively small differences in task format and test procedures can have major effects on the outcomes of comparative cognition studies, it is essential to report the testing procedure and individual results in detail, ideally supplemented with video clips. While basic aspects of internal task validity can thus be improved by establishing that the tasks measure at least cognition *per se*, the more specific issue of construct validity remains. Since all cognitive abilities represent constructs and as such have to be inferred from cognitive test performance, each construct needs to be carefully defined before its validity can be established. Although establishing construct validity in a top-down (theory-based) rather than a bottom-up approach (statistically) would be ideal, doing so is extremely challenging (but see Burkart et al., 2017a,b for recommendations on how to achieve this).

In sum, test design remains the major issue of contemporary comparative psychology and it is essential that researchers validate and redesign cognitive tests, if needed, in order to ascertain that the tasks accurately capture cognitive abilities. Once the sensitivity, reliability, and internal validity of cognitive tasks have been established, these tasks can then be integrated into test batteries and applied to an increasingly wide range of species. This will also help establishing their external validity, i.e., if they measure the same cognitive abilities in different species. Such evaluated test batteries that only include tasks with established internal and external validity will then hopefully provide a solid base for future cognitive comparisons and further our understanding of the evolution of mind, including the human mind.

AUTHOR CONTRIBUTIONS

MS, CF, and JB wrote the review together. All authors contributed to the article and approved the submitted version.

FUNDING

This research was financially supported by the Swiss National Science Foundation (project funding to JB; project number 310030-120383) and the A.H. Schultz-Foundation (two individual grants to MS).

REFERENCES

Addessi, E., Crescimbene, L., and Visalberghi, E. (2008). Food and token quantity discrimination in capuchin monkeys (*Cebus apella*). *Anim. Cogn.* 11, 275–282. doi: 10.1007/s10071-007-0111-116

Amici, F., Aureli, F., and Call, C. (2010). Monkeys and apes: are their cognitive skills really so different? *Am. J. Phys. Anthropol.* 143, 188–197. doi: 10.1002/ajpa.21305

Amici, F., Aureli, F., and Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Curr. Biol.* 18, 1415–1419. doi: 10.1016/j.cub.2008.08.020

Amici, F., Barney, B., Johnson, V. E., Call, J., and Aureli, F. (2012). A modular mind? A test using individual data from seven primate species. *PLoS One* 7:e51918. doi: 10.1371/journal.pone.0051918

Amodio, P., Boeckle, M., Schnell, A. K., Ostojíc, L., Fiorito, G., and Clayton, N. S. (2019). Grow smart and die young: why did cephalopods evolve intelligence? *Trends Ecol. Evol.* 34, 45–56. doi: 10.1016/j.tree.2018.10.010

Anderson, J. R., Sallaberry, P., and Barbier, H. (1995). Use of experimenter-given cues during object-choice tasks by capuchin monkeys. *Anim. Behav.* 49, 201–208. doi: 10.1016/00033473(95)80168

Banerjee, K., Chabris, C. F., Johnson, V. E., Lee, J. J., Tsao, F., and Hauser, M. D. (2009). General intelligence in another primate: individual differences across cognitive task performance in a new world monkey (*Saguinus oedipus*). *PLoS One* 4:e5883. doi: 10.1371/journal.pone.0005883

Bard, K. A., Bakeman, R., Boysen, S. T., and Leavens, D. A. (2014). Emotional engagements predict and enhance social cognition in young chimpanzees. *Devel. Sci.* 17, 682–696. doi: 10.1111/desc.12145

Barth, J., Reaux, J. E., and Povinelli, D. J. (2005). Chimpanzees (Pan troglodytes) use of gaze cues in object-choice tasks: different methods yield different results. *Anim. Cogn.* 8, 84–92. doi: 10.1007/s10071-004-0235-x

Bayley, N. (1969). *Bayley Scales of Infant Development*. San Antonio, TX: Psychological Corporation.

Bensky, M. K., Gosling, S. D., and Sinn, D. L. (2013). The world from a dog's point of view: a review and synthesis of dog cognition research. *Adv. Study Behav.* 45, 209–406. doi: 10.1016/B978-0-12-407186-5.00005-7

Beran, M. J., Perdue, B. M., Church, B. A., and Smith, J. D. (2016). Capuchin monkeys (*Cebus apella*) modulate their use of an uncertainty response depending on risk. *J. Exp. Psychol. Anim. L.* 42, 32–43. doi: 10.1037/xan0000080

Beran, M. J., Perdue, B. M., and Smith, J. D. (2014). What are my chances? Closing the gap in uncertainty monitoring between rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*). *J. Exp. Psychol. Anim. L.* 40, 303–316. doi: 10.1037/xan0000020

Bhsary, R., and Brown, C. (2014). Fish cognition. *Curr. Biol.* 24, R947–R950. doi: 10.1016/j.cub.2014.08.043

Bräuer, J., Kaminski, J., Riedel, J., Call, J., and Tomasello, M. (2006). Making inferences about the location of hidden food: social dog, causal ape. *J. Comp. Psychol.* 120, 38–47. doi: 10.1037/0735-7036.120.1.38

Burkart, J., and Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *J. Comp. Psychol.* 120, 120–130. doi: 10.1037/0735-7036.120.2.120

Burkart, J. M. (2015). Opposite effects of male and female helpers on social tolerance and proactive prosociality in callitrichid family groups. *Sci. Rep.* 5, 1–9.

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., et al. (2014). The evolutionary origin of human hyper-cooperation. *Nat. Commun.* 5, 1–9.

Burkart, J. M., and Heschl, A. (2007). Understanding visual access in common marmosets, *Callithrix jacchus*: perspective taking or behaviour reading? *Anim. Behav.* 73, 457–469. doi: 10.1016/j.anbehav.2006.05.019

Burkart, J. M., Schubiger, M. N., and van Schaik, C. P. (2017a). Future directions for studying the evolution of general intelligence. *Behav. Brain Sci.* 40, 1–24.

Burkart, J. M., Schubiger, M. N., and van Schaik, C. P. (2017b). The evolution of general intelligence. *Behav. Brain Sci.* 40, 50–67.

Burkart, J. M., and van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Anim. Cogn.* 13, 1–19. doi: 10.1007/s10071-009-0263-7

Burkart, J. M., and van Schaik, C. P. (2016). Revisiting the consequences of cooperative breeding. *J. Zool.* 299, 77–83. doi: 10.1111/jzo.12322

Byrne, R. W., Bates, L. A., and Moss, C. J. (2009). Elephant cognition in primate perspective. *Comp. Cogn. Behav. Rev.* 4, 65–79. doi: 10.3819/ccbr.2009.40009

Carducci, P., Schwing, R., Huber, L., and Truppa, V. (2018). Tactile information improves visual object discrimination in kea, Nestor notabilis, and capuchin monkeys, *Sapajus* spp. *Anim. Behav.* 135, 199–207. doi: 10.1016/j.anbehav.2017.11.018

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., et al. (2018). The repeatability of cognitive performance: a meta-analysis. *Phil. Trans. R. Soc. B* 373:20170281. doi: 10.1098/rstb.2017.0281

Chittka, L. (2017). Bee cognition. *Curr. Biol.* 27, R1037–R1059.

Cronin, K. A., Van Leeuwen, E. J., Vreeman, V., and Haun, D. B. M. (2014). Population-level variability in the social climates of four chimpanzee societies. *Evol. Hum.Behav.* 35, 389–396. doi: 10.1016/j.evolhumbehav.2014.05.004

Dale, R., Despraz, M.-N., Marshall-Pescini, S., and Range, F. (2019). Piloting a new prosociality paradigm in dogs and wolves_ The location choice task. *Behav. Process.* 162, 79–85. doi: 10.1016/j.beproc.2019.01.004

Damerius, L. A., Burkart, J. M., van Nordwijk, M. A., Haun, D. B. M., Kosonen, Z. K., Galdikas, B. M. F., et al. (2019). General cognitive abilities in orangutans (*Pongo abelii* and *Pongo pygmaeus*). *Intelligence* 74, 3–11. doi: 10.1016/j.intell.2018.10.008

Damerius, L. A., Forss, S. I. F., Kosonen, Z. K., Willems, E. P., Burkart, J. M., Call, J., et al. (2017a). Orientation toward humans predicts cognitive performance in orang-utans. *Sci. Rep.* 7:40052. doi: 10.1038/srep40052

Damerius, L. A., Graber, S., Willems, E. P., and van Schaik, C. P. (2017b). Curiosity boosts orang-utan problem-solving ability. *Anim. Behav.* 134, 57–70. doi: 10.1016/j.anbehav.2017.10.005

Davenport, R. K., Rogers, C. M., and Rumbaugh, D. M. (1973). Long-term cognitive deficits in chimpanzees associated with early impoverished rearing. *Dev. Psychol.* 9:343. doi: 10.1037/h0034877

Deaner, R. O., Johnson, V. E., and van Schaik, C. P. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evol. Psychol.* 4, 149–196. doi: 10.1177/147470490600400114

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., et al. (2009). Age-associated cognitive decline. *Br. Med. Bull.* 92, 135–152. doi: 10.1093/bmb/ldp033

Dunbar, R. I. M. (1988). *Primate Social Systems*. Ithaca: Cornell Univ. Press.

Earp, B. D., and Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Front. Psychol.* 6:621. doi: 10.3389/fpsyg.2015.00621

Emery, N. J., and Clayton, N. S. (2004). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306, 1903–1907. doi: 10.1126/science.1098410

Farrar, B. G., Boeckle, M., and Clayton, N. S. (2020). Replications in comparative cognition: what should we expect and how can we improve. *Anim. Behav. Cogn.* 7, 1–22. doi: 10.26451/abc.07.01.02.2020

Fichtel, C., Dinter, K., and Kappeler, P. (2020). The lemur baseline: how lemurs compare to monkeys and apes in the primate cognition test battery. *BioRxiv* [Preprint]. doi: 10.1101/2020.04.21.052852

Fichtel, C., Schnoell, A. V., and Kappeler, P. M. (2018). Measuring social tolerance – an experimental approach in two lemurid primates. *Ethology* 124, 65–73. doi: 10.1111/eth.12706

Freitag, F. B., and Almeida Pessoa, D. M. (2012). Effect of luminosity on color discrimination of dichromatic marmosets (*Callithrix jacchus*). *J. Opt. Soc. Am. A* 29, A216–A222. doi: 10.1364/JOSAA.29.00A216

Gallup, G. G. J., McClure, M. K., Hill, S. D., and Bundy, R. A. (1971). Capacity for self-recognition in differentially reared chimpanzees. *Psychol. Rec.* 21, 69–74. doi: 10.1007/bf03393991

Gazes, R. P., Billas, A. R., and Schmitt, V. (2018). Impact of stimulus format and reward value on quantity discrimination in capuchin and squirrel monkeys. *Learn. Behav.* 46, 89–100. doi: 10.3758/s13420-017-0295-299

Giret, N., Miklósi, Á, Kreutzer, M., and Bovet, D. (2009). Use of experimenter-given cues by African gray parrots (*Psittacus erithacus*). *Anim. Cogn.* 12:e10.

Gokcekus, S. Y. (2020). *Behavioral Flexibility, Curiosity, and Cooperative Breeding: Dealing with Complex Concepts and Paradigms*. Durham theses, Durham University, Durham.

Güntürkün, O. (2014). Is dolphin cognition special? *Brain. Behav. Evol.* 83, 177–180. doi: 10.1159/000357551

Güntürkün, O., and Bugnyar, T. (2016). Cognition without cortex. *Trends Cogn. Sci.* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Hare, B., Addessi, E., Call, J., Tomasello, M., and Visalberghi, E. (2003). Do capuchin monkeys, *Cebus apella*, know what conspecifics do and do not see? *Anim. Behav.* 65, 131–142. doi: 10.1006/anbe.2002.2017

Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Anim. Behav.* 59, 771–785. doi: 10.1006/anbe.1999.1377

Hare, B., Melis, A. P., Woods, V., Hastings, S., and Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Curr. Biol.* 17, 619–623. doi: 10.1016/j.cub.2007.02.040

Hare, B., and Tomasello, M. (2004). Chimpanzees are more skilful in competitive than cooperative cognitive tasks. *Anim. Behav.* 68, 571–581. doi: 10.1016/j.anbehav.2003.11.011

Heldstab, S. A., Kosonen, Z. K., Koski, S. E., Burkart, J. M., van Schaik, C. P., and Isler, K. (2016). Manipulation complexity in primates coevolved with brain size and terrestriality. *Sci. Rep.* 6:24528. doi: 10.1038/srep24528

Herrmann, E., Call, J., Hernàndez-Lloreda, M. V., Hare, B., and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317, 1360–1366. doi: 10.1126/science.1146282

Herrmann, E., Hare, B., Call, J., and Tomasello, M. (2010). Differences in the cognitive skills of bonobos and chimpanzees. *PLoS One* 5:e12438. doi: 10.1371/journal.pone.0012438

Holekamp, K. E., and Benson-Amram, S. (2017). The evolution of intelligence in mammalian carnivores. *Interface Focus* 7:20160108. doi: 10.1098/rsfs.2016.0108

Hopkins, W. D., Russel, J. L., and Schaeffer, J. (2014). Chimpanzee intelligence is heritable. *Curr. Biol.* 24, 1649–1652. doi: 10.1016/j.cub.2014.05.076

Horner, V., and Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Anim. Cogn.* 8, 164–181. doi: 10.1007/s10071-004-0239-6

Huber, L. (2017). Where is the evidence for general intelligence in nonhuman animals? *Behav. Brain Sci.* 40, 33–34.

Jaeggi, A. V., Burkart, J. M., and van Schaik, C. P. (2010a). On the psychology of cooperation in humans and other primates: combining the natural history and experimental evidence of prosociality. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 2723–2735. doi: 10.1098/rstb.2010.0118

Jaeggi, A. V., Stevens, J. M., and Van Schaik, C. P. (2010b). Tolerant food sharing and reciprocity is precluded by despotism among bonobos but not chimpanzees. *Am. J. Phys. Anthropo.* 143, 41–51. doi: 10.1002/ajpa.21288

Jaeggi, A. V., and van Schaik, C. P. (2011). The evolution of food sharing in primates. *Behav. Ecol. Sociobiol.* 65, 2125–2140. doi: 10.1007/s00265-011-1221-3

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159.

Joly, M., Micheletta, J., De Marco, A., Langermans, J. A., Sterck, E. H. M., and Waller, B. M. (2017). Comparing physical and social cognitive skills in macaque species with different degrees of social tolerance. *Proc. Biol. Sci.* 284:20162738. doi: 10.1098/rspb.2016.2738

Kappeler, P. M., and van Schaik, C. P. (2002). Evolution of primate social systems. *Int. J. Primatol.* 23, 707–740. doi: 10.1023/A:1015520830318

Karin-D'Arcy, R. M., and Povinelli, D. J. (2002). Do chimpanzees know what each other see? A closer look. *Int. J. Comp. Psychol.* 15, 21–54.

King, G. E. (2016). *Primate Behavior and Human Origins*. London: Routledge.

King, J. E., and Figuereo, A. J. (1997). The five-factor model plus dominance in chimpanzee personality. *J. Res. Pers.* 31, 257–271. doi: 10.1006/jrpe.1997.2179

Kittler, K., Kappeler, P. M., and Fichtel, C. (2018). Instrumental problem solving in three lemur species. *J. Comp. Psychol.* 302, 306–314. doi: 10.1037/com0000113

Koenig, A. (1998). Visual scanning by common marmosets (*Callithrix jacchus*): functional aspects and the special role of adult males. *Primates* 39, 85–90. doi: 10.1007/bf02557746

Krasheninnikova, A., Berardi, R., Lind, M.-A., and O'Neill, L. (2019). Primate cognition test battery in parrots. *Behaviour* 156, 721–761. doi: 10.1163/1568539x-00003549

Lacreuse, A., Russell, J. L., Hopkins, W. D., and Herndon, J. G. (2014). Cognitive and motor aging in female chimpanzees. *Neurobiol. Aging* 35, 1–22.

Lambert, M., Jacobs, I., Osvath, M., and von Bayern, A. M. P. (2018). Birds of a feather? Parrot and corvid cognition compared. *Behaviour* 156, 505–594. doi: 10.1163/1568539X-00003527

Limongelli, L., Boysen, S. T., and Visalberghi, E. (1995). Comprehension of cause-effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 109, 18–26. doi: 10.1037/0735-7036.109.1.18

MacLean, E. L., Hare, B., Nunn, C. L., Adessi, E., Amici, F., Anderson, J. R., et al. (2014). The evolution of self-control. *PNAS* 111, E2140–E2148. doi: 10.1073/pnas.132533111

MacPhail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. B* 308, 37–51. doi: 10.1098/rstb.1985.0008

MacPhail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–656.

Maestripieri, D. (2012). "Comparative primate psychology," in *Encyclopedia of Human Behavior*, ed. V. S. Ramachandran (Amsterdam: Elsevier), 550–555. doi: 10.1016/b978-0-12-375000-6.00101-4

Manger, P. R. (2013). Questioning the interpretations of behavioral observations of cetaceans: is there really support for a special intellectual status for this mammalian order? *Neuroscience* 250, 664–696. doi: 10.1016/j.neuroscience.2013.07.041

ManyPrimates, Altschul, D. M., Beran, M. J., Bohn, M., Call, J., DeTroy, S., et al. (2019a). Establishing an infrastructure for collaboration in primate cognition research. *PLoS One* 14:e0223675. doi: 10.1371/journal.pone.0223675

ManyPrimates, Altschul, D. M., Beran, M. J., Bohn, M., Caspar, K. R., Fichtel, C., et al. (2019b). Collaborative open science as a way to reproducibility and new insights in primate cognition research. *Jpn. Psychol. Rev.* 62, 205–220.

Mather, J. A., and Kuba, M. J. (2013). The cephalopod specialties: complex nervous system, learning, and cognition. *Can. J. Zool.* 91, 31–49.

Matzel, L. D., and Sauce, B. (2017). Individual differences: case studies of rodent and primate intelligence. *J. Exp. Psychol. Anim. Learn. Cogn.* 43, 325–340. doi: 10.1037/xan0000152

Morton, F. B., Lee, P. C., and Buchanan-Smith, H. M. (2013). Taking personality selection bias seriously in animal cognition research: a case study in capuchin monkeys (*Sapajus apella*). *Anim. Cogn.* 16, 677–684. doi: 10.1007/s10071-013-0603-605

Mulcahy, N. J., and Call, J. (2006). How great apes perform on a modified trap-tube task. *Anim. Cogn.* 9, 193–199. doi: 10.1007/s10071-006-0019-6

Mulcahy, N. J., and Call, J. (2009). The performance of bonobos (*Pan paniscus*), chimpanzees (*Pan troglodytes*), and orangutans (*Pongo pygmaeus*) in two versions of an object-choice task. *J. Comp. Psychol.* 123, 304–309. doi: 10.1037/a0016222

Mulcahy, N. J., and Hedge, V. (2012). Are great apes tested with an abject object-choice task? *Anim. Behav.* 83, 313–321. doi: 10.1016/j.anbehav.2011.11.019

Mulcahy, N. J., and Suddendorf, T. (2011). An obedient orangutan (*Pongo abelii*) performs perfectly in peripheral object-choice tasks but fails the standard centrally presented versions. *J. Comp. Psychol.* 125, 112–115. doi: 10.1037/a0020905

Open Science Collaboration (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspect. Psychol. Sci.* 7, 657–660. doi: 10.1177/1745691612462588

Pashler, H., and Wagenmaekers, E. J. (2012). A crisis of confidence? *Perspect. Psychol. Sci.* 7, 528–530. doi: 10.1177/1745691612465253

Patterson, F., and Cohn, R. (1994). "Self-recognition and self-awareness in lowland gorillas," in *Self-Awareness in Animals and Humans Developmental Perspectives*, eds T. S. Parker, R. W. Mitchell, and L. Boccia (Cambridge: University Press).

Pepperberg, I. M. (2002). Cognitive and communicative abilities of grey parrots. *Curr. Dir. Psychol. Sci.* 11, 83–87.

Pepperberg, I. M. (2006). Grey parrot numerical competence: a review. *Anim. Cogn.* 9, 377–391. doi: 10.1007/s10071-006-0034-7

Pessoa, D. M. A., Tomaz, C., and Pessoa, V. F. (2005). Color vision in marmosets and tamarins: behavioural evidence. *Am. J. Primatol.* 67, 487–495. doi: 10.1002/ajp.20202

Petit, O., Desportes, C., and Thierry, B. (1992). Differential probability of "coproduction" in two species of macaque (*Macaca tonkeana*, *M. mulatta)*. *Ethology* 90, 107–120. doi: 10.1111/j.1439-0310.1992.tb00825.x

Plotnik, J. M., Waal, F. B. M. D., and Reiss, D. (2006). Self-recognition in an Asian elephant. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17053–17057. doi: 10.1073/pnas.0608062103

Posada, S., and Colell, M. (2007). Another gorilla (*Gorilla gorilla* gorilla) recognizes himself in a mirror. *Am. J. Primatol.* 69, 1–8.

Povinelli, D. J. (2000). *Folk Physics For Apes: A Chimpanzee's Theory of How the Mind Works*. Oxford: Oxford University Press.

Ramus, F. (2017). General intelligence is an emerging property, not an evolutionary puzzle. *Behav. Brain Sci.* 40, 43–44.

Reader, S. M., Hager, Y., and Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1017–1027. doi: 10.1098/rstb.2010.0342

Rumbaugh, D. M. (1971). Evidence of qualitative differences in learning processes among primates. *J. Comp. Physiol. Psychol.* 76, 250–255. doi: 10.1037/h0031401

Scarf, D., Hayne, H., and Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science* 334:1664. doi: 10.1126/science.1213357

Schmitt, V., and Fischer, J. (2011). Representational format determines numerical competence in monkeys. *Nat. Commun.* 2:257. doi: 10.1038/ncomms1262

Schmitt, V., Pankau, B., and Fischer, J. (2012). Old world monkeys compare to apes in the primate cognition test battery. *PLoS One* 7:e32024. doi: 10.1371/journal.pone.0032024

Schmon, R. (2011). *Action Understanding and Theory of Mind in Common Marmosets (Callithrix jacchus)*. Master thesis, University of Zurich, Zurich.

Schubiger, M. N. (2019). *Optimising Tests of Primate Cognition - Towards Valid Species Comparisons*. Zürich: University of Zurich.

Schubiger, M. N., Kissling, A., and Burkart, J. M. (2016). How task format affects cognitive performance: a memory test with two species of New World monkeys. *Anim. Beh.* 121, 33–39. doi: 10.1016/j.anbehav.2016.08.005

Schubiger, M. N., Kissling, A., and Burkart, J. M. (2019). Does opportunistic testing bias cognitive performance in primates? Learning from drop-outs. *PLoS One* 14:e0213727. doi: 10.1371/journal.pone.0213727

Schubiger, M. N., Wüstholz, F. L., Wunder, A., and Burkart, J. M. (2015). Emotional reactivity affects participation, but not performance in cognitive tests with common marmosets (*Callithrix jacchus*). *Anim. Cogn.* 18, 701–712. doi: 10.1007/s10071-015-0837-5

Seed, A. M., Call, J., Emery, N. J., and Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *J. Exp. Psychol. Anim. Behav. Process.* 35, 23–34. doi: 10.1037/a0012925

Smith, P. L., and Little, D. R. (2018). Small is beautiful: in defense of the small-N design. *Psychon. Bull. Rev.* 25, 2083–2101. doi: 10.3758/s13423-018-1451-8

Solvi, C., Gutierrez Al-Khudhairy, S., and Chittka, L. (2020). Bumble bees display cross-modal object recognition between visual and tactile senses. *Science* 21, 910–912. doi: 10.1126/science.aay8064

Strasser, A., and Burkart, J. M. (2012). Can we measure brain efficiency? An empirical test with common marmosets (*Callithrix jacchus*). *Brain Behav. Evol.* 80, 26–40. doi: 10.1159/000338014

Tomasello, M., and Call, J. (1997). *Primate Cognition*. Oxford: Oxford University Press.

Torigoe, T. (1985). Comparison of object manipulation among species of non-human primates. *Primates* 26, 182–194. doi: 10.1007/bf02382017

Townsend, S. W., Allen, C., and Manser, M. B. A. (2012). Simple test of vocal individual recognition in wild meerkats. *Biol. Lett.* 8, 179–182. doi: 10.1098/rsbl.2011.0844

Toxopeus, I. B., Sterck, E. H. M., van Hooff, J. A. R. A. M., Spruijt, B. M., and Heeren, T. J. (2005). Effects of trait anxiety on performance of socially housed monkeys in a learning test. *Behaviour* 142, 1269–1287. doi: 10.1163/156853905774539373

van Schaik, C. P., and Burkart, J. M. (2011). Social learning and evolution: the cultural intelligence hypothesis. *Phil. Trans. R. Soc. B* 366, 1008–1016. doi: 10.1098/rstb.2010.0304

Visalberghi, E., and Limongelli, L. (1994). Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 108, 15–22. doi: 10.1037/0735-7036.108.1.15

Vonk, J., and Leete, J. A. (2017). Carnivore concepts: categorization in carnivores "bears" further study. *Int. J. Comp. Psychol.* 30:32707.

Weiss, A. (2017). "Exploring factor space (and other adventures) with the hominoid personality questionnaire," in *Personality in Nonhuman Animals*, eds J. Vonk, A. Weiss, and S. Kuczaj (Cham: Springer).

Wilkinson, A., and Huber, L. (2012). "Cold-blooded cognition: reptilian cognitive abilities," in *The Oxford Handbook of Comparative Evolutionary Psychology*, eds J. Vonk and T. K. Shackelford (New York, NY: Oxford), 129–141. doi: 10.1073/pnas.93.10.5122

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 08 September 2020
doi: 10.3389/fpsyg.2020.02053

The Relevance of Ecological Transitions to Intelligence in Marine Mammals

Gordon B. Bauer[1,2], Peter F. Cook[1,2] and Heidi E. Harley[1,2,3]*

[1]Division of Social Sciences, New College of Florida, Sarasota, FL, United States, [2]Mote Marine Laboratory, Sarasota, FL, United States, [3]The Seas, Epcot®, Walt Disney World® Resorts, Lake Buena Vista, FL, United States

Macphail's comparative approach to intelligence focused on associative processes, an orientation inconsistent with more multifaceted lay and scientific understandings of the term. His ultimate emphasis on associative processes indicated few differences in intelligence among vertebrates. We explore options more attuned to common definitions by considering intelligence in terms of richness of representations of the world, the interconnectivity of those representations, the ability to flexibly change those connections, and knowledge. We focus on marine mammals, represented by the amphibious pinnipeds and the aquatic cetaceans and sirenians, as animals that transitioned from a terrestrial existence to an aquatic one, experiencing major changes in ecological pressures. They adapted with morphological transformations related to streamlining the body, physiological changes in respiration and thermoregulation, and sensory/perceptual changes, including echolocation capabilities and diminished olfaction in many cetaceans, both in-air and underwater visual focus, and enhanced senses of touch in pinnipeds and sirenians. Having a terrestrial foundation on which aquatic capacities were overlaid likely affected their cognitive abilities, especially as a new reliance on sound and touch, and the need to surface to breath changed their interactions with the world. Vocal and behavioral observational learning capabilities in the wild and in laboratory experiments suggest versatility in group coordination. Empirical reports on aspects of intelligent behavior like problem-solving, spatial learning, and concept learning by various species of cetaceans and pinnipeds suggest rich cognitive abilities. The high energy demands of the brain suggest that brain-intelligence relationships might be fruitful areas for study when specific hypotheses are considered, e.g., brain mapping indicates hypertrophy of specific sensory areas in marine mammals. Modern neuroimaging techniques provide ways to study neural connectivity, and the patterns of connections between sensory, motor, and other cortical regions provide a biological framework for exploring how animals represent and flexibly use information in navigating and learning about their environment. At this stage of marine mammal research, it would still be prudent to follow Macphail's caution that it is premature to make strong comparative statements without more empirical evidence, but an approach that includes learning more about how animals flexibly link information across multiple representations could be a productive way of comparing species by allowing them to use their specific strengths within comparative tasks.

Keywords: intelligence, marine mammals, cetaceans, pinnipeds, sirenians, brain

OPEN ACCESS

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Eduardo Mercado,
University at Buffalo, United States
Fiona Cross,
University of Canterbury, New Zealand

***Correspondence:**
Gordon B. Bauer
bauer@ncf.edu

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *29 April 2020*
Accepted: *24 July 2020*
Published: *08 September 2020*

Citation:
Bauer GB, Cook PF and Harley HE (2020) The Relevance of Ecological Transitions to Intelligence in Marine Mammals. Front. Psychol. 11:2053. doi: 10.3389/fpsyg.2020.02053

THE RELEVANCE OF ECOLOGICAL TRANSITIONS TO INTELLIGENCE IN MARINE MAMMALS

Since the birth of psychology, scientists have debated the power of associationism as the central mechanism behind "the science of mental life" (James, 1890/1952, p. 1). James began his seminal psychological work by contrasting three different possibilities for describing the human mind: soul, associative processes dictated by experience, and the innate and developed structure of the human mind as a framework that constrains how we process information. Macphail (1982, 1987), after a comprehensive review, concluded that an objective assessment of the vertebrate literature indicated that learning only occurred through a limited set of processes, primarily associative (i.e., habituation, classical, and instrumental conditioning). He further indicated that, restricting comparisons to associative processes, there were no differences in intelligence among vertebrates with the exception of humans, and he speculated that difference might be attributable to language, essentially discounting specific cognitive adaptations to distinct niches. Of course, the circumstances in which animals perform and the underlying mechanisms they use are actually the province of psychology, and Macphail's assertion does not inspire a productive comparative research agenda going forward. Here, we explore a broader framework for interpreting intelligent behavior in animals using widely studied marine mammals (bottlenose dolphins, sea lions, harbor seals, and West Indian manatees) as examples. Their evolutionary history, notably the transition of terrestrial mammals into marine species, forced many adaptations including unique sensory systems, complex social organization, and neurobiological extremes. In addition, many of these animals show flexible cognition, at least, comparable to what has been observed in primates.

Macphail's definition of intelligence, limited to associative processes, is not consistent with common conceptualizations of human intelligence by experts or the lay public. Expert conceptualizations of human intelligence are multifaceted and include adaptation to the environment, mental processes, and higher order thinking (e.g., reasoning, problem solving, decision-making, and metacognition; review in Sternberg, 2003). Studies of lay ideas of intelligence in the United States identify factors such as speed of processing, practical problem solving, verbal ability, non-verbal reasoning, numerical reasoning, and social competence (Sternberg et al., 1981; Chen and Chen, 1988). Unfortunately, these definitional factors do not hold up well cross-culturally (Nagoshi, 1987), and there is no strong consensus among psychologists on what the components of intelligence should be, illustrating the problem of generalizing from any single human tradition. The origin of the concept of human intelligence and practical applications in individual differences (Binet and Simon, 1916; McNemar, 1964) suggest questionable utility for interspecies comparisons, unless we use individual variability itself, which may be a hallmark of an intelligent species, as a comparative measure. In addition, Mackintosh (1998) notes that associative learning as described by Macphail bears striking similarities to human implicit learning (e.g., Reber, 1993), an area typically not addressed on intelligence tests. This sets a conundrum for comparative psychologists because intelligence defined for humans excludes implicit (associative) processes, and so intelligence would then appear to lie outside the realm of comparative psychology and Darwinian evolution.

Comparative researchers responding to Macphail's null hypothesis of no species differences in intelligence emphasized that intelligence consists of multiple facets including sensory and perceptual processes, memory, spatial relations, concept formation, rule learning, and tool use (Goldman-Rakic and Preuss 1987; Hodos, 1987; Shettleworth, 1987; Rilling, 1990; Walker, 1990). Bullock (1986) suggested that candidates for investigation might include flexibility in interacting with the environment, social interactions, communication, and difficult, higher forms of cognition, plus problem solving across all the categories. In addition, he considered acquired knowledge essential to considerations of intelligence. Goldman-Rakic and Preuss (1987) and Vauclair (1990) among others also suggested that representation rather than association might be a more productive approach. Representations, based on the environmental information that animals extract through their sensory-motor systems and then organize perceptually and cognitively, vary widely across species and facilitate intelligent behavior. Associations between representations and the breadth and flexibility of those representations may be especially relevant for marine mammals, who become interesting due to their operating so effectively in two vastly different perceptual media – water and air – where they must recruit sensory-motor systems developed differentially for this split life. More recent approaches to animal intelligence retain a multifaceted approach (e.g., Roth and Dicke, 2017).

In considering the intelligence of marine mammals, we start with the assumption that marine mammals have the basic associative processes identified by Macphail (1982, 1987) and demonstrated ubiquitously in marine mammal training (Pepper and Defran, 1975), entertainments at commercial oceanaria, and numerous studies. We take an evolutionary stance that the transition from a terrestrial to an aquatic environment modified sensory and perceptual processes, as well as the flexibility and processing speed of other cognitive processes contributing to intelligence. We also investigate the implications of marine mammal neurobiology in the manifestation of intelligent behavior. For us, intelligence is the effectiveness by which one deploys cognitive processes including sensation and perception, instantiated in the central and peripheral nervous systems, and studied through investigations of behavior.

MARINE MAMMALS

Marine mammals can be characterized as the mammals that depend primarily on the marine environment for survival (Rice, 1998). This list could include marine otters (*Lontra felina*), polar bears (*Ursus maritimus*), Arctic foxes (*Vulpes lagopus*), and fishing bats (*Noctilio leporinus*), which feed on marine prey, but much more is known about the senses and cognition of cetaceans (whales and dolphins), pinnipeds (seals, sea lions, and walruses), and sirenians (manatees and dugongs), so our focus will be on these orders. There are species within these orders that are exclusively freshwater such as the river dolphins, family Iniidae and Platinistidae, and some of the manatee species such as Amazonian manatees (*Trichechus inunguis*), but by and large these orders are marine. Within each order only a few species have been studied, so some caution needs to be observed in generalizing across species, but these few species serve to provide a working base with which to compare other species. In addition, sample sizes for laboratory experimentation are small, frequently only one or two subjects, so conclusions are likely to be modified as more subjects are studied.

The ancestors of cetaceans (whales and dolphins) and sirenians (sea cows) made a major transition from a terrestrial to an aquatic environment during the Eocene (~50 million years ago). The pinnipeds (seals, sea lions, and walruses) made a partial transition more recently during the late Oligocene (~26–23 million years ago) and remain amphibious, feeding at sea but reproducing on land. These three orders, which constitute the most studied groups of marine mammals, responded to new ecological pressures with numerous adaptive changes in morphology, physiology, behavior, and sensory/perceptual processes, thereby shifting the information they could gain about the world and thus their representations of it.

Morphological changes included streamlining the body to reduce drag, including loss or reduction in hind limbs and modification of forelimbs, various other skeletal modifications, loss, reduction or modification in pelage, and internalization of male reproductive organs. Respiratory mechanisms had to meet simultaneous demands for the combination of in-air breathing with diving, often to great depths under great physical pressure. Circulatory systems were modified to maintain warm body temperature in cold water environments. Many species developed group social structures and cooperative systems for foraging and defense in an environment with few places to hide either for purposes of prey ambush or predator avoidance, especially for animals coupled to the surface for respiration. Communication systems emphasized auditory and tactile channels, while de-emphasizing or modifying visual systems, which were limited by low light and turbidity underwater and sharp transitions in brightness at the surface. Olfactory systems that evolved on land had limited utility underwater. Novel sensory processes, such as echolocation and exquisite senses of touch, were shaped by natural selection to facilitate foraging and orientation in a marine environment. An evolutionary perspective suggests that adaptation to the aquatic realm overlain on a terrestrial foundation likely affected an array of behavioral and cognitive intellectual processes, preserving some attributes while modifying others.

Bullock (1986) provides an entry to a comparative assessment of intelligence beyond associationism by presenting a broad palette of candidate domains for the investigation of animal intelligence. We have selected from that palette to emphasize flexibility in problem solving, the neural plasticity that underlies flexibility, and knowledge. Curiously, Macphail (1987) also emphasized the generality of human intelligence, as well as its dependence on experience. Knowledge is a little studied topic in marine mammal science, but we can identify mechanisms that would allow the accumulation of knowledge: the resolution (perceptual detail) of sensory systems, the speed of information transfer by imitation, retention over long time periods, and facility at problem solving. We have organized this information into four categories: sensation and perception, social learning, flexibility of cognitive processes, and the brain.

SENSATION AND PERCEPTION

Sensory-motor experiences provide a foundation for intelligent thought by providing insight to the quality, range, and resolution of animal worlds or *Umwelten* (von Uexküll, 1934/1957), the detail creating the representations operated on during cognitive processing to produce intelligent behavior. Early researchers (Galton, 1883; Cattell, 1890; Spearman, 1904) considered sensory discrimination as integral to human intelligence, but their view failed to gain traction in mental measurement (Deary, 1994; Sternberg, 2003). Nevertheless, subsequent research provided support for this sensory hypothesis: for example, Deary et al. (2004) reported a high correlation between general sensory discrimination (representing shared variance across several modalities) and fluid intelligence (Cattell, 1963), which is closely related to working memory (Kyllonen and Christal, 1990; Salthouse and Pink, 2008). Intelligence differences between animal species may even more strongly reflect sensory processing, since sensory differences between species are more likely to be greater than differences within a single species, such as humans.

The marine environment places specific demands on sensory perception. The slow rate of diffusion of chemical compounds in water limits their utility to marine mammals compared to terrestrial mammals living where olfactants are rapidly dispersed. The olfactory systems of terrestrial mammals, designed to detect and discriminate airborne compounds, are of reduced importance to animals spending substantial time underwater. Taste may be relevant but the overall sense of flavor (i.e., combined effect of taste and smell) is probably lessened to the extent that olfaction is unavailable. Underwater vision is constrained by the limits of photic transmission in water, and it loses much of its relevance at depth or in turbid environments where light is limited. Touch provides advantages underwater for sensing hydrodynamic movement caused by currents or distortions in water flowing past objects, as well as for close contact investigation of items. Sound in water travels close to 4.5 times as fast as sound in air and can be conveyed with fidelity over great distances. The long wavelengths of lower frequency sounds allow them to pass around objects that block light transmission, and high frequency sounds are capable of transmitting detailed information over shorter distances. Adaptations for enhanced acoustic and tactile processing required for life underwater not only fostered new sensory mechanisms for gaining important information but also pushed speed and range of processing to new heights due to the physics of sound transmission and pressure changes in water.

Cetaceans

There are over 80 species of odontocetes (toothed whales) living in diverse environments. The river dolphins, who live in muddy waters thick with particulate matter, have extremely poor eyesight. For example, Platanistidae, the Southeast Asian river dolphins, are probably capable of seeing only degrees of brightness and the *Inia*, the South American river dolphins, have visual acuity of over 40 arc min (Mass and Supin, 1989). Bottlenose dolphins (*Tursiops truncatus*) have considerably better resolution, useful in the frequently more transparent water of coastal regions. Underwater visual acuity for bottlenose dolphins is about 8.5 min and in-air is 12.5 min (Herman et al., 1975). This reasonably good acuity underwater and in-air is surprising because of the differential role the cornea plays in refraction underwater (practically none) and in air where it is the primary refractive component of the eye. An eye adapted for vision underwater should not be able to focus in air and vice-versa, without specific adaptive mechanisms, which dolphins have (Herman et al., 1975). Dolphins, like other marine mammals tested, are monochromats (Ahnelt and Kolb, 2000) who see the world in shades of gray (Madsen and Herman, 1980), although there is evidence that they may have some color perception, presumably mediated by the differential sensitivity of rods and the single cone-type (Griebel and Schmid, 2002). Dolphins depend on vision to build their representations of the world, their *Umwelten*. For example, they integrate information from multiple sensory systems, like vision and echolocation, to represent objects (e.g., Harley et al., 1996), and they can discriminate among photographs and video of fish underwater using vision alone, likely an important ability for stealthy foraging (Harley et al., 2019).

Dolphin hearing is exceptional (reviews in Supin et al., 2001; Au and Hastings, 2008) ranging from 0.15 to 200 kHz, an upper limit over three octaves higher than that of humans. They are excellent at sound localization with 0.5–4 degrees of resolution. They have a temporal processing rate, the ability of the nervous system to map sound intervals, as measured by auditory brainstem responses, over 1,500 Hz for amplitude modulated sounds (compared to a rate of 50 Hz for humans). These evoked potential measures provide only indirect measures of temporal processing. Behavioral tests, direct tests of the ability of dolphins to discriminate sound intervals, indicate that they have a temporal integration time of an order of magnitude less than humans do (Supin et al., 2001). Dolphins are also active echolocators that have the ability to make subtle distinctions among object characteristics, e.g., they can discriminate cylinders that vary in wall thickness by less than a millimeter (Au and Pawloski, 1992). They can also recognize an equivalence between their visual and echoic experiences of objects (Harley et al., 2003) and share echoic information with nearby eavesdropping dolphins about object identity (Xitco and Roitblat, 1996). Clearly, their representations of objects are fine-tuned and flexible.

The sense of touch in dolphins has not been investigated to the same extent as in other marine mammals but electrophysiological measures of skin response show greatest sensitivity around the head (Ridgway and Carder, 1990, 1993) with sensitivity comparable to human lips and fingers, sufficient to detect underwater turbulence (Kolchin and Belkovich, 1973). Hair, important for touch in other marine mammals, has not been investigated well in cetaceans, probably because of its infrequent appearance among odontocetes, where it is found only on the rostrum of river dolphins and some neonates of other species. Sensory hairs are found on the rostrums of baleen whales, but they are difficult to study in these large, pelagic animals, although the structure of hairs of right whales appear to be adapted for detection of small prey such as plankton (Murphy et al., 2015).

Cetaceans have missing or greatly reduced olfactory bulbs and ethmoturbinates. Their taste buds are few. Nevertheless, they have low thresholds for sour (citric acid) and bitter (Nachtigall and Hall, 1984; Friedl et al., 1990; Kuznetsov, 1990). They also can detect salt.

Pinnipeds

Although a large number of species comprise the pinnipeds (seals, sea lions, and walruses), most sensory research has been conducted on the California sea lion (*Zalophus californianis*) and harbor seal (*Phoca vitulina*). The vision of the pinnipeds may be most notable for relatively high acuity both in air and underwater. The large, curved orbit of the lens focuses light on the retina underwater. This would lead to myopic (near-sighted) vision in air, except the cornea in pinnipeds contains a flattened area over the pupil reducing or eliminating corneal refraction in air (West et al., 1991; Miller et al., 2010). The underwater and in-air acuity of the sea lion are equivalent at moderate and high brightness at 4.7–7 arc min, but underwater vision is better under dim light conditions. Seal vision is similar at 5–8 min. Pinnipeds are monochromats, and therefore, do not have dichromatic color vision as do most terrestrial mammals (Ahnelt and Kolb, 2000), although a weak form of mesopic color vision in seals has also been reported (Oppermann et al., 2016). These reports of a weak form of color vision based on rod-cone spectral sensitivity differences (Griebel and Peichel, 2003) have been challenged (Scholtyssek et al., 2015).

Audiograms for pinnipeds tend to have considerable variability among studies, perhaps attributable to small sample sizes (frequently just one animal) and individual differences, but in general, the frequency range for harbor seals is about 0.2–72 kHz (Kastelein et al., 2009), with sea lions having a somewhat lower upper limit. Early reports of hearing by pinnipeds suggested that underwater hearing was superior, but more recent evidence suggests that they are similar with both having low threshold levels (Reichmuth et al., 2013), a more understandable relationship given pinnipeds' amphibious existence and terrestrial ancestry. Pinnipeds demonstrate sensitive mechanoreception both in the active (haptic) and passive modes, which they use for detecting hydrodynamic stimuli. They can discriminate size and shape by active touch (Dehnhardt, 1994; Dehnhardt and Dücker, 1996) and detect water movement at detection thresholds under a micron of particle displacement (Dehnhardt et al., 1998; Dehnhardt and Mauck, 2008). Their high sensitivity to hydrodynamic stimuli allows both seals and sea lions to track fish by the turbulence they generate in swimming. While sea lion vibrissae appear to be more sensitive than those of phocids to relatively low frequency vibrations in the water, harbor seals have shown greater ability at following complex wakes over longer periods of time (Gläser et al., 2011), perhaps due to differences in vibrissal structure (Hanke et al., 2010; Witte, 2012).

The olfactory bulbs of pinnipeds are reduced in size, and there are fewer nasal turbinates. Nonetheless, scent recognition is a demonstrated feature of individual recognition in pinnipeds, particularly well-studied in mother-pup identification, and likely relevant for reproductive behavior in some species (Pitcher et al., 2011). Gustation has hardly been studied. There are taste buds on the tongue, albeit fewer than among terrestrial mammals. Despite the unimpressive anatomy associated with the chemical senses, sea lions detect sour, bitter, and salt (Friedl et al., 1990). They also have low discrimination thresholds for saline solutions (Sticken and Dehnhardt, 2000).

Sirenians

West Indian manatees have modest visual acuity of approximately 20 arc min (Mass et al., 1997; Bauer et al., 2003) and probably limited visual tracking capabilities (Samuelson et al., 2012). Unlike many of the cetaceans and pinnipeds studied, they lack a tapetum lucidum for enhancing light sensitivity, but also unlike them have dichromatic color vision (Griebel and Schmid, 1996; Ahnelt and Kolb, 2000; Newman and Robinson, 2006). Preliminary evidence from streak retinoscopy indicates emmetropic to hyperopic vision both underwater and in-air (Samuelson et al., 2012). They lack a vomeronasal organ and their neurophysiology suggests modest olfaction (review in Reep and Bonde, 2006). They have a higher density of taste receptors than dolphins (Yamasaki et al., 1980), but the psychophysics of taste and other chemical senses has not been investigated. Auditory capabilities include about an eight-octave frequency range extending from about 0.25 kHz into the ultrasonic range over 70 kHz (Gaspard et al., 2012), a high temporal processing rate (Mann et al., 2005), and good sound localization (Colbert-Luke et al., 2015). Manatees are the only mammal known to have exclusively sensory hairs (vibrissae) covering their entire body. Manatees' sense of touch is highly sensitive with Weber fractions between 0.025 and 0.14 (Bachteler and Dehnhardt, 1999; Bauer et al., 2012). At low frequencies, they can detect hydrodynamic particle movement under a micron with an order of magnitude greater sensitivity rostrally (Gaspard et al., 2013, 2017).

Although formal, behavioral experiments have not been done, the sensitivity and resolution of the manatee senses of hearing and touch suggest the ability to discriminate fine detail, which might allow for orientation by auditory and tactile scene analysis. Masking experiments reveal enhanced hearing in noise as indicated by low critical ratios, especially within the range of the second and third harmonic (Gaspard et al., 2012), which in conjunction with field studies identifying signature vocalizations, suggest that manatees might acoustically differentiate among individuals (Sousa-Lima et al., 2002). Although the physiology of chemoreception is unimpressive, naturalistic observations of tracking estrus females and locating fresh water in a saltwater milieu suggest that chemical senses might be more prominent than expected.

In summary, the three orders of marine mammals display visual modifications appropriate for maintaining an adaptive level of visual acuity in underwater and in-air environments. Their sense of hearing allows detailed temporal perception, exquisite in the case of echolocating cetaceans. The active touch sense facilitates fine textural discrimination in pinnipeds and manatees. In the passive touch mode, harbor seals and sea lions can follow the trail of residual turbulence left by swimming fish. The sensitive mechanosensory systems of manatees and dolphins are likely to be similarly sensitive to water movement. The chemical senses remain to be explored more thoroughly.

One way to think about many of these sensory characteristics (e.g., high frequency hearing) are as adaptations for particular niches. Byrne (1995, p. 34), however, argued that viewing adaptations as intelligence adds nothing explanatory, so suggests that "intelligence" be reserved for something more restricted, a "… quality of flexibility that allows individuals to find their own solutions to problems." We agree that sensory adaptations by themselves are not intelligence, but when integrated with systems that connect senses to motor responses (c.f., von Uexküll, 1934/1957), and when these connections generate complex behavior, intelligence emerges. Furthermore, sensory systems that are multimodal can be linked by common representations, which might provide a useful avenue for considering intelligence. For example, a pit viper that uses heat as a single indicator of prey or predator and strikes at it, whether it is a mouse or a warm water-filled balloon, has a narrow perceptual world. In contrast, a cat might integrate its good visual resolution, keen sense of smell, and high frequency hearing to represent the warm object as a mouse. An *Umwelt* built at this level of complexity provides more tools for problem solving and adaptability – more opportunities to build a better mousetrap.

The marine environment promoted the development of high resolution auditory and tactile senses in marine mammals, and in the case of the former, it fostered high speed processing. These adaptations, in conjunction with good visual acuity found in many, but not all, species, facilitated a general sensory foundation for multimodal, rapid, integrated information processing. Furthermore, the selective pressures of an aquatic environment to develop general sensory systems suggest the possibility of generating richer representations and perhaps something akin to the fluid intelligence capacity described in humans (Deary et al., 2004). Fluid intelligence, per se, has not been assessed to our knowledge in marine mammals, but its correlate, working memory, has been well-investigated (e.g., Thompson and Herman, 1977, 1981).

SOCIAL LEARNING

Dim light and the efficiency of sound transmission in the underwater environment favor acoustic (and possibly tactile) communication among marine mammals. The structural characteristics of vocalizations by marine mammals are well described in the literature, but it is only among the cetaceans that we find substantial investigation of the cognitive aspects of communication, especially vocal mimicry. Dolphins also demonstrate flexible behavioral mimicry which may be unique among non-human mammals in the variety and flexibility of both vocal and behavioral copying, although these capabilities have been demonstrated to some extent in an African gray parrot (Moore, 1992), as well. Little is known of the cognitive aspects of pinniped vocalizations, although a single case study of a harbor seal that spoke several phrases in English (Ralls et al., 1985; Deacon, 1997) suggests that it is an area worth greater attention. The ability to engage in social learning not only expands avenues for gaining new information and skills, but also pushes individuals to decode the actions of conspecifics, a rich area for cognitive growth. When social learning occurs through mimicry, this decoding requires a representation of a social partner that applies in a fine-tuned way to oneself. Although some behavioral copying can be learned slowly through trial and error, consistent with Macphail's perspective, laboratory evidence of rapid acquisition, including single trial learning, suggests more efficient mechanisms.

Cetaceans

Herman (1980, 2002) and Whitehead and Rendell (2015) provide several, thorough reviews of both vocal and behavioral mimicry. Therefore, in this section we will provide brief summaries of research previously reviewed and updates of more recent literature.

Vocal

Marine mammals show remarkable flexibility in vocal copying, e.g., with human-like spontaneous vocalizations in beluga whales (*Delphinapterous leucas*; Ridgway et al., 2012) and dolphin-like vocalizations by killer whales (*Orcinus orca*) who had dolphin pool-mates (Musser et al., 2014). Wild social groups of killer whales share call types (Ford, 1991). Young dolphins born in human environments incorporate trainer's whistles into their repertoires (Miksis et al., 2002). Dolphins also spontaneously mimic computer-generated sounds (Herman, 1980; Richards et al., 1984; Richards, 1986; Reiss and McCowan, 1993), both narrow and broadband. Dolphins naturally copy each other's identifying whistles, individually distinctive signature whistles that serve as contact signals (Caldwell and Caldwell, 1965; Tyack, 1986; Caldwell et al., 1990). These whistles are learned, unique identifiers discriminable by other dolphins (Harley, 2008), and used on meeting in the wild (Janik et al., 2006). Dolphins can vocally mimic on command in controlled laboratory settings, including the sound bursts of human speech (Lilly, 1965; Lilly et al., 1968), and sine waves, similar to natural sounds, but also atypical sounds like triangular and square wave tonal patterns, sometimes going beyond copying to mimic amplitude modulations and transients at the onset of tonal stimuli, as well as transposing sounds by an octave (Richards, et al., 1984).

Vocalizations are not the only behaviors showing evidence of dolphin mimicry and perhaps other forms of social learning. Synchrony in swimming, respiration, and leaping is a common feature of wild dolphin behavior (Connor et al., 1992a, 2006b). Synchrony occurs immediately after birth (Cockcroft and Ross, 1990) as dolphin calves swim continuously (Lyamin et al., 2005; Sekiguchi et al., 2006) in the slipstreams of their mothers. Calf synchronous swimming with other dolphins in the social group appears later in development (Fellner et al., 2012). The early development of synchrony may support social learning capabilities (Whiten, 2001; Fellner et al., 2006; Hastie et al., 2006; Whitehead and Rendell, 2015) and act as a means of social affiliation (Connor et al., 2006a,b; Perelberg and Schuster, 2009) and cultural transmission of information (Bauer and Harley, 2001; Whiten, 2001; Fellner et al., 2012).

There is a rich anecdotal literature on cetacean behavioral imitation, for example, captive bottlenose dolphins (*Tursiops aduncus*) copying the sleeping posture of a Cape fur seal (*Arctocephalus pussilus*); recruiting feathers, expelling bubbles and making scuba noises to mimic human divers cleaning; and acquiring and releasing a mouthful of milk, like a smoke cloud, at smokers standing by a pool window (Tayler and Saayman, 1973). In commercial shows, a false killer whale (*Pseudorca crassidens*) learned the routines of a pilot whale and two rough-toothed dolphins by observation (no training involved; Brown et al., 1966), and a bottlenose dolphin copied the unique, spiraling leap of a spinner dolphin (*Stenella longirostris*) introduced to its tank, atypical for a bottlenose. Another example suggesting emulation of a routine occurred at Kewalo Basin Marine Mammal Laboratory (Bauer, personal observations, 1979–1980). The routine for training the dolphins for object "names" and actions included a tonal secondary reinforcer for correct behaviors and then fish reinforcement. During the sessions, the dolphins would drop fish to the tank bottom, and occasionally, for incorrect trials, the dolphin itself would whistle the secondary reinforcer and eat a stockpiled fish.

Ostensibly insightful or otherwise intelligent behavior frequently attracts human attention, despite absence of knowledge of how these behaviors developed. Often, trial and error mechanisms explain the behaviors (Macphail, 1982; Shettleworth, 2010). Here, controlled experimental studies support the anecdotal evidence highlighting the flexibility of dolphin cognition. Young bottlenose dolphins in a "Do this…" paradigm mimicked humans modeling a diverse array of behaviors, some on the first trial (Xitco, 1988; Herman, 2002), even with dramatic differences in morphology (e.g., legs vs. fluke and flipper vs. arms), which present a concordance problem. Xitco et al. (1998) later brought imitation under control of a hand signal designating mimic in a study of dolphin-dolphin imitation with two dolphins. Importantly, the model was given hand signals to do other behaviors in addition to mimic, so that mimicry was clearly under stimulus control of an arbitrary signal. The experimental design included training on a set of behaviors and testing on a set of different, untrained but familiar behaviors, and finally on a set of novel behaviors. Both untrained familiar behaviors and novel behaviors were copied, some on the first trial. Xitco also demonstrated that the dolphins could successfully copy behaviors after delays up to 80 s. The mimicry of novel behaviors met Thorpe (1963) criterion for imitation: learning a new behavior by copying. Bauer and Johnson (1994) partially replicated this study, although without demonstrating mimicry of novel behaviors. Major differences in subject experience could have easily accounted for this discrepancy.

Later, Herman et al. demonstrated that dolphins could copy a human model standing in air and a previously performed behavior (self-mimicry; Mercado et al., 1998). The experiments indicated that dolphins were responding to visual cues but left open the possibility that the dolphins might, in addition, respond to auditory and tactile (water flow) cues. Jaakola et al. (2010, 2013) demonstrated that dolphins could perform modeled behaviors even when they were wearing eyecups blocking vision, using passive listening with dolphin models and echolocating human models.

In summary, dolphins exhibit robust mimetic abilities, both vocally and behaviorally, an apparently unique combination among non-human mammals. They copy sounds of conspecifics, computer generated sounds, and qualities of human speech. They copy a rich variety of behaviors modeled by different species with different morphologies. They mimic models in water and out of water. They mimic spontaneously and under stimulus control. Their mimicry is exhibited to visual stimuli alone and to acoustic and possibly tactile stimuli. They mimic synchronously and after delay, demonstrating the persistence of the representation. All of these factors argue that dolphins have a conceptual understanding of imitation. Herman (2002, p. 100), in a review of dolphin imitation, asks:

> What does it mean to have a generalized concept of "imitate"? It implies that the capacity is not reserved or restricted to functionally significant events, or to events tied to the organism's natural repertoire, ecology, or habitat, but is broadly understood as applicable to any arbitrary experienced event. The dolphin is obviously an imitative generalist…

Ascertaining how copying behavior functions in the wild is difficult because of the problem of controlling alternative explanations of behavioral acquisition. For example, copying behaviors might reflect contagion, social facilitation, stimulus or response enhancement, observational conditioning, or matched-dependent behavior, which are expressions of already existing behaviors or behaviors easily explained by trial and error learning. These are difficult to discriminate from true imitation that requires actually learning new behaviors (reviews in Whiten and Ham, 1992; Zentall, 1996). Nevertheless, we can propose promising candidates for social learning in all its forms by looking at wild behavior.

Wild marine mammals are highly flexible foragers. Foraging techniques found in limited groups of the same species present interesting examples of cooperation that suggest social learning. For example, symbiotic fishing between humans and dolphins was reported by Pliny the Elder (~70 AD) and more recently in Australia, Brazil, Myanmar, India, and Mauritania (review in Whitehead and Rendell, 2015). Typically, dolphins herd fish toward fishermen who capture them in nets and wild dolphins capture the fish concentrated between them and humans. The origins of these cooperative fishing ventures are unknown, but the outcome appears to be beneficial for both species. There is also some evidence that California sea lions use dolphins to locate large schools of fish for predation (Bearzi, 2006). Another example is provided by small groups of sponge feeding dolphins (*T. aduncus*) in Shark Bay Western Australia, where these dolphins carry sponges on their rostrums, presumably as protection from fish spikes on the murky bottom (Mann et al., 2012). At least two unrelated subgroups share the behavior, suggesting some social learning, although there is some familial relatedness within each subgroup. These candidates for acquisition of knowledge through social learning might be explained by vertical transmission, parent to offspring. A case broadening the sources of knowledge within a group has recently been provided in a study of the unusual behavior of "shelling," also by bottlenose dolphins (*T. aduncus*) in Shark Bay. In "shelling" a dolphin drives fish into large shells, takes the shell to the surface, and then shakes the fish out into its mouth. Integrating behavioral, genetic, and environmental data, Wild et al. (2020) demonstrated that the behavior is transmitted horizontally (i.e., relationships other than parent-offspring). Both vertical and horizontal transmission of foraging behavior enhances the dispersion of knowledge and increases flexible responding.

Killer whales are apex predators feeding on a wide variety of prey (e.g., beaked whales, salmon, herring, seals, cephalopods, gentoo and chinstrap penguins, humpback whales, gray whales, gray seals, blue whales, sea turtles, minke whales, emperor penguins, elephant seals, sharks, deer, and moose). Different prey require different hunting techniques including corralling, swimming onto beaches, and collaboratively creating waves to wash prey off ice floes (Visser et al., 2008). In another social sphere, male dolphins synchronize and coordinate both vocal (Moore et al., 2020) and motor behavior to control and protect access to females (Connor et al., 2006b).

Although we do not have controlled, laboratory experimental data on baleen whale behavior, in the wild they engage in a variety of cooperative behaviors such as synchronous swimming, cooperative foraging, and memory for migratory destinations that suggest the possibility of social learning, but in most cases, instinctual responding cannot be ruled out. A notable exception is evidence that humpback whale song (*Megaptera novaeangliae*) is learned socially (for reviews of humpback whale song, see Payne, 1983; Whitehead and Rendell, 2015). Humpback males sing at the tropical/subtropical termini of their annual migrations from polar regions. Elevated testosterone and increased mating behavior in these regions suggest that the songs have reproductive functions. The songs range over seven octaves (~30 Hz to 4 kHz) and have units, phrases, and themes organized in a hierarchical structure. The units include tonal whistles and broadband sounds lasting from 0.15 to 8 s. Generally, there are fewer than 10 themes in a song, but for any one song, the order and number of themes are fixed, although the number of phrases may vary. The song, which may last from 10 to 30 min, is repeated, continuing for many hours.

Social learning of songs is indicated by several factors. Humpback whale populations are discrete, with relatively little exchange with other populations. The songs are the same for all members of a population and they change over the course of a season and between seasons (i.e., annually) by dropping or adding themes.

Furthermore, evidence from the South Pacific indicates that songs are transmitted east to west, while at the same time, there is little east to west movement by individual whales. Therefore, the movement of song represents a transfer of information, not movement of individuals (Garland et al., 2011).

Learning, remembering, and producing these complex, changing songs suggest substantial cognitive demands on male whales. Interestingly, Guinee and Payne (1988) reported that they had found multi-themed sub-phrases that formed similar beginnings and endings of adjacent phrases, a phenomenon they characterized as rhyme-like. These rhyme-like patterns were positively correlated with the number of themes (i.e., the amount of material to be remembered) but not duration, suggesting a mnemonic function like that found in human recitation of long, complex works.

Pinnipeds

Less is known regarding vocal learning in pinnipeds than in cetaceans, and there have been very few experiments probing behavioral motor imitation. However, there are emerging observations that suggest that pinniped species may demonstrate a rich range of vocal learning capabilities.

As reviewed by Reichmuth and Casey (2014), there is growing evidence – predominantly gathered from observational field studies – for vocal learning in phocid pinnipeds. This includes regional variability in vocalizations of Weddell seals, leopard seals, harbor seals, harp seals, and bearded seals and raises the possibility of social learning influencing development and production of vocalizations in the wild. Implementing the types of developmental cross-fostering studies that have illuminated vocal learning in birds using pinnipeds is logistically and ethically difficult. However, opportunistic observations of a female Northern elephant seal raised in social isolation suggest species atypical call-types, as has been observed in songbirds reared in similar situations.

The famous but singular case of Hoover the harbor seal (Ralls et al., 1985; Deacon, 1997) continues to stoke interest in vocal imitation and flexible learning in phocids. Hoover was orphaned as a pup and rescued by a fisherman, who raised him until he became too difficult to maintain. Hoover was then transferred to the New England Aquarium where he surprised staff and visitors by speaking English phrases, which included, “Hey!, Hey!, Hey!, Hey!”; “Hoova!” (Hoover with a New England accent); “Hey!, Hey!, Get outa there!,” “Hello there,” and “Come over here”; and some speech-like, but indecipherable sounds. It is not clear where and how he learned to “speak.” Based on Hoover’s accent and other factors, Deacon (1997) has suggested that he had learned speech from the fisherman. Since the origin of Hoover’s speech is unknown, we cannot determine if it was copied or learned by trial and error. Hoover is apparently unique among pinnipeds regarding the quality and specificity of his mimicry. What is clear is that for now Hoover is unique among pinnipeds in his mimicry of human speech. Laboratory research probing vocal ontogeny in phocid pinnipeds is ongoing (Ravignani, 2019; Ravignani et al., 2019).

Walruses, which are a separate clade from the phocid (true seal) and otariid (sea lion) pinnipeds, have also been suggested as potential vocal imitators, although more data are needed. In the wild, adult male walruses have been shown to alter song types substantially over subsequent breeding seasons, much as humpback whales do (Sjare et al., 2003).

There are scant data on behavioral motor imitation in pinnipeds, but there is reason to further explore their capabilities. While phocid pinnipeds are typically weaned very rapidly (within a month of birth) and do not have extensive social interactions during development, most of the otariid pinnipeds spend far longer with their mothers prior to weaning, up to 2 years for Steller sea lions (Trites et al., 2006). Walrus pups may spend even longer with their mothers, up to 5 years in some cases (Fay, 1982). The young of most otariid species live in large, crowded, hyper-social rookeries, where they spend much of their day engaging in play behavior with other young. Play is a rich context for social learning, and, indeed, there is some evidence of social learning during Steller pup play (Gentry, 1974). This extended weaning period, during which otariid young achieve significant mastery of open ocean swimming well before they need to hunt on their own, may also allow a period of social observational learning related to hunting behavior of adults. Fur seal pups have been observed overlapping with hunting adults months before they begin hunting on their own (Lowther and Goldsworthy, 2012). The apparent vocal flexibility of phocid pinnipeds, and the extended juvenile period and active play of otariid and odobenid pinnipeds, provide reason to further probe social learning and imitation in pinniped species.

The apparent profusion of social learning and mimicry across cetaceans and pinnipeds is noteworthy, given the frequent difficulty of proving these abilities in laboratory studies with terrestrial mammals. Social learning among sirenians has not been reported to our knowledge. The ecology of marine mammals has generally favored long lives and large group size, both of which may privilege accumulation of social learning across the lifespan. Though each instance of such apparent learning must be investigated, cognitive flexibility is broadly believed to support such rapid and variable learning.

FLEXIBLE COGNITIVE PROCESSING

Although some cognitive abilities found in marine mammals were modified to adapt to the aquatic environment, many useful attributes were no doubt conserved in the transition from a terrestrial environment. Cetaceans display flexibility across a broad array of learning, memory, and problem solving tasks (reviews in Herman, 1980, 1986; Marino et al., 2007; Mercado and DeLong, 2010; Pack, 2015; Harley and Bauer, 2017), as do pinnipeds (reviewed in Schusterman et al., 2002; Cook et al., 2020).

Cetaceans

Many odontocete species, e.g., *T. truncatus*, *Delphinus delphis*, *Phocoena phocoena*, *Inia geoffrensis*, and *Lagenorhynchus obliquidens*, display basic discrimination learning abilities, frequently exhibited in studies of sensory detection discrimination thresholds, particularly perceptual systems in the auditory and visual domains (reviews in Nachtigall, 1980, and Au, 1993, of echolocation discrimination learning). Cetaceans have been tested broadly on other cognitive tasks showing that they are proficient at abstract rule learning. Dolphins demonstrated facility in auditory learning sets using hundreds of novel pairs and in reversals of the same pairs (Herman and Arbeit, 1973; reviews in Herman, 1980, 1986; Herman et al., 1993). Both procedures require learning a win stay/lose shift rule. Early efforts with training visual stimuli were not successful (Herman et al., 1969) suggesting a bias toward audition, a possible adaptation to an aquatic environment, but later work suggested that dolphins were capable in both domains. For example, dolphins tested with auditory (echolocated) and visual 2d and 3d planar stimuli successfully solved same/different discrimination problems (Mercado et al., 2000). They demonstrated generalization of the concept by correctly classifying pairs of novel targets in air on the first trial and then transferring this ability to unfamiliar targets presented underwater. They also transferred the same/difference concept from pairs of objects to objects presented in groups of three, in which "same" was represented by three identical objects and different by two identical objects and one different object.

Retention of information by cetaceans has been tested using a variety of short-term and long-term memory procedures. Many of the memory findings are broadly found among species, but it is nevertheless important to establish similarities in intelligent behavior, as well as differences. Dolphins do well on tests of short-term working memory, typically assessed in matching tasks in which a sample stimulus is presented followed by a recognition test in which two or more stimuli, one of which matches the sample, are presented after a delay. Dolphins in artificial "language" testing also performed what were essentially conditional matching-to-sample problems (Herman et al., 1984), in which the sample stimulus was symbolically represented in the test as a sound or hand-sign (A to A') that was paired with an object choice presented among object alternatives (A', B', C', etc.). Dolphins also showed a recency effect for lists of sounds (Thompson and Herman, 1977) and good memory for relative spatial positions (Herman et al., 1984). Long-term memory has not been well-studied. In an investigation of captive dolphins using a habituation-dishabituation design, subjects apparently remembered signature whistles over a period of 15 years (Bruck, 2013). Memory for signature whistles would be an adaptive characteristic for long-lived dolphins living in fission-fusion societies in the wild.

In a creative use of memory, dolphins can acquire an "innovate" rule by correctly doing a self-selected new behavior when signaled to do so. Pryor et al. (1969) reported an early instance of this rule with rough-toothed dolphins (*Steno bredanensis*). On command, the subjects were reinforced for executing a behavior not previously done in the innovation training sessions. The experiment was terminated after 16 reinforced innovative behaviors, when trainers found it hard to discriminate between novel and familiar behaviors. In similar studies, trainers brought novel or not recently performed behaviors under stimulus control (reviews in Kuczaj and Yeater, 2006; Mercado and DeLong, 2010). Difficulty in ascertaining novel behaviors in long-term projects may make "rare behaviors" a better term. Under any circumstances, the "innovate" procedures required learning the opposite of the learning set rule; a win shift strategy is required.

More evidence of dolphins' flexible cognitive powers occurred at the Kewalo Basin Marine Mammal Laboratory where two dolphins learned an artificial "language," in which experimenter-created "words" were presented to the dolphins by hand signals or arbitrary computer-generated sounds. The dolphins successfully learned signals representing objects, actions, and modifiers. They also learned that the order of "words" in language-like sequences could indicate different actions. For example, the sequence "hoop pipe fetch," meaning take the pipe to the hoop, required a different action than "pipe hoop fetch" meaning take the hoop to the pipe. Five-word sentences were created by adding relational modifiers (e.g., surface vs. bottom and right vs. left). Evidence that the dolphins learned a specific grammar, not simply memorizing specific sequences, was indicated by the fact that they could correctly follow behavioral instructions when novel terms were introduced. There is some disagreement concerning whether these dolphins actually displayed language-like learning with syntax (Herman, 1988, 1989) or an associative process (Schusterman and Gisiner, 1988). In either case, the behaviors displayed by these dolphins indicated flexible, complex, sequence rule learning. The "language" also allowed testing of a variety of concepts: presence vs. absence of objects, identification of body parts, memory for action events such as repetitions after delays of behaviors, combinations of behaviors, and actions on a specific object in the presence of many other objects (Mercado and Delong, 2010).

Dolphins have also learned sequences in other contexts. They can recognize relative number magnitudes in ordered sequences and novel melody sequences as ascending or descending (review in Pack, 2015). They can recognize specific rhythms, transfer them across frequency and tempo shifts, and produce them (Harley et al., 2005). In a task in which they were originally trained to produce rhythms using a pneumatic switch, they spontaneously transferred the rhythms to vocalizations. The transfer suggests abstract representation of the rhythm and/or ability to copy a tonal rhythm.

Dolphins interpret and produce referential pointing gestures when engaging with another species. They follow referential human points (i.e., pointing to an object; Herman, et al., 1999), as well as using pointing gestures themselves to direct humans (Xitco et al., 2001a,b). In these studies, dolphins pointed with their rostrums at fish in jars placed at various locations in a captive habitat in the presence of humans but not in their absence. The dolphins would also engage in joint attention behaviors by turning toward human swimmers and pointing back at the jars. The humans responded by opening the jars.

The flexible referential quality of dolphin cognition, in contrast to perception of the simple, physical stimulus character of objects, is illustrated in cross-modal experiments. Harley et al. (1996, 2003) and Pack and colleagues (Pack and Herman, 1995; Herman et al., 1998) trained dolphins to identify objects in one modality (vision) and then identify them in a different one (echolocation) and vice versa using a matching-to-sample format. For example, a dolphin wearing eyecups to block vision investigated an object echoically but then successfully matched that object to an identical alternative presented visually in air where dolphin echolocation does not work. Since the visual and echoic (hearing) experiences are obviously physically different, the dolphin had to represent the stimuli in such a way as to allow recognition in either modality. Dolphins clearly have a plastic hierarchical object representation system that includes attributes gleaned through multiple, high-resolution sensory modalities.

Flexibility is also required to identify objects using echolocation alone in that the echoes from different aspects of a single object can vary more than those between different objects (review in Harley and DeLong, 2008). For objects that vary only in size, like different-sized disks, dolphins can use differences in amplitude to discriminate among the disks. For objects that vary only in material, like an aluminum vs. a stainless-steel cylinder, dolphins may use different pitches to tell the objects apart. Once objects get more complex and vary across many features simultaneously, it is more difficult to know how the dolphins manage recognition tasks because echoes from different attributes interact, but they can do it. This same quality of the elasticity of dolphin cognition is evident in the interpretation of shared echoes. Xitco and Roitblat (1996), using a three-choice delayed matching-to-sample task, demonstrated that a dolphin who had only heard the echoes returning from a sample object to its neighbor could choose the identical alternative at above-chance levels. The import of this ability to share information directly is that it may allow a group of dolphins to act as a sensory integration unit (Norris and Johnson, 1994), surpassing the experience of any individual.

Anecdotal evidence, corroborated by experimental work, suggests cetacean planning abilities. For example, killer whale foraging and dolphin sponge fishing discussed earlier indicate some preparation. Two experimental studies provided more easily verifiable evidence of planning (review in Kuczaj and Walker, 2006). In the first experiment, dolphins learned a task in which four weights placed in a device within a given timespan released a fish. Dolphins learned the task by observing a human placing the weights one at a time and then executed the task in the same way when weights were close to the device. When weights were further away, the dolphins switched strategies to carry multiple weights to the device, a more efficient approach suggesting planning. In the second experiment, three separate devices released a fish when one weight was inserted. Two of the three devices allowed the weight to pass through, so it could be reused. The third did not release the weight, preventing the dolphin from getting fish from the other devices. The best strategy, therefore, was to select this device last, which the dolphins learned to do.

Pinnipeds

Although behavioral and cognitive studies of pinnipeds have featured small sample sizes, and have focused predominantly on California sea lions, impressive results have been obtained in the realm of language learning, memory, concept formation, and rhythmic capability. As with the cetaceans, the number of apparently unique and rare abilities observed in pinnipeds is striking given how few studies with how few subjects there have been.

While studies examining human-like language learning in animals have mostly featured apes and cetaceans, there were a series of studies in the 80s and 90s with sea lions. As reviewed in Schusterman and Gisiner (1997), several sea lions, having learned to respond to gestures indicating objects (e.g., cones and balls), descriptors (e.g., large and white), and actions (e.g., fetch and bring to), responded appropriately to novel combinations of those gestures (e.g., bring the small black ball to the large white cone). This suggests something akin to receptive syntax, which has been shown in very few non-human species.

Further studies at the same laboratory probed the ability of sea lions to group arbitrary stimuli into concept classes and then to use logical reasoning to add new stimuli into one class or another with one-trial learning. For example, having learned that A, B, and C go together, and 1, 2, and 3 go together, sea lions were able to add D to the correct class following one exposure. In other words, if D goes with A, B and C go with A, D must also go with B and C. Further, because A, B, and C do not go with 1, 2, and 3, if D goes with A then D does not go with 1, 2, and 3. This represents a type of transitive inference rarely demonstrated in non-human animals. Impressively, sea lions have demonstrated robust long-term (10+ years) memory not just for the stimuli involved in these experiments but for their logical relations to each other (Schusterman et al., 2002).

Sea lions are also unique among non-human animals for having shown the ability to move in time to a musical rhythm and then to generalize it to novel stimuli and tempi in transfer tests (Cook et al., 2013). This capacity was previously believed unique to humans until demonstrated in some parrot-type birds (Patel et al., 2009), leading to the theory that brain circuits involved in complex vocal production learning were necessary to learn to match movement timing to complex auditory rhythms as in human dance. Sea lions, who show very limited vocal learning, challenge this theory. It remains to be seen how widely this faculty is distributed. It may be that, as a motivated animal with strong motor control, sea lions have an easier time demonstrating certain complex sensory-motor behaviors than many other species (Wilson and Cook, 2016).

A patchwork of studies over the last 15 years has probed a number of "higher" cognitive abilities in sea lions related to self-control, working memory, and mental manipulation of representation. While more work is needed, sea lions have shown strong inhibition of pre-potent motor responses, besting primates in their capacity to inhibit selection of a lesser reward for later receipt of a greater (Genty and Roeder, 2006; though see Beran and Hopkins, 2018). Sea lions can also mentally rotate shapes in matching tasks. While orientation-invariant matching is not rare in tested animals, sea lions are unusual in that their response times scale with the degree of mismatch between the stimuli and their comparisons (Mauck and Dehnhardt, 1997; Stich et al., 2003). One explanation is that, as humans are believed to do, they are actually rotating a mental representation in working memory. They have also demonstrated the ability to locate objects based solely on mirrored visual representations (Hill et al., 2015). In addition, they show a capability to follow ostensive pointing gestures with high success, potentially relating to an ability to decouple a local visual stimulus from its immediate surroundings (Scheumann and Call, 2004). Finally, South American sea lions have been shown to have primate-like capabilities for discriminating stimuli based on numerosity, a skill generally believed reliant on some degree of working memory function (Abramson et al., 2011). Each of these abilities could be considered to be related to "executive function," a general set of neurobehavioral processes relying on prefrontal and parietal associative "control" regions in humans and primates.

Sirenians

Sirenians, primarily manatees, display basic discrimination learning abilities in studies of sensory detection and discrimination thresholds in the tactile, auditory, and visual domains (review in Bauer and Reep, 2017). There has been no formal research on long-term memory in manatees. Anecdotal evidence from Florida manatees in the wild is suggestive. Reep and Bonde (2006) report that manatees recall the location of freshwater hoses between seasons. In a captive situation, two manatees remembered an active touch discrimination of textures procedure with 100% performance accuracy after 14 and 29 months, respectively (Bauer et al., 2012). Cognitive investigations limited primarily to response shaping and discrimination learning does not provide an adequate basis for conclusions characterizing sirenian intelligence.

Cetaceans and pinnipeds display a wide range of cognitive abilities. Perhaps the issue is not so much their ability at any one of the procedures on which they have been tested, many of which have been displayed by diverse other species, but in the range of abilities demonstrated. Of significance for intelligence is the complexity of representation and the transfer of complex skill sets across contexts highlighting flexible intelligence, e.g., in cross-modal tasks. It is also striking that a relatively small number of marine mammal subjects has demonstrated such an expansive list of abilities.

THE BRAIN

To the extent that conceptions of intelligence rely on association, sensory processing, representation, and manipulation of information, intelligence can be understood to be a general feature of the nervous system, or, at least, a general potential for the nervous system to produce certain outcomes in different environmental contexts. Human neuroscience, bolstered by functional neuroimaging technology, has done much to unravel the neurobiological mechanisms undergirding human cognition. We now have a strong understanding of which brain regions represent sensory information, which brain regions code motor behavior, and the relation between these sets of regions that allow us to respond to our environment (Power et al., 2011). These primary brain regions, directly connected to body sensors and effectors, are evolutionarily conserved and provide the foundation for the brain's higher processes. We further have delved into how non-primary "association" regions in the brain, with no direct connection to body sensors or effectors, work to regulate, control, and manipulate primary brain regions to support complex cognition (Goldman-Rakic, 1988). The human brain is composed of parallel hierarchies of motor and sensory processing (Fuster, 1997). The primary motor and sensory regions are cortical brain regions directly connected to body sensors and effectors. These areas share information with, and are regulated by, secondary cortical brain regions with no direct connections to body effectors, the premotor cortex and unimodal association areas, respectively. These secondary regions in turn share information with and are regulated by tertiary regions that influence the secondary regions, and, typically, through those secondary regions, the primary regions connected to the body. These are the prefrontal cortex in the motor hierarchy and the polymodal association cortex in the sensory hierarchy, and they can be thought of as sitting atop the neural hierarchy, exerting disproportionate control over the other brain regions. The influence of prefrontal and polymodal association areas is strongly correlated with "higher" cognitive function in humans, allowing the formation, maintenance, storage, and manipulation of complex representations (Yeo et al., 2015).

In well understood examples drawn from human neuroscience, primary sensory regions can be recruited by the hippocampus and prefrontal cortex to support experiential memory (Preston and Eichenbaum, 2013). Motor regions can be inhibited by frontal control regions to stop immediate response to stimuli, opening up time for slower, more deliberate responses and planning (Ridderinkhof et al., 2004). Subcortical regions processing reward can be activated in concert with memory supporting and motor control regions to support complex learning and planning based on prior and simulated outcomes (Pasupathy and Miller, 2005). In these and essentially all other circumstances of higher cognition in humans, our current neuroscientific understanding relies on *connectivity* (see Rubinov and Sporns, 2010). Brain regions influence each other through connections, the patterns of these connections are a map of potential interactions and thus potential neurobehavioral outcomes, and the dynamic interaction and plasticity of these regions and their connections support complex and changing behavior across a range of situations.

Comparative neuroscience now also increasingly operates on a connectionist framework (Mars et al., 2016, 2018), and studies of rodents and primates seek to find the similarities and differences related to network connectivity in humans, in order to better understand the functional relevance of these connections and how they influence behavior, both typical and atypical as in disease states. Marine mammals have long been of interest to comparative neuroscientists for a number of reasons, but the bulk of interest has been driven by the grossest features of their neurobiology. First, size – marine mammal brains are large in comparison to those of terrestrial animals, both in absolute terms, but also, for some species such as dolphins, in relation to body size (Marino, 1998). Second, gyrification – the pattern of folds (including bumps, or gyri, and grooves, or sulci). Pinnipeds and cetaceans have remarkably folded brains in comparison to terrestrial mammals, while sirenians have remarkably smooth (lissencephalic) brains (Reidenberg, 2007). Most research on marine mammal brains to date has addressed two general features – size and wrinkliness, and we will briefly review that literature below. A number of researchers have suggested that the large size of cetacean brains, when viewed from the perspective of their often impressive behavior in the laboratory, is a clear indicator of extreme intelligence (Marino et al., 2007). Fewer hypotheses have been advanced regarding the functional relevance of gyrification, and recent research suggests gyrification is predominantly a product of brain size and neuron proliferation early in development (Mota and Herculano-Houzel, 2015).

It must be emphasized that, from the perspective of modern neuroscience, while size does matter, this is mostly as it relates to the number of neuron units (Herculano-Houzel, 2009). Neurons are the general information processing unit of the nervous system (Shepherd, 2015). More neurons mean the potential for more processing power. It turns out that the correlation between brain size and neuron number across species is, while present, fairly variable (Herculano-Houzel et al., 2014). Further, the number of neurons, while important, is no more (and possibly less) important than the patterns of connections between those neurons and the regions they compose (e.g., in humans, Emerson and Cantlon, 2012; Xiang et al., 2012; Xiao et al., 2018). Here, research into marine mammal neurobiology is still in its infancy. We discuss preliminary efforts to characterize marine mammal brains from the perspective of functional processing, and suggest some potentially fruitful and achievable future directions that will better enable us to understand in which ways marine mammal nervous systems are like and unlike those of their terrestrial relatives. In line with the Jamesian principle that cognition and intelligence are reliant on features of neurobiology, this may help support a framework for assessing the general intelligence of these species.

Marine Mammal Brain Size

On the topic of sheer size, marine mammals are notable for featuring the species with earth's largest brain, the sperm whale. Weighing in at up to 8 kg (Povinelli et al., 2014), it dwarfs the human's 1.3 kg brain. In addition, marine mammals include two of the four clades in the "over 700 g club" (Manger et al., 2013) with numerous whales, and four separate pinniped species besting this brain weight. Generally speaking, animals with bigger bodies have bigger brains, so it is perhaps not surprising that, freed from the constraints of gravity in a terrestrial environment, marine mammals evolved bigger bodies, and, thus, bigger brains. The terrestrial members of the club are apes and elephants. Apes may seem an exception compared to other club members in the relatively small size of their bodies. Indeed, apes are particularly notable for their "encephalization quotient (EQ)," a measure comparing brain-to-body-size ratio against the typical cross-species trend (Jerison, 1977). Some have suggested that EQ can serve as a predictor of a species' intelligence (Jerison, 1985). The human EQ is up to seven times what would be expected for their body size. Some of the toothed whales, particularly dolphin species, also have very high EQs, up to four times what would be expected based on body size (Marino et al., 2004). Despite their large brains, baleen whales and pinnipeds suffer on EQ measures, tending to fall close to the average brain size predicted by body size (Worthy and Hickie, 1986). Just as an unusually large brain increases EQ, so does an unusually large body decrease EQ. This may be the case with manatees. The attributions for sirenian intelligence based on brain size may suffer from the simplistic view that the small relative brain/body ratio of manatees and dugongs (Jerison, 1973) implies a dim intellect. O'Shea and Reep (1990) argue that this is a misrepresentation that does not take into consideration ecological (herbivory) and physiological (heat conservation) pressures driving large body size. That is, sirenians do not have excessively small brains, rather they have disproportionately large bodies.

Reliance on EQ as a predictor of intelligence has faded, with some evidence suggesting that, within related clades of animals, overall brain size is a stronger predictor of cognitive capability than EQ (Deaner et al., 2007). This returns us to the importance of overall brain size, but, again, the size of the brain is most notable as it predicts neuron count (processing power). Just because a brain has evolved to be bigger does not mean it will have more neurons. In fact, animals with bigger bodies tend to have less dense "neuronal packing." For example, some bird species have as many neurons packed into their forebrains (analogous to mammalian cortex) as some primate species (Olkowicz et al., 2016). A gray parrot's brain weighs no more than 20 g, while a lion's brain is over 10 times that size. But the gray parrot has twice the cortical neuron count of a lion. These birds have a much higher measure of processing power per unit of brain volume than do mammals. Obtaining neuron counts used to be prohibitively time consuming, but newer methods allow much more rapid counting (von Bartheld et al., 2016). Notably, the current record holder for total number of cortical neurons is the killer whale, with over 40 billion cortical neurons (Ridgway et al., 2019). The pilot whale is a close second, with 32 billion (Mortensen et al., 2014), twice what humans average at 16 billion (Herculano-Houzel et al., 2015). Most other cetaceans measured, including dolphins, have 10–12 billion cortical neurons, slightly more than the 7–10 billion found in non-human ape species (Herculano-Houzel, 2019). Pinnipeds have fewer, although still relatively high numbers compared to terrestrial mammals. Elephant seals and walruses, the biggest pinnipeds, have in the range of 4 billion cortical neurons. Contrast this to a horse, with approximately 1 billion (Haug, 1987), or a dog, with 500 million (Jardim-Messeder et al., 2017). These numbers help contextualize prior debates about the relevance of cetacean brain size. While a controversial hypothesis (Manger, 2006) has suggested that cetacean brain size is largely due to thermoregulation, the high neuronal cell counts better match other theories emphasizing cetacean cognitive capability (Marino et al., 2007) – not because brains are bigger, but because we now know they likely have more processing power.

However, it is not just the number of neurons that matters, it is where they are, and how they are connected. More work is needed to obtain neuron counts from specific structures in marine mammal brains. For example, much has been made of the small gross volume of the dolphin hippocampus, a region involved in explicit memory processes in mammals (Oelschläger, 2008). Cell counts of different regions will provide a better indicator of those regions' importance and functional capabilities. It does appear that toothed whales have densely packed brains with many cortical neurons. But what regions are those neurons in? Great expansion of primary processing regions, as seen in cortical enlargement of motor control systems in human hand cortex and visual cortex in primates (Krubitzer, 2007; Kaas, 2008), can increase brain size, relative brain size, and total cell count but may have quite different relevance for assessing global intelligence than parallel increases in association cortices and other brain regions "higher" in the neural processing hierarchy.

As noted, it is this pattern of connections that allows a region to engage in a specific function. Indeed, regional definition depends predominantly on connection profile. The cortical region receiving the bulk of direct projections from auditory receiving structures will be the primary auditory cortex, regardless of where it is in the brain. Contemporary neuroscientists believe these patterns of connection represent the possibility space for a nervous system. Everything that a nervous system can do, including information processing of the sort we tend to consider "cognitive" or intelligent (memory, decision making, self-control, and learning) relies on communication across specific connections between different brain and body regions.

Human neuroscience has placed particular emphasis on corticocortical connections – communication pathways between different cortical areas that support dynamic and flexible information processing. Here, as in the literature on size, whale and dolphin brains have received the most attention among marine mammals. Early histological work characterized the whale cortex as "primitive," meaning similar in some ways to non-placental mammals like monotremes and marsupials, taken to be emblematic of early mammal neurobiology (Morgane et al., 1985). Unlike most extant terrestrial mammals, whale cortex has five instead of six discrete cell layers. Ancestral mammals were believed to have five, while the vast majority of extant species have six, suggesting whales lost a layer somewhere in their evolutionary history after returning to the water (Barbas and Rempel-Clower, 1997).

The missing layer, cortical layer 4, is essential for connecting distributed cortical regions in terrestrial mammals (Dantzker and Callaway, 2000), and its absence, in addition to the sparse cross-hemispheric connections in cetaceans, has been taken as evidence for generally low corticocortical connectivity in the whales and dolphins. Importantly, cross-hemispheric connectivity may be reduced in part to allow for unihemispheric sleep (Tarpley and Ridgway, 1994). More recent histological examination of whale and dolphin cortex has indicated unusual patterns of dense local connectivity (Hof et al., 2005). In addition, whales do have some features associated with complex long-distance brain connectivity, such as giant "spindle" neurons also found in elephants and primates (Butti et al., 2013; Raghanti et al., 2015). Hof et al. have suggested that whale brains are not under-connected but, rather, *differently* connected. What the cognitive ramifications of this altered connectivity might be remains to be determined.

Decoding and interpreting the patterns of connectivity in whale brains will require identification of functional processing regions – as stated, it is the connection between these regions that forms the basis of brain architecture. Traditionally, neuroscientists have conducted careful cell staining studies (histology) to characterize different neural populations associated with different processing regions. One of the potential mysteries of cetacean neurobiology is the apparent lack of differentiation in cortical cell type across regions (Morgane et al., 1980; although see again Hof et al., 2005), frustrating attempts to localize functional processing regions by cell type. There have been fewer attempts to conduct these types of analyses in pinnipeds, but recent studies have successfully delineated somatosensory and visual cortex in pinniped species (Sawyer et al., 2016; Turner et al., 2017). Pinniped somatosensory cortex is large, well developed and has a high proportion of cells involved in receiving and processing touch signals from vibrissae (whiskers). When possible, these types of histological analyses can speak to the characteristics of primary sensory and motor regions and may help determine the volume and type of information processing these regions can afford.

Functional brain regions can also be identified *via* tracing studies. Historically, tracing has been conducted with chemical agents that are injected directly into a brain region, and then transported (forward or backwards along axonal connections, depending on the agent used) to connected regions (Oztas, 2003). These injections are administered to a live animal that is then killed, the brain removed and sectioned to find transport sites. Such work is no longer conducted in marine mammal species for ethical and regulatory reasons, but early work with cetaceans did seek to identify cortical processing regions for auditory information (obviously of interest given complex vocal communication and echolocation in many cetacean species). These studies indicated that primary auditory cortex was in the dorsal posterior portion of dolphin cortex, in or adjacent to the cortical regions where primary visual processing typically occurs in mammal brains (Sokolov et al., 1972; Popov et al., 1986). These studies have typically indicated reduced or absent association cortex separating these primary processing regions, which would suggest a very unusual overall pattern of brain organization, potentially relevant to how cetaceans process and integrate echoic and visual signals.

While transport tracing is no longer plausible for use in marine mammals, a set of non-invasive neuroimaging techniques can provide similar information about connection between different brain regions. Diffusion tensor imaging (DTI) is an application of magnetic resonance imaging (MRI) technology, relying on determination of direction and magnitude of water movement in the brain (Le Bihan et al., 2001). Water moves most reliably and easily along large axons, which form the primary pathway for neural communication in the brain. Thus, DTI can provide a map of the structural connections in the brain. These images can be acquired from live animals (although this requires anesthesia, which comes with risks, particularly in some of the marine mammals with non-obligate breath control). They can also be obtained from dead brains. If the images are acquired soon after the brain is removed (e.g., following planned euthanasia or a stranding death), the images can be as good or better than those obtained in live brains (Seehaus et al., 2015). Recent applications of post-mortem DTI have yielded tantalizing new information about the connectivity of dolphin auditory systems. Berns et al. (2015) traced connections from the inferior colliculus in the dolphin midbrain, the primary midbrain waystation of ascending auditory information, and found strong projections to the superior temporal lobe, significantly less dorsal than transport tracing studies have indicated the location of primary auditory cortex to be, and in line with primary auditory projections in terrestrial mammals. In addition, researchers have begun to map auditory-motor pathways in dolphin brains that may be analogous to the arcuate connections supporting vocal learning in humans (Wright et al., 2018).

These techniques are now being applied to pinniped brains in an effort to map out auditory-motor connections (relevant to ongoing debate over vocal learning capabilities in pinniped species). They have also been used to map specific neurological damage in wild sea lions exposed to algal toxins (Cook et al., 2018).

Mapping connection patterns in marine mammal brains will help us understand the functional architecture of these brains and determine to what extent it differs from that of terrestrial mammals. For example, if we can delineate the dolphin auditory cortex based on patterns of connectivity with lower brain regions, we can begin to determine to what extent auditory expansion accounts for overall brain expansion. More importantly, we can assess whether the patterns of connections with auditory regions support the complex, multi-region hierarchical processing we associate with higher cognition in humans. For example, the apparent lack of dolphin frontal cortex (cortical regions anterior to motor regions) has been commented on frequently in the literature. Berns et al. (2015) used projections from basal ganglia regions to map out brain regions functionally analogous to prefrontal regions in humans and found that they largely paralleled those observed in other species, although the gross location of regions was somewhat more lateral. Anecdotal assessments of corticocortical connectivity in pinnipeds (high) and manatees (low) may also lead to quantifications related to the capabilities of those species and can be used to assess potential functional relevance of gyrification patterns. The density and patterns of these connections in marine mammals, and how they compare to those in humans and other terrestrial mammals, will provide a biological framework for considering behavioral measures associated with intelligence and flexible cognition. In addition, by collecting neurobiological data from more individuals, we can begin to assess inter-individual variability in brain organization, which should bear directly on individual differences in cognition and behavior.

DISCUSSION

We suggest that the transition from a terrestrial to a marine environment encouraged an emphasis on high-resolution auditory and tactile senses, while reducing the importance of visual and chemical modalities. The high resolution of hearing and touch promoted stimulus discrimination capabilities. Furthermore, the high speed of sound in water required faster information processing as reflected in high temporal resolution, rapid sound integration, and good sound localization. The absence of solid physical structures for hiding from predators and prey in three-dimensional aquatic space facilitated the development of social grouping for many marine mammals for defense and foraging. Group living, in turn, fostered the development of an array of social learning skills, particularly mimetic behaviors, unsurpassed by other mammals other than humans. Marine mammals also demonstrate a wide array of other flexible cognitive capacities, perhaps surprising and notable given the relatively small number of animals tested a small number of times. What does this tell us about marine mammal intelligence?

The diversity of definitions or characterizations of "intelligence" makes this a difficult question to answer. Macphail's (1982) characterization, which limits animal intelligence to associative processes, simply does not capture the way "intelligence" is used by the lay public or professional researchers of human intelligence. Mackintosh (1998) notes that the types of associative processes Macphail describes are largely subsumed under implicit learning in human research and are not assessed on intelligence tests. That is, human intelligence research segregates intelligence from the very characteristics that Macphail characterizes as animal intelligence, a difficult position to reconcile with an evolutionary perspective. In many ways, "intelligence" seems to be a folk psychology term that maps poorly on natural psychological and biological processes, and therefore, lends itself to a wide range of often-inconsistent interpretations.

In the absence of a generally agreed upon theory of intelligence, we ought to ask what approaches lead to fruitful lines of inquiry, inquiries that might lead to theory development. Researchers investigating questions of marine mammal behavior have generally found it productive to address brain-behavior relationships, ecological adaptations, and comparisons among species. Early efforts to study cetacean cognition engaged in an overly simplistic attempt to confirm a speculative hypothesis that dolphins and whales, because of their large brains, must be highly intelligent (Lilly, 1967). This approach was given some credibility by the fact that some species of cetaceans had large brain-to-body relationships (Jerison, 1973) and that the largest of the toothed whales, the sperm whale, had the largest brain in terms of absolute size of all animals (Oelschlager and Kemp, 1999; Povinelli et al., 2014). While it is true that, within humans, larger brains have been correlated with higher measures on intelligence tests (Lee et al., 2019), this may be due to the correlation between brain size and neuron number within a species. Across species, the correlation between number of neurons and brain size is only moderate, as "neuronal packing" density can differ greatly. Adherence to the big brain-high intelligence hypothesis has yielded to more fine-tuned approaches emphasizing structure, organization, and function of brains. Cell counts suggest that some marine mammal brains are unique in terms of the sheer number of neurons they contain, e.g., killer whales and pilot whales have more than twice the number of cortical neurons than humans do. We have begun to explore overall patterns of histology and connectivity to identify different brain regions and map the patterns of connections between them with the goal of determining functionality. These data are getting easier to acquire and manipulate and could lead to specific hypotheses concerning what types of information processing are strengths of these animals. Hof et al. (2005) have suggested that in addition to large brains, cetaceans have unique patterns of cortical connectivity that, although different in structure from that of terrestrial mammals, may allow for formation of complex associations and manipulation of complex representations. While connectivity has been linked with variability in intelligence in humans (Song et al., 2008), it is important to note that at the most simple level, brain connections allow associations between different regions. Thus, differences in connectivity alone are not enough to refute an account such as Macphail's, where conserved associative learning is broadly shared across species. That said, different connection patterns may allow vastly different behavioral and cognitive outcomes, so must be considered. Although considerations of brain-behavior relationships in marine mammals are still emerging, we nevertheless think consideration of brain structure and function as opposed to mere size is the appropriate approach for better understanding cognitive/behavioral attributes.

The benefits of an ecological approach to intelligence can be seen most clearly in the sensory realm, where marine mammals demonstrate acute sensitivity and discrimination of sound and tactile stimuli. The capacities of the visual senses are more varied among marine mammal species, although for many their acuity is also quite good. The resolving capabilities of marine mammals provide a rich *Umwelt* in the acoustic and tactile realms, one that implies a detailed perceptual texture to their lives. The ecology of these animals drove these sensory changes to allow them to marshal their cognitive power to respond flexibly to their new surroundings. The quantitative precision with which sensory sensitivity, discrimination, and identification can be measured also facilitates comparisons to other species. The high correlation between general sensory discrimination and fluid intelligence (and perhaps working memory) in humans also suggests an avenue for further intelligence research in animals.

In addition to sensory adaptations, the transition to the ocean also facilitated social adaptations. For many species, social grouping fostered the sensory integration and behavioral coordination among members necessary for successful hunting, defense, and other activities (Norris and Johnson, 1994) in the absence of much of the physical scaffolding used by terrestrial animals. Group coordination placed a premium on social learning among marine mammals, and it is in social mimicry where a clear difference is found between cetaceans and terrestrial mammals. They are the only mammals other than humans reported to demonstrate vocal and behavioral copying behavior beyond mere rudiments. This copying behavior is strikingly flexible, characterized by learning novel skills, demonstrating both accurate mimicry of physical movement and emulation of end goals, mimicking the behavior of other species, mimicking computer-generated sounds, and copying behaviors of other species, even humans in air. Evidence for social learning in the wild, although not as tightly controlled as in the laboratory, indicates that the abstract learning situations tested in the laboratory have practical generality to the natural environment. For example, vocal mimicry is reported from observations and experiments with bottlenose dolphins in the laboratory (Richards et al., 1984; Richards, 1986) and in the wild (Janik et al., 2006). Acquisition of novel motor behavior is also reported in the laboratory (e.g., Xitco, 1988; Xitco et al., 1998) and in the wild (Wild et al., 2020). Social behavior and vocal imitation provide another rich area for comparative work on intelligence.

While some of the cognitive skills tested in cetaceans and pinnipeds are found in other species, the breadth in marine mammals is marked. For example, although animals with few neurons in their nervous systems, like honeybees with a million neurons, can do delayed matching-to-sample tasks, marine mammals' neuronal tool kit (supported by perhaps a million times that number in killer whales, for example) seems to be expanded. An approach that considers intelligence to be multifaceted considers a wide range of test performances; intelligence might be assessed over procedures testing myriad capacities, for example, perceptual resolution, short and long term memory capacity, imitation, problem solving, and the many other attributes suggested by Bullock (1986). Both cetaceans and pinnipeds have demonstrated successful performance on a broad array of tasks. The relative range compared to other species remains to be evaluated.

Several other factors could be incorporated into a model of intelligence: (1) Analysis of cognitive representation in addition to measurement of stimulus features can provide insight to the way animals make connections. For example, showing that a dolphin can identify visually an object that has previously only been identified through audition and vice-versa indicates a representation independent of modality. (2) The intelligence of a species might be indicated by its ability to learn from experience. In this case, we are talking about more than just a learning set type of experiment but rather the changes that occur over days, weeks, and years showing learning built on previous experiences. Many researchers report anecdotally that marine mammals who engage in years and decades of cognitive work improve in their ability to learn new test procedures over time. Such long-term growth and change are fundamental to our understanding of human intelligence, and the long developmental course of many marine mammals suggests extended neural and behavioral plasticity, as seen in humans. There is now some evidence that behavioral plasticity is, indeed, adaptive (Ducatez et al., 2020), allowing some species to better adjust to and survive in rapidly changing environments. If flexibility and the knowledge attainment it supports are adaptive, then they are subject to evolutionary pressures and will necessarily vary across species. It is possible that comparative psychologists have unintentionally gone out of their way to ignore these factors by focusing study on naive animals placed in impoverished contexts; this method might squelch our ability to find differences across species and between individuals. (3) Anatomical and physiological techniques can greatly enhance the collection efficiency of experimental data. One of the big problems of marine mammal behavioral research is the length of time it takes to collect data with all its attendant costs, small sample sizes, and limits on questions to be asked. For example, a visual acuity test or audiogram for a naïve animal might take a year. Alternatively, good estimates of visual acuity can be determined from measures of retinal ganglion cell density and axial length of the eye (Mass and Supin, 1989), measurements that can be quickly made post-mortem. Good audiogram approximations can be made through evoked potential techniques in less than an hour (Finneran and Houser, 2006). As neural function and organization measurement improves, we may be able to explore valid cognitive characteristics through widely available anatomical techniques like post-mortem DTI. (4) Differences in individual intelligence are a major focus of human intelligence testing, but we do not usually consider this quality in comparing intelligence among species. This is certainly something that animal trainers encounter, when they find major differences in trainability among subjects, although it may not be something that is formally assessed and reported. Variability in intelligence among individuals might reflect the cognitive flexibility of a species better than a static measure of average performance.

Just because comparative psychologists have yet to successfully characterize and delineate all the processes and situations that govern animal thought and behavior does not mean that there are not significant differences in how animals gather information in the world, process it, and act on it across multiple contexts. As indicated here, there are numerous comparisons we could make that might be more fruitful for delineating differences in intelligence than the foundational processes targeted by Macphail. Clearly, these foundational processes exist, but they are recruited differentially across species as their ecologies drive shifts in other systems (e.g., sensory-motor ones) bringing new information to their *Umwelten* and expanding fundamental areas of cognition (e.g., through requiring much faster temporal processing to deal with sound in the water). Marine mammal species transitioned, over the course of evolutionary history, between markedly different ecological settings, and continue to transition between these settings on a daily basis. These transitions may have promoted neural, sensory, and cognitive flexibility reflected in their behavior in the wild and in the laboratory. As long-lived animals who perform well in experimental settings, they are excellently situated to provide insight into the link between ecological and cognitive flexibility and how this may bear on a comparative understanding of intelligence.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The work was a collaborative effort among HH, PC, and GB. All parts of the manuscript reflect a group effort with the exception of The Brain, which was written primarily by PC. All authors contributed to the article and approved the submitted version.

REFERENCES

Abramson, J. Z., Hernández-Lloreda, V., Call, J., and Colmenares, F. (2011). Relative quantity judgments in South American sea lions (*Otaria flavescens*). *Anim. Cogn.* 14, 695–706. doi: 10.1007/s10071-011-0404-7

Ahnelt, P. K., and Kolb, H. (2000). The mammalian photoreceptor mosaic-adaptive design. *Prog. Retin. Eye Res.* 19, 711–777. doi: 10.1016/S1350-9462(00)00012-4

Au, W. W. L. (1993). *The sonar of dolphins*. New York: Springer-Verlag.

Au, W. W. L., and Hastings, M. C. (2008). *Principles of marine acoustics*. New York: Springer.

Au, W. W. L., and Pawloski, D. A. (1992). Cylinder wall thickness difference discrimination by an echolocating Atlantic bottlenose dolphin. *J. Comp. Physiol. A.* 170, 41–47. doi: 10.1007/BF00190399

Bachteler, D., and Dehnhardt, G. (1999). Active touch performance in the Antillean manatee: evidence for a functional differentiation of the facial tactile hairs. *Zoology* 102, 61–69.

Barbas, H., and Rempel-Clower, N. (1997). Cortical structure predicts the pattern of corticocortical connections. *Cereb. Cortex* 7, 635–646. doi: 10.1093/cercor/7.7.635

Bauer, G. B., Colbert, D., Fellner, W., Gaspard, J., and Littlefield, B. (2003). Underwater visual acuity of Florida manatees (*Trichechus manatus latirostris*). *Int. J. Comp. Psychol.* 16, 130–142. doi: 10.1037/e604042013-038

Bauer, G. B., Gaspard, J. C. 3rd, Colbert, D. E., Leach, J. B., Stamper, S. A., Mann, D. A., et al. (2012). Tactile discrimination of textures by Florida manatees (*Trichechus manatus latirostris*). *Mar. Mamm. Sci.* 28, 456–471. doi: 10.1111/j.1748-7692.2012.00565.x

Bauer, G. B., and Harley, H. E. (2001). The mimetic dolphin. *Behav. Brain Sci.* 24, 326–327. doi: 10.1017/S0140525X01243969

Bauer, G. B., and Johnson, C. M. (1994). Trained motor imitation by bottlenose dolphin (*Tursiops truncatus*). *Percept. Mot. Skills* 79, 1307–1315. doi: 10.2466/pms.1994.79.3.1307

Bauer, G. B., and Reep, R. (2017). "Sirenian sensory processes" in *Encyclopedia of animal cognition and behavior*. eds. J. Vonk and T. K. Shackelford (Cham: Springer International Publishing), 1–7.

Bearzi, M. (2006). California sea lions use dolphins to locate food. *J. Mammal.* 87, 606–617. doi: 10.1644/04-MAMM-A-115R4.1

Beran, M. J., and Hopkins, W. D. (2018). Self-control in chimpanzees relates to general intelligence. *Curr. Biol.* 28, 574–579. doi: 10.1016/j.cub.2017.12.043

Berns, G. S., Cook, P. F., Foxley, S., Jbabdi, S., Miller, K. L., and Marino, L. (2015). Diffusion tensor imaging of dolphin brains reveals direct auditory pathway to temporal lobe. *Proc. Biol. Sci.* 282:20151203. doi: 10.1098/rspb.2015.1203

Binet, A., and Simon, T. (1916). *The development of intelligence in children*. Baltimore: Williams & Wilkins.

Brown, D. H., Caldwell, D. K., and Caldwell, M. C. (1966). Observations on the behavior of wild and captive false killer whales, with notes on associated behavior and other genera of captive delphinids. *Los Angeles County Mus. Contr. Sci.* 95, 1–32.

Bruck, J. N. (2013). Decades-long social memory in bottlenose dolphins. *Proc. Biol. Sci.* 280, 1–6. doi: 10.1098/rspb.2013.1726

Bullock, T. H. (1986). "Suggestions for research on ethological and comparative cognition" in *Dolphin cognition and behavior: A comparative approach*. eds. R. J. Schusterman, J. A. Thomas and F. G. Wood (Hillsdale: Lawrence Erlbaum Associates).

Butti, C., Santos, M., Uppal, N., and Hof, P. R. (2013). Von Economo neurons: clinical and evolutionary perspectives. *Cortex* 49, 312–326. doi: 10.1016/j.cortex.2011.10.004

Byrne, R. (1995). *The thinking ape: Evolutionary origins of intelligence*. New York: Oxford University Press.

Caldwell, M. C., and Caldwell, D. K. (1965). Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*). *Nature* 207, 434–435. doi: 10.1038/207434a0

Caldwell, M. C., Caldwell, D. K., and Tyack, P. L. (1990). "A review of the signature whistle hypothesis for the Atlantic bottlenose dolphin, *Tursiops truncatus*" in *The bottlenose dolphin: Recent progress in research*. eds. S. Leatherwood and R. R. Reeves (San Diego: Academic Press).

Cattell, J. M. (1890). Mental tests and measurements. *Mind* 15, 373–380. doi: 10.1093/mind/os-XV.59.373

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.* 54, 1–22. doi: 10.1037/h0046743

Chen, M. J., and Chen, H. C. (1988). Concepts of intelligence: a comparison of Chinese graduates from Chinese and English schools in Hong Kong. *Int. J. Psychol.* 23, 471–487. doi: 10.1080/00207598808247780

Cockcroft, V. G., and Ross, G. J. B. (1990). "Observations on the early development of a captive bottlenose dolphin calf" in *The bottlenose dolphin*. eds. S. Leatherwood and R. R. Reeves (San Diego: Academic Press), 461–478.

Colbert-Luke, D. E., Gaspard, J. C. 3rd, Reep, R. L., Bauer, G. B., Dziuk, K., Cardwell, A., et al. (2015). Eight-choice sound localization by manatees: performance abilities and head related transfer functions. *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 201, 249–259. doi: 10.1007/s00359-014-0973-4

Connor, R. C., Mann, J., and Watson-Capp, J. (2006a). A sex-specific affiliative contact behavior in Indian Ocean bottlenose dolphins, *Tursiops* sp. *Ethology* 112, 631–638. doi: 10.1111/j.1439-0310.2006.01203.x

Connor, R. C., Smolker, R. A., and Bejder, L. (2006b). Synchrony, social behavior and alliance affiliation in Indian Ocean bottlenose dolphins, *Tursiops aduncus*. *Anim. Behav.* 72, 1371–1378. doi: 10.1016/j.anbehav.2006.03.014

Connor, R. C., Smolker, R. A., and Richards, A. F. (1992a). "Dolphin alliances and coalitions" in *Coalitions and alliances in humans and other animals*. eds. H. Alexander and F. B. M. de Waal (Oxford: Oxford University Press), 415–443.

Cook, P. F., Berns, G. S., Colegrove, K., Johnson, S., and Gulland, F. (2018). Postmortem DTI reveals altered hippocampal connectivity in wild sea lions diagnosed with chronic toxicosis from algal exposure. *J. Comp. Neurol.* 526, 216–228. doi: 10.1002/cne.24317

Cook, P. F., Hanke, F., and Reichmuth, C. (2020). "The mind of the Sea Lion" in *Ethology and behavioral ecology of Eared Seales and the Walrus, the Otariids and Odobenid*. New York: Springer. (in press)

Cook, P., Rouse, A., Wilson, M., and Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* 127, 412–427. doi: 10.1037/a0032345

Dantzker, J. L., and Callaway, E. M. (2000). Laminar sources of synaptic input to cortical inhibitory interneurons and pyramidal neurons. *Nat. Neurosci.* 3, 701–707. doi: 10.1038/76656

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W. W. Norton.

Deaner, R. O., Isler, K., Burkart, J., and van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain Behav. Evol.* 70, 115–124. doi: 10.1159/000102973

Deary, I. J. (1994). Sensory discrimination and intelligence: postmortem or resurrection? *Am. J. Psychol.* 107, 95–115. doi: 10.2307/1423292

Deary, I. J., Bell, P. J., Bell, A. J., Campbell, M. L., and Fazal, N. D. (2004). Sensory discrimination and intelligence: testing Spearman's other hypothesis. *Am. J. Psychol.* 117, 1–18. doi: 10.2307/1423593

Dehnhardt, G. (1994). Tactile size discrimination by a California sea lion (*Zalophus californianus*) using its mystacial vibrissae. *J. Comp. Physiol. A.* 175, 791–800. doi: 10.1007/BF00191851

Dehnhardt, G., and Dücker, G. (1996). Tactual discrimination of size and shape by a California sea lion (*Zalophus californianus*). *Anim. Learn. Behav.* 24, 366–374. doi: 10.3758/BF03199008

Dehnhardt, G., and Mauck, B. (2008). "Mechanoreception in secondarily aquatic vertebrates" in *Sensory evolution on the threshold: Adaptations in secondarily aquatic vertebrates*. eds. J. G. M. Thewissen and S. Nummela (Berkeley: University of California Press), 295–314.

Dehnhardt, G., Mauck, B., and Bleckmann, H. (1998). Seal whiskers detect water movements. *Nature* 394, 235–236. doi: 10.1038/28303

Ducatez, S., Sol, D., Sayol, F., and Lefebvre, L. (2020). Behavioural plasticity is associated with reduced extinction risk in birds. *Nat. Ecol. Evol.* 4, 788–793. doi: 10.1038/s41559-020-1168-8

Emerson, R. W., and Cantlon, J. F. (2012). Early math achievement and functional connectivity in the fronto-parietal network. *Dev. Cogn. Neurosci.* 2, S139–S151. doi: 10.1016/j.dcn.2011.11.003

Fay, F. H. (1982). Ecology and biology of the Pacific walrus, *Odobenus rosmarus divergens* Illiger. *N. Am. Fauna* 74, 1–279. doi: 10.3996/nafa.74.0001

Fellner, W., Bauer, G. B., and Harley, H. E. (2006). Cognitive implications of synchrony in dolphins: a review. *Aquat. Mamm.* 32, 511–516. doi: 10.1578/AM.32.4.2006.511

Fellner, W., Bauer, G. B., Stamper, S. A., Losch, B. A., and Dahood, A. (2012). The development of synchronous movement by bottlenose dolphins (*Tursiops truncatus*). *Mar. Mamm. Sci.* 29, E203–E225. doi: 10.1111/j.1748-7692.2012.00609.x

Finneran, J. J., and Houser, D. S. (2006). Comparison of in-air evoked potential and underwater behavioral hearing thresholds in four bottlenose dolphins (*Tursiops truncatus*). *J. Acoust. Soc. Am.* 119, 3181–3192. doi: 10.1121/1.2180208

Ford, J. K. B. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Can. J. Zool.* 69, 1454–1483. doi: 10.1139/z91-206

Friedl, W. A., Nachtigall, P. E., Moore, P. B., Chun, N. K., Haun, J. E., Hall, R. W., et al. (1990). "Taste reception in the Pacific bottlenose dolphin (*Tursiops truncatus gilli*) and California sea lion (*Zalophus californianus*)" in *Sensory abilities of cetaceans: Laboratory and field evidence*. eds. J. A. Thomas and R. A. Kastelein (New York: Plenum Press).

Fuster, J. M. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe. 3rd Edn*. San Diego: Lippincott-Raven.

Galton, F. (1883). *Inquiry into human faculty and its development*. London: Macmillan.

Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Daeschler Hauser, N., et al. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Curr. Biol.* 21, 687–691. doi: 10.1016/j.cub.2011.03.019

Gaspard, J. C. 3rd, Bauer, G. B., Mann, D. A., Boerner, K., Denum, L., Frances, C., et al. (2017). Detection of hydrodynamic stimuli by the postcranial sensory body of Florida manatees (*Trichechus manatus latirostris*). *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 203, 111–120. doi: 10.1007/s00359-016-1142-8

Gaspard, J. C. 3rd, Bauer, G. B., Reep, R. L., Dziuk, K., and Mann, D. A. (2012). Audiogram and auditory critical ratios of two Florida manatees (*Trichechus manatus latirostris*). *J. Exp. Biol.* 215, 1442–1447. doi: 10.1242/jeb.065649

Gaspard, J. C. 3rd, Bauer, G. B., Reep, R. L., Dziuk, K., Read, L., and Mann, D. A. (2013). Detection of hydrodynamic stimuli by the Florida manatee (*Trichechus manatus latirostris*). *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 199, 441–450. doi: 10.1007/s00359-013-0822-x

Gentry, R. L. (1974). The development of social behavior through play in the Steller sea lion. *Am. Zool.* 14, 391–403. doi: 10.1093/icb/14.1.391

Genty, E., and Roeder, J. J. (2006). Self-control: why should sea lions, *Zalophus californianus*, perform better than primates? *Anim. Behav.* 72, 1241–1247. doi: 10.1016/j.anbehav.2006.02.023

Gläser, N., Wieskotten, S., Otter, C., Dehnhardt, G., and Hanke, W. (2011). Hydrodynamic trail following in a California sea lion (*Zalophus californianus*). *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 197, 141–151. doi: 10.1007/s00359-010-0594-5

Goldman-Rakic, P. S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. *Annu. Rev. Neurosci.* 11, 137–156. doi: 10.1146/annurev.ne.11.030188.001033

Goldman-Rakic, P. S., and Preuss, T. M. (1987). Whither comparative psychology. *Behav. Brain Sci.* 10, 666–667. doi: 10.1017/S0140525X00055102

Griebel, U., and Peichel, L. (2003). Colour vision in aquatic mammals—facts and open questions. *Aquat. Mamm.* 29, 18–30. doi: 10.1578/016754203101024040

Griebel, U., and Schmid, A. (1996). Color vision in the manatee (*Trichechus manatus*). *Vis. Res.* 36, 2747–2757. doi: 10.1016/0042-6989(95)00323-1

Griebel, U., and Schmid, A. (2002). Spectral sensitivity and color vision in the bottlenose dolphin (*Tursiops truncatus*). *Mar. Freshw. Behav. Physiol.* 35, 129–137. doi: 10.1080/1023624021000014716

Guinee, L. N., and Payne, K. (1988). Rhyme-like repetitions in songs of humpback whales. *Ethology* 79, 295–306. doi: 10.1111/j.1439-0310.1988.tb00718.x

Hanke, W., Witte, M., Miersch, L., Brede, M., Oeffner, J., Michael, M., et al. (2010). Harbor seal vibrissa morphology suppresses vortex-induced vibrations. *J. Exp. Biol.* 213, 2665–2672. doi: 10.1242/jeb.043216

Harley, H. E. (2008). Whistle discrimination and categorization by the bottlenose dolphin (*Tursiops truncatus*): a review of the signature whistle framework and a perceptual test. *Behav. Process.* 77, 243–268. doi: 10.1016/j.beproc.2007.11.002

Harley, H. E., and Bauer, G. B. (2017). "Cetacean cognition" in *Encyclopedia of animal cognition and behavior*. eds. J. Vonk and T. K. Shackelford (Cham: Springer International Publishing).

Harley, H. E., and DeLong, C. M. (2008). Echoic object representation by the bottlenose dolphin. *Comp. Cogn. Behav. Rev.* 3, 46–65. doi: 10.3819/ccbr.2008.30003

Harley, H. E., Fellner, W., Odell, K., Larsen-Plott, L., and Crowell, S. (2005). Rhythm perception and production by the bottlenose dolphin. *J. Acoust. Soc. Am.* 118:1906. doi: 10.1121/1.4780115

Harley, H. E., Fellner, W., Odell, K., and Stamper, M. A. (2019). "Dolphins can discriminate among marine species using vision alone." in *World Marine Mammal Conference*; December 9–12, 2019; Barcelona, Spain.

Harley, H. E., Putman, E. A., and Roitblat, H. L. (2003). Bottlenose dolphins perceive object features through echolocation. *Nature* 424, 467–469. doi: 10.1038/nature01846

Harley, H. E., Roitblat, H. L., and Nachtigall, P. E. (1996). Object representation in the bottlenose dolphin (*Tursiops truncatus*): integration of visual and echoic information. *J. Exp. Psychol. Anim. Behav. Process.* 22, 164–174. doi: 10.1037//0097-7403.22.2.164

Hastie, G. D., Wilson, B., Tufft, L. H., and Thompson, P. M. (2006). Bottlenose dolphins increase breathing synchrony in response to boat traffic. *Mar. Mamm. Sci.* 19, 74–84. doi: 10.1111/j.1748-7692.2003.tb01093.x

Haug, H. (1987). Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: a stereological investigation of man and his variability and a comparison with some mammals (primates, whales, marsupials, insectivores, and one elephant). *Am. J. Anat.* 180, 126–142. doi: 10.1002/aja.1001800203

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* 3:31. doi: 10.3389/neuro.09.031.2009

Herculano-Houzel, S. (2019). Longevity and sexual maturity vary across species with number of cortical neurons, and humans are no exception. *J. Comp. Neurol.* 527, 1689–1705. doi: 10.1002/cne.24564

Herculano-Houzel, S., Catania, K., Manger, P. R., and Kaas, J. H. (2015). Mammalian brains are made of these: a dataset of the numbers and densities of neuronal and nonneuronal cells in the brain of Glires, Primates, Scandentia, Eulipotyphlans, Afrotherians and Artiodactyls, and their relationship with body mass. *Brain Behav. Evol.* 86, 145–163. doi: 10.1159/000437413

Herculano-Houzel, S., Manger, P. R., and Kaas, J. H. (2014). Brain scaling in mammalian brain evolution as a consequence of concerted and mosaic changes in numbers of neurons and average neuronal cell size. *Front. Neuroanat.* 8:77. doi: 10.3389/fnana.2014.00077

Herman, L. M. (ed.) (1980). "Cognitive characteristics of dolphins" in *Cetacean behavior: Mechanisms and functions* (Malabar: Krieger Publishing Company).

Herman, L. M. (1986). "Cognition and language competencies of bottlenosed dolphins" in *Dolphin cognition and behavior: A comparative approach*. eds. R. J. Schusterman, J. A. Thomas and F. G. Wood (Hillsdale: Lawrence Erlbaum Associates).

Herman, L. M. (1988). The language of animal language research: reply to Schusterman and Gisiner. *Psychol. Rec.* 38, 349–362. doi: 10.1007/BF03395028

Herman, L. M. (1989). In which Procrustean bed does the sea lion sleep tonight? *Psychol. Rec.* 39, 19–49. doi: 10.1007/BF03395052

Herman, L. M. (2002). "Vocal, social, and self-imitation" in *Imitation in animals and artifacts*. eds. K. Dautenbahn and C. L. Nehaniv (Cambridge: MIT Press).

Herman, L. M., Abichandani, S. L., Elhajj, A. N., Herman, E. Y., Sanchez, J. L., and Pack, A. A. (1999). Dolphins (*Tursiops truncatus*) comprehend the referential character of the human pointing gesture. *J. Comp. Psychol.* 113, 347–364. doi: 10.1037/0735-7036.113.4.347

Herman, L. M., and Arbeit, W. R. (1973). Stimulus control and auditory discrimination learning sets in the bottlenose dolphin. *J. Exp. Anal. Behav.* 19, 379–394. doi: 10.1901/jeab.1973.19-379

Herman, L. M., Beach, F. A. 3rd, Pepper, R. L., and Stalling, R. B. (1969). Learning-set formation in the bottlenose dolphin. *Psychon. Sci.* 14, 98–99. doi: 10.3758/BF03332723

Herman, L. M., Pack, A. A., and Hoffmann-Kuhnt, M. (1998). Seeing through sound: dolphins (*Tursiops truncatus*) perceive the spatial structure of objects through echolocation. *J. Comp. Psychol.* 112, 292–305. doi: 10.1037/0735-7036.112.3.292

Herman, L. M., Pack, A. A., and Morrel-Samuels, P. (1993). "Representational and conceptual skills of dolphins" in *Language and communication: Comparative perspectives*. eds. H. R. Roitblat, L. M. Herman and P. Nachtigall (Hillsdale: Lawrence Erlbaum), 273–298.

Herman, L. M., Peacock, M. F., Yunker, M. P., and Madsen, C. J. (1975). Bottle-nosed dolphin: double slit pupil yields equivalent aerial and underwater diurnal acuity. *Science* 189, 650–652. doi: 10.1126/science.1162351

Herman, L. M., Richards, D. G., and Wolz, J. P. (1984). Comprehension of sentences by bottlenose dolphins. *Cognition* 16, 129–219. doi: 10.1016/0010-0277(84)90003-9

Hill, H. M., Webber, K., Kemery, A., Garcia, M., Kuczaj, I. I., and Stan, A. (2015). Can sea lions' (*Zalophus californianus*) use mirrors to locate an object? *Int. J. Comp. Psychol.* 28, 1–12.

Hodos, W. (1987). Animal general intelligence: an idea ahead of its time. *Behav. Brain Sci.* 10:668. doi: 10.1017/S0140525X00055138

Hof, P. R., Chanis, R., and Marino, L. (2005). Cortical complexity in cetacean brains. *Anat. Rec. A. Discov. Mol. Cell. Evol. Biol.* 287, 1142–1152. doi: 10.1002/ar.a.20258

Jaakola, K., Guarino, E., and Rodriguez, M. (2010). Blindfolded imitation in a bottlenose dolphin (*Tursiops truncatus*). *Int. J. Comp. Psychol.* 23, 671–688. doi: 10.1037/e598112013-132

Jaakola, K., Guarino, E., Rodriguez, M., and Hecksher, J. (2013). Switching strategies: a dolphin's use of passive and active acoustics to imitate motor actions. *Anim. Cogn.* 16, 701–709. doi: 10.1007/s10071-013-0605-3

James, W. (1890/1952). *The principles of psychology*. Chicago: Encyclopedia Britannica.

Janik, V. M., Sayigh, L. S., and Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8293–8297. doi: 10.1073/pnas.0509918103

Jardim-Messeder, D., Lambert, K., Noctor, S., Pestana, F. M., de Castro Leal, M. E., Bertelsen, M. F., et al. (2017). Dogs have the most neurons, though not the largest brain: trade-off between body mass and number of neurons in the cerebral cortex of large carnivoran species. *Front. Neuroanat.* 11:118. doi: 10.3389/fnana.2017.00118

Jerison, H. J. (1973). *Evolution of the brain and intelligence*. New York: Academic Press.

Jerison, H. J. (1977). The theory of encephalization. *Ann. N. Y. Acad. Sci.* 299, 146–160. doi: 10.1111/j.1749-6632.1977.tb41903.x

Jerison, H. J. (1985). Animal intelligence as encephalization. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 308, 21–35. doi: 10.1098/rstb.1985.0007

Kaas, J. H. (2008). The evolution of the complex sensory and motor systems of the human brain. *Brain Res. Bull.* 75, 384–390. doi: 10.1016/j.brainresbull.2007.10.009

Kastelein, R. A., Wensveen, P., Hoek, L., and Terhune, J. M. (2009). Underwater hearing sensitivity of harbor seals (*Phoca vitulina*) for narrow noise bands between 0.2 and 80 kHz. *J. Acoust. Soc. Am.* 126, 476–483. doi: 10.1121/1.3132522

Kolchin, S. P., and Belkovich, M. (1973). Tactile sensitivity in *Delphinis delphis*. *Zool. Zhurnal* 52, 620–622.

Krubitzer, L. (2007). The magnificent compromise: cortical field evolution in mammals. *Neuron* 56, 201–208. doi: 10.1016/j.neuron.2007.10.002

Kuczaj, S. A. 2nd, and Walker, R. T. (2006). "How do dolphins solve problems?" in *Comparative cognition: Experimental explorations of animal intelligence*. eds. E. A. Wasserman and T. R. Zentall (New York: Oxford University Press), 580–601.

Kuczaj, S. A. 2nd, and Yeater, D. B. (2006). Dolphin imitation: who, what, when, and why? *Aquat. Mamm.* 32, 413–422. doi: 10.1578/AM.32.4.2006.413

Kuznetsov, V. B. (1990). "Chemical senses of dolphins: quasi-olfaction" in *Sensory abilities of cetaceans: Laboratory and field evidence*. eds. J. A. Thomas and R. A. Kastelein (Kusnetzov: Plenum Press), 481–503.

Kyllonen, P. C., and Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence* 14, 389–433. doi: 10.1016/S0160-2896(05)80012-1

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., et al. (2001). Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imaging* 13, 534–546. doi: 10.1002/jmri.1076

Lee, J. J., McGue, M., Iacono, W. G., Michael, A. M., and Chabris, C. F. (2019). The causal influence of brain size on human intelligence: evidence from within-family phenotypic associations and GWAS modeling. *Intelligence* 75, 48–58. doi: 10.1016/j.intell.2019.01.011

Lilly, J. C. (1965). Vocal mimicry in *Tursiops*: ability to match numbers and durations of human vocal bursts. *Science* 147, 300–301. doi: 10.1126/science.147.3655.300

Lilly, J. C. (1967). *The mind of the dolphin: A nonhuman intelligence*. Garden City: Doubleday.

Lilly, J. C., Miller, A. M., and Truby, H. M. (1968). Reprogramming of the sonic output of the dolphin: sonic burst count matching. *J. Acoust. Soc. Am.* 43, 1412–1424. doi: 10.1121/1.1911001

Lowther, A. D., and Goldsworthy, S. D. (2012). Head start: Australian sea lion pups gain experience of adult foraging grounds before weaning. *Mar. Biol.* 159, 2687–2696. doi: 10.1007/s00227-012-2026-2

Lyamin, O., Pryaslova, J., Lance, V., and Siegel, J. (2005). Animal behavior: continuous activity in cetaceans after birth. *Nature* 435:1177. doi: 10.1038/4351177a

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford University Press.

Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Clarendon Press.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695. doi: 10.1017/S0140525X00054984

Madsen, C. J., and Herman, L. M. (1980). "Social and ecological correlates of cetacean vision and visual appearance" in *Cetacean behavior: Mechanisms and functions*. ed. L. M. Herman (New York: Wiley), 101–147.

Manger, P. R. (2006). An examination of cetacean brain structure with a novel hypothesis correlating thermogenesis to the evolution of a big brain. *Biol. Rev.* 81, 293–338. doi: 10.1017/S1464793106007019

Manger, P. R., Spocter, M. A., and Patzke, N. (2013). The evolutions of large brain size in mammals: the 'over-700-gram club quartet'. *Brain Behav. Evol.* 82, 68–78. doi: 10.1159/000352056

Mann, D., Hill, M., Casper, B., Colbert, D., Gaspard, J., and Bauer, G. B. (2005). Temporal resolution of the Florida manatee (*Trichechus manatus latirostris*) auditory system. *J. Comp. Physiol.* 191, 903–908. doi: 10.1007/s00359-005-0016-2

Mann, J., Stanton, M., Patterson, E. M., Bienenstock, E. J., and Singh, L. O. (2012). Social networks reveal cultural behaviour in tool-using dolphins. *Nat. Commun.* 3:980. doi: 10.1038/ncomms1983

Marino, L. (1998). A comparison of encephalization between odontocete cetaceans and anthropoid primates. *Brain Behav. Evol.* 51, 230–238. doi: 10.1159/000006540

Marino, L., Connor, R. C., Fordyce, R. E., Herman, L. M., Hop, P. R., Lefebvre, L., et al. (2007). Cetaceans have complex brains and complex cognition. *PLoS Biol.* 5:e139. doi: 10.1371/journal.pbio.0050139

Marino, L., McShea, D. W., and Uhen, M. D. (2004). Origin and evolution of large brains in toothed whales. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 281, 1247–1255. doi: 10.1002/ar.a.20128

Mars, R. B., Sotiropoulos, S. N., Passingham, R. E., Sallet, J., Verhagen, L., Khrapitchev, A. A., et al. (2018). Whole brain comparative anatomy using connectivity blueprints. *Elife* 7:e35237. doi: 10.7554/eLife.35237

Mars, R. B., Verhagen, L., Gladwin, T. E., Neubert, F. X., Sallet, J., and Rushworth, M. F. (2016). Comparing brains by matching connectivity profiles. *Neurosci. Biobehav. Rev.* 60, 90–97. doi: 10.1016/j.neubiorev.2015.10.008

Mass, A. M., Odell, D. K., Ketten, D. R., and Supin, A. Y. (1997). Ganglion layer topography and retinal resolution of the Caribbean manatee, *Trichechus manatus latirostris*. *Dokl. Biol. Sci.* 355, 392–394.

Mass, A. M., and Supin, A. Y. (1989). Distribution of ganglion cells in the retina of an Amazon river dolphin (*Inia geoffrensis*). *Aquat. Mamm.* 15, 49–56.

Mauck, B., and Dehnhardt, G. (1997). Mental rotation in a California sea lion (*Zalophus californianus*). *J. Exp. Biol.* 200, 1309–1316.

McNemar, Q. (1964). Lost: our intelligence? Why? *Am. Psychol.* 19, 871–882. doi: 10.1037/h0042008

Mercado, E. 3rd, and Delong, C. M. (2010). Dolphin cognition: representations and processes in memory and perception. *Int. J. Comp. Psychol.* 23, 344–378.

Mercado, E. 3rd, Killebrew, D. A., Pack, A. A., Macha, I. V. B., and Herman, L. M. (2000). Generalization of "same-different" classification abilities in bottlenosed dolphins. *Behav. Process.* 50, 79–94. doi: 10.1016/S0376-6357(00)00100-5

Mercado, E. 3rd, Murray, S. O., Uyeyama, R. K., Pack, A. A., and Herman, L. M. (1998). Memory for recent actions in the bottlenose dolphin (*Tursiops truncatus*): repetition of arbitrary behaviors using an abstract rule. *Anim. Learn. Behav.* 26, 210–218. doi: 10.3758/BF03199213

Miksis, J. L., Tyack, P. L., and Buck, J. R. (2002). Captive dolphins (*Tursiops truncatus*) develop signature whistles that match acoustic features of human-made model sounds. *J. Acoust. Soc. Am.* 112, 728–739. doi: 10.1121/1.1496079

Miller, S. N., Colitz, C. M. H., and Dubielzig, R. R. (2010). Anatomy of the California sea lion globe. *Vet. Ophthalmol.* 13, 1–9. doi: 10.1111/j.1463-5224.2010.00815.x

Moore, B. R. (1992). Avian movement imitation and a new form of mimicry: tracing the evolution of a complex form of learning. *Behaviour* 122, 231–263. doi: 10.1163/156853992X00525

Moore, B. L., Connor, R. C., Allen, S. J., Krützen, M., and King, S. L. (2020). Acoustic coordination by allied male dolphins in a cooperative context. *Proc. Biol. Sci.* 287:20192944. doi: 10.1098/rspb.2019.2944

Morgane, P. J., Jacobs, M. S., and Galaburda, A. (1985). Conservative features of neocortical evolution in dolphin brain. *Brain Behav. Evol.* 26, 176–184. doi: 10.1159/000118774

Morgane, P. J., Jacobs, M. S., and McFarland, W. L. (1980). The anatomy of the brain of the bottlenose dolphin (*Tursiops truncatus*). Surface configurations of the telencephalon of the bottlenose dolphin with comparative anatomical observations in four other cetacean species. *Brain Res. Bull.* 5, 1–107. doi: 10.1016/0361-9230(80)90272-5

Mortensen, H. S., Pakkenberg, B., Dam, M., Dietz, R., Sonne, C., Mikkelsen, B., et al. (2014). Quantitative relationships in delphinid neocortex. *Front. Neuroanat.* 8:132. doi: 10.3389/fnana.2014.00132

Mota, B., and Herculano-Houzel, S. (2015). Cortical folding scales universally with surface area and thickness, not number of neurons. *Science* 349, 74–77. doi: 10.1126/science.aaa9101

Murphy, C. T., French, A. N., Werth, A. J., Marx, M., Martin, W. N., Simmons, N., et al. (2015). "The vibrissal system of the North Atlantic right whale: insight into Sensory Ecology and Inspiration for Biomimetic Technology." in *North Atlantic Right Whale Consortium Annual Meeting*; November 4–5, 2015; New Bedford, Massachusetts.

Musser, W. B., Bowles, A. E., Grebner, D. M., and Crance, J. L. (2014). Differences in acoustic features of vocalizations produced by killer whales crosssocialized with bottlenose dolphins. *J. Acoust. Soc. Am.* 136, 1990–2002. doi: 10.1121/1.4893906

Nachtigall, P. E. (1980). "Odontocete echolocation performance on object size, shape and material" in *Animal sonar systems*. eds. R. G. Busnel and J. F. Fish (New York: Springer).

Nachtigall, P. E., and Hall, R. W. (1984). Taste reception in the bottlenosed dolphin. *Acta Zool.* 172, 147–148.

Nagoshi, C. T. (1987). The epistemology of intelligence: contextual variables, tautologies, and external referents. *Behav. Brain Sci.* 10, 675–676. doi: 10.1017/S0140525X00055229

Newman, L. A., and Robinson, P. R. (2006). The visual pigments of the West Indian manatee (*Trichechus manatus*). *Vis. Res.* 46, 3326–3330. doi: 10.1016/j.visres.2006.03.010

Norris, K. S., and Johnson, C. M. (1994). "Schools and schooling" in *The Hawaiian spinner dolphin*. eds. K. S. Norris, B. Würsig, R. S. Wells and M. Würsig (Berkeley: University of California), 232–242.

Oelschläger, H. H. (2008). The dolphin brain—a challenge for synthetic neurobiology. *Brain Res. Bull.* 75, 450–459. doi: 10.1016/j.brainresbull.2007.10.051

Oelschlager, H. H. A., and Kemp, B. (1999). Ontogenesis of the sperm whale brain. *J. Comp. Neurol.* 399, 210–228. doi: 10.1002/(sici)1096-9861(19980921)399:2<210::aid-cne5>3.0.co;2-3

Olkowicz, S., Kocourek, M., Lučan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., et al. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proc. Natl. Acad. Sci. U. S. A.* 113, 7255–7260. doi: 10.1073/pnas.1517131113

Oppermann, D., Schramme, J., and Neumeyer, C. (2016). Rod-cone based color vision in seals under photopic conditions. *Vis. Res.* 125, 30–40. doi: 10.1016/j.visres.2016.04.009

O'Shea, T. J., and Reep, R. L. (1990). Encephalization quotients and life-history traits in the Sirenia. *J. Mammal.* 71, 534–543. doi: 10.2307/1381792

Oztas, E. (2003). Neuronal tracing. *Neuroanatomy* 2:5.

Pack, A. (2015). "Experimental studies of dolphin cognitive abilities" in *Dolphin communication and cognition: Past, present, and future*. eds. D. L. Herzing and C. M. Johnson (Cambridge: MIT Press).

Pack, A. A., and Herman, L. M. (1995). Sensory integration in the bottlenosed dolphin: intermediate recognition of complex shapes across the senses of echolocation and vision. *J. Acoust. Soc. Am.* 98, 722–733. doi: 10.1121/1.413566

Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876. doi: 10.1038/nature03287

Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* 19, 827–830. doi: 10.1016/j.cub.2009.03.038

Payne, R. (1983). *Communication and behavior of whales*. Boulder: Westview Press.

Pepper, R. L., and Defran, R. H. (1975). *Dolphin trainer's handbook: Basic training*. San Diego: Naval Undersea Center.

Perelberg, A., and Schuster, R. (2009). Bottlenose dolphins (*Tursiops truncates*) prefer to cooperate when petted: integrating proximate and ultimate explanations ii. *J. Comp. Psychol.* 123, 45–55. doi: 10.1037/a0013585

Pitcher, B. J., Harcourt, R. G., Schaal, B., and Charrier, I. (2011). Social olfaction in marine mammals: wild female Australian sea lions can identify their pup's scent. *Biol. Lett.* 7, 60–62. doi: 10.1098/rsbl.2010.0569

Popov, V. V., Ladygina, T. F., and Supin, A. Y. (1986). Evoked potentials of the auditory cortex of the porpoise, *Phocoena phocoena*. *J. Comp. Physiol. A.* 158, 705–711. doi: 10.1007/BF00603828

Povinelli, M., Panin, M., Mazzariol, S., Giurisato, M., Ballarin, C., Roncon, G., et al. (2014). Notes on the brain and encephalization quotient of two sperm whales with a synthesis of the literature and indications of a new method of extraction. *Nat. Hist. Sci.* 1, 131–138. doi: 10.4081/nhs.2014.202

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678. doi: 10.1016/j.neuron.2011.09.006

Preston, A. R., and Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, R764–R773. doi: 10.1016/j.cub.2013.05.041

Pryor, K. W., Haag, R., and O'Reilly, J. (1969). The creative porpoise: training for novel behavior. *J. Exp. Anal. Behav.* 12, 653–661. doi: 10.1901/jeab.1969.12-653

Raghanti, M. A., Spurlock, L. B., Treichler, F. R., Weigel, S. E., Stimmelmayr, R., Butti, C., et al. (2015). An analysis of von Economo neurons in the cerebral cortex of cetaceans, artiodactyls, and perissodactyls. *Brain Struct. Funct.* 220, 2303–2314. doi: 10.1007/s00429-014-0792-y

Ralls, K., Fiorelli, P., and Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Can. J. Zool.* 63, 1050–1056. doi: 10.1139/z85-157

Ravignani, A. (2019). Singing seals imitate human speech. *J. Exp. Biol.* 222:jeb208447. doi: 10.1242/jeb.208447

Ravignani, A., Kello, C. T., de Reus, K., Kotz, S. A., Dalla Bella, S., Méndez-Aróstegui, M., et al. (2019). Ontogeny of vocal rhythms in harbor seal pups: an exploratory study. *Curr. Zool.* 65, 107–120. doi: 10.1093/cz/zoy055

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Oxford University Press.

Reep, R. L., and Bonde, R. K. (2006). *The Florida manatee: Biology and conservation*. Gainesville: University Press of Florida.

Reichmuth, C., and Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Curr. Opin. Neurobiol.* 28, 66–71. doi: 10.1016/j.conb.2014.06.011

Reichmuth, C., Holt, M. M., Mulsow, J., Sills, J. M., and Southall, B. L. (2013). Comparative assessment of amphibious hearing in pinnipeds. *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 199, 491–507. doi: 10.1007/s00359-013-0813-y

Reidenberg, J. S. (2007). Anatomical adaptations of aquatic mammals. *Anat. Rec.* 290, 507–513. doi: 10.1002/ar.20541

Reiss, D., and McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): evidence for vocal learning. *J. Comp. Psychol.* 107, 301–312. doi: 10.1037/0735-7036.107.3.301

Rice, D. W. (1998). *Marine mammals of the world: Systematics and distribution*. Lawrence: The Society for Marine Mammalogy.

Richards, D. G. (1986). "Dolphin vocal mimicry and vocal object labelling" in *Dolphin cognition and behavior: A comparative approach*. eds. R. J. Schusterman, J. A. Thomas and F. G. Wood (Hillsdale: Lawrence Erlbaum Associates).

Richards, D. G., Wolz, J. P., and Herman, L. M. (1984). Vocal mimicry of computer generated sounds and vocal labeling by a bottlenose dolphin, *Tursiops truncatus*. *J. Comp. Psychol.* 98, 10–28. doi: 10.1037/0735-7036.98.1.10

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016

Ridgway, S. H., Brownson, R. H., Van Alstyne, K. R., and Hauser, R. A. (2019). Higher neuron densities in the cerebral cortex and larger cerebellums may limit dive times of delphinids compared to deep-diving toothed whales. *PLoS One* 14:e0226206. doi: 10.1371/journal.pone.0226206

Ridgway, S. H., and Carder, D. A. (1990). "Tactile sensitivity, somatosensory responses, skin vibrations, and skin surface ridges of the bottlenose dolphin, *Tursiops truncatus*" in *Sensory abilities of cetaceans: Laboratory and field evidence*. eds. J. A. Thomas and R. A. Kastelein (New York: Plenum Press), 163–179.

Ridgway, S. H., and Carder, D. A. (1993). Features of dolphin skin with potential hydrodynamic importance. *IEEE Eng. Med. Biol.* 12, 83–88. doi: 10.1109/51.232347

Ridgway, S., Carder, D., Jeffries, M., and Todd, M. (2012). Spontaneous human speech mimicry by a cetacean. *Curr. Biol.* 22, R860–R861. doi: 10.1016/j.cub.2012.08.044

Rilling, M. (1990). Comparing representations between species intelligently. *Behav. Brain Sci.* 13, 392–393. doi: 10.1017/S0140525X00079358

Roth, G., and Dicke, U. (2017). "Evolution of cognitive brains: mammals" in *Evolution of the brain, cognition, emotion in vertebrates*. eds. S. Watanabe, M. A. Hofman and T. Shimizu (Tokyo: Springer).

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Salthouse, T. E., and Pink, J. E. (2008). Why is working memory related to fluid intelligence. *Psychon. Bull. Rev.* 15, 364–371. doi: 10.3758/PBR.15.2.364

Samuelson, D., McGee, J., Ben-Shlomo, G., Bauer, G., Murphy, C., Howland, H., et al. (2012). "An overview of the visual system of the Florida manatee (*Trichechus manatus latirostris*)." in *Florida Marine Mammal Health Conference IV*; April 24–27; Mote Marine Laboratory, Florida.

Sawyer, E. K., Turner, E. C., and Kaas, J. H. (2016). Somatosensory brainstem, thalamus, and cortex of the California sea lion (*Zalophus californianus*). *J. Comp. Neurol.* 524, 1957–1975. doi: 10.1002/cne.23984

Scheumann, M., and Call, J. (2004). The use of experimenter-given cues by South African fur seals (*Arctocephalus pusillus*). *Anim. Cogn.* 7, 224–230. doi: 10.1007/s10071-004-0216-0

Scholtyssek, C., Kelber, A., and Dehnhardt, G. (2015). Why do seals have cones? Behavioural evidence for colour blindness in harbour seals. *Anim. Cogn.* 18, 551–560. doi: 10.1007/s10071-014-0823-3

Schusterman, R. J., and Gisiner, R. (1988). Artificial language comprehension in dolphins and sea lions: the essential cognitive skills. *Psychol. Rec.* 38, 311–348. doi: 10.1007/BF03395027

Schusterman, R. J., and Gisiner, R. (1997). "Pinnipeds, porpoises and parsimony: animal language research viewed from a bottom-up perspective" in *Anthropomorphism, anecdotes and animals: The emperor's new clothes?* eds. R. W. Mitchell, N. S. Thompson and H. L. Miles (Albany, New York: State University of New York Press), 370–382.

Schusterman, R. J., Kastak, C. R., and Kastak, D. (2002). "The cognitive sea lion: meaning and memory in the laboratory and in nature" in *The cognitive animal: Empirical and theoretical perspectives on animal cognition*. eds. M. Bekoff, C. Allen and G.M. Burghardt (Cambridge, MA: The MIT Press), 217–228.

Seehaus, A., Roebroeck, A., Bastiani, M., Fonseca, L., Bratzke, H., Lori, N., et al. (2015). Histological validation of high-resolution DTI in human post mortem tissue. *Front. Neuroanat.* 9:98. doi: 10.3389/fnana.2015.00098

Sekiguchi, Y., Arai, K., and Kohshima, S. (2006). Sleep in continuously active dolphins. *Nature* 441, E9–E10. doi: 10.1038/nature04898

Shepherd, G. M. (2015). *Foundations of the neuron doctrine*. Oxford: Oxford University Press.

Shettleworth, S. J. (1987). Intelligence: more than a matter of associations. *Behav. Brain Sci.* 10:679. doi: 10.1017/S0140525X00055266

Shettleworth, S. J. (2010). Clever animals and killjoy explanations in comparative psychology. *Trends Cogn. Sci.* 14, 477–481. doi: 10.1016/j.tics.2010.07.002

Sjare, B., Stirling, I., and Spencer, C. (2003). Structural variation in the songs of Atlantic walruses breeding in the Canadian high Arctic. *Aquat. Mamm.* 29, 297–318. doi: 10.1578/016754203101024121

Sokolov, V. E., Ladygina, T., and Supin, A. (1972). Localization of sensory zones in the dolphin cerebral cortex. *Dokl. Akad. Nauk SSSR* 202, 490–493.

Song, M., Zhou, Y., Li, J., Liu, Y., Tian, L., Yu, C., et al. (2008). Brain spontaneous functional connectivity and intelligence. *NeuroImage* 41, 1168–1176. doi: 10.1016/j.neuroimage.2008.02.036

Sousa-Lima, R. S., Paglia, A. P., and De Fonseca, G. A. B. (2002). Signature information and individual recognition in the isolation calls of Amazonian manatees, *Trichechus inunguis* (Mammalia: Sirenia). *Anim. Behav.* 63, 301–310. doi: 10.1006/anbe.2001.1873

Spearman, C. (1904). "General intelligence," objectively determined and measured. *Am. J. Psychol.* 15, 201–293. doi: 10.2307/1412107

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge: Cambridge University Press.

Sternberg, R. J., Conway, B. E., Ketron, J. L., and Bernstein, M. (1981). People's conceptions of intelligence. *J. Pers. Soc. Psychol.* 41, 37–55. doi: 10.1037/0022-3514.41.1.37

Stich, K. P., Dehnhardt, G., and Mauck, B. (2003). Mental rotation of perspective stimuli in a California sea lion (*Zalophus californianus*). *Brain Behav. Evol.* 61, 102–112. doi: 10.1159/000069355

Sticken, J., and Dehnhardt, G. (2000). Salinity discrimination in harbour seals: a sensory basis for spatial orientation in the marine environment? *Naturwissenschaften* 87, 499–502. doi: 10.1007/s001140050767

Supin, A. Y., Popov, V. V., and Mass, A. M. (2001). *The sensory physiology of aquatic mammals*. Boston: Kluwer Academic.

Tarpley, R. J., and Ridgway, S. H. (1994). Corpus callosum size in delphinid cetaceans. *Brain Behav. Evol.* 44, 156–165. doi: 10.1159/000113587

Tayler, C. K., and Saayman, G. S. (1973). Imitative behavior by Indian Ocean bottlenose dolphins (*Tursiops aduncus*) in captivity. *Behaviour* 44, 286–297. doi: 10.1163/156853973X00436

Thompson, R. K. R., and Herman, L. M. (1977). Memory for lists of sounds by the bottle-nosed dolphin: convergence of memory processes with humans? *Science* 195, 501–503. doi: 10.1126/science.835012

Thompson, R. K. R., and Herman, L. M. (1981). Auditory delayed discriminations by the dolphin: nonequivalence with delayed-matching performance. *Anim. Learn. Behav.* 9, 9–15. doi: 10.3758/BF03212019

Thorpe, W. H. (1963). *Learning and instinct in animals*. London: Methuen.

Trites, A. W., Porter, B. P., Deecke, V. B., Coombs, A. P., Marcotte, M. L., and Rosen, D. A. (2006). Insights into the timing of weaning and the attendance patterns of lactating Steller sea lions (*Eumetopias jubatus*) in Alaska during winter, spring, and summer. *Aquat. Mamm.* 32, 85–97. doi: 10.1578/AM.32.1.2006.85

Turner, E. C., Sawyer, E. K., and Kaas, J. H. (2017). Optic nerve, superior colliculus, visual thalamus, and primary visual cortex of the northern elephant seal (*Mirounga angustirostris*) and California sea lion (*Zalophus californianus*). *J. Comp. Neurol.* 525, 2109–2132. doi: 10.1002/cne.24188

Tyack, P. (1986). Whistle repertoires of two bottlenosed dolphins, *Tursiops truncatus*: mimicry of signature whistles? *Behav. Ecol. Sociobiol.* 18, 251–257. doi: 10.1007/BF00300001

Vauclair, J. (1990). Wanted: cognition. *Behav. Brain Sci.* 13, 393–394. doi: 10.1017/S0140525X0007936X

Visser, I. N., Smith, T. G., Bullock, I. D., Green, G. D., Carlsson, O. G. L., and Imberti, S. (2008). Antarctic peninsula killer whales (*Orcinus orca*) hunt seals and a penguin on floating ice. *Mar. Mamm. Sci.* 24, 225–234. doi: 10.1111/j.1748-7692.2007.00163.x

von Bartheld, C. S., Bahney, J., and Herculano-Houzel, S. (2016). The search for true numbers of neurons and glial cells in the human brain: a review of 150 years of cell counting. *J. Comp. Neurol.* 524, 3865–3895. doi: 10.1002/cne.24040

von Uexküll, J. (1934/1957). "A stroll through the worlds of animals and men: a picture book of invisible worlds" in *Instinctive behavior: The development of a modern concept*. ed. C. H. Schiller (New York: International Universities Press, Inc), 5–80.

Walker, S. F. (1990). Species comparisons versus comparative epistemology. *Behav. Brain Sci.* 13, 394–395. doi: 10.1017/S0140525X00079371

West, J. A., Sivak, J. G., Murphy, C. J., and Kovacs, K. M. (1991). A comparative study of the anatomy of the iris and ciliary body in aquatic mammals. *Can. J. Zool.* 69, 2594–2607. doi: 10.1139/z91-366

Whitehead, H., and Rendell, L. (2015). *The cultural lives of whales and dolphins*. Chicago: University of Chicago Press.

Whiten, A. (2001). Imitation and cultural transmission in apes and cetaceans. *Behav. Brain Sci.* 24, 359–360. doi: 10.1017/S0140525X01603960

Whiten, A., and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. *Adv. Study Behav.* 21, 239–283. doi: 10.1016/S0065-3454(08)60146-1

Wild, S., Hoppitt, W. J. E., Allen, S. J., and Krützen, M. (2020). Integrating genetic, environmental and social networks to reveal transmission pathways of a dolphin foraging innovation. *Curr. Biol.* 30, 1–7. doi: 10.1016/j.cub.2020.05.069

Wilson, M., and Cook, P. F. (2016). Rhythmic entrainment: why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychon. Bull. Rev.* 23, 1647–1659. doi: 10.3758/s13423-016-1013-x

Witte, M. (2012). "On the wake flow dynamics behind harbor seal vibrissae—a fluid mechanical explanation for an extraordinary capability" in *Nature-inspired fluid mechanics. Notes on numerical fluid mechanics and multidisciplinary design*. *Vol. 119*. eds. C. Tropea and H. Bleckmann (New York: Springer).

Worthy, G. A., and Hickie, J. P. (1986). Relative brain size in marine mammals. *Am. Nat.* 128, 445–459. doi: 10.1086/284579

Wright, A. K., Theilmann, R. J., Ridgway, S. H., and Scadeng, M. (2018). Diffusion tractography reveals pervasive asymmetry of cerebral white matter tracts in the bottlenose dolphin (*Tursiops truncatus*). *Brain Struct. Funct.* 223, 1697–1711. doi: 10.1007/s00429-017-1525-9

Xiang, H., Dediu, D., Roberts, L., Oort, E. V., Norris, D. G., and Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Lang. Learn.* 62, 110–130. doi: 10.1111/j.1467-9922.2012.00708.x

Xiao, L., Stephen, J. M., Wilson, T. W., Calhoun, V. D., and Wang, Y. P. (2018). Alternating diffusion map based fusion of multimodal brain connectivity networks for IQ prediction. *IEEE Trans. Biomed. Eng.* 66, 2140–2151. doi: 10.1109/TBME.2018.2884129

Xitco, M. J. (1988). Mimicry of modeled behaviors by bottlenose dolphins. (Unpublished master's thesis). University of Hawaii.

Xitco, M. J., Gory, J. D., and Kuczaj, S. (2001a). Spontaneous pointing by bottlenose dolphins (*Tursiops truncatus*). *Anim. Cogn.* 4, 115–123. doi: 10.1007/s100710100107

Xitco, M. J., Gory, J. D., and Kuczaj, S. (2001b). Dolphin pointing is linked to the attentional behavior of a receiver. *Anim. Cogn.* 7, 231–238. doi: 10.1007/s10071-004-0217-z

Xitco, M. J. Jr., Harley, H. E., and Brill, R. L. (1998). "Action level imitation by bottlenose dolphins." in *Napoli Social Learning Conference*; June 30–July 5, 1998; Naples, Italy.

Xitco, M. J., and Roitblat, H. L. (1996). Object recognition through eavesdropping: passive echolocation in bottlenose dolphins. *Anim. Learn. Behav.* 24, 355–365. doi: 10.3758/BF03199007

Yamasaki, F., Komatsu, S., and Kamiya, T. (1980). A comparative morphological study on the tongues of manatee and dugong (*Sirenia*). *Sci. Rep. Whales Res. Ins. 32*, 127–144.

Yeo, B. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., et al. (2015). Functional specialization and flexibility in human association cortex. *Cereb. Cortex* 25, 3654–3672. doi: 10.1093/cercor/bhu217

Zentall, T. R. (1996). "An analysis of imitative learning in animals" in *Social learning in animals: The roots of culture*. eds. C. Heyes and B. G. Galef (San Diego: Academic Press), 221–243.

Conflict of Interest: HH participates in scientific and welfare collaborations with Walt Disney World's Animals, Science, and Education. The Seas, Epcot®, Walt Disney World® Resorts was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

The remaining authors declare that the research was conducted in the absence of any commercial conflicts of interest.

frontiers
in Psychology

REVIEW
published: 23 September 2020
doi: 10.3389/fpsyg.2020.560669

Neural Processes Underlying Tool Use in Humans, Macaques, and Corvids

María J. Cabrera-Álvarez and Nicola S. Clayton*

Department of Psychology, University of Cambridge, Cambridge, United Kingdom

It was thought that tool use in animals is an adaptive specialization. Recent studies, however, have shown that some non-tool-users, such as rooks and jays, can use and manufacture tools in laboratory settings. Despite the abundant evidence of tool use in corvids, little is known about the neural mechanisms underlying tool use in this family of birds. This review summarizes the current knowledge on the neural processes underlying tool use in humans, macaques and corvids. We suggest a possible neural network for tool use in macaques and hope this might inspire research to discover a similar brain network in corvids. We hope to establish a framework to elucidate the neural mechanisms that supported the convergent evolution of tool use in birds and mammals.

Keywords: tool use, neural mechanisms, neural network, causal reasoning, macaques, corvids

Edited by:
Michael Colombo,
University of Otago, New Zealand

Reviewed by:
Andrew N. Iwaniuk,
University of Lethbridge, Canada
Jean-Baptiste Leca,
University of Lethbridge, Canada
Camilla Cenni,
Department of Psychology, University of Lethbridge, Canada, in collaboration with reviewer J-BL

****Correspondence:***
María J. Cabrera-Álvarez
CabreraResearch@gmail.com

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *09 May 2020*
Accepted: *31 August 2020*
Published: *23 September 2020*

Citation:
Cabrera-Álvarez MJ and Clayton NS (2020) Neural Processes Underlying Tool Use in Humans, Macaques, and Corvids. Front. Psychol. 11:560669. doi: 10.3389/fpsyg.2020.560669

INTRODUCTION

The classic definition of tool use in animals is "the use of an external object as a functional extension of mouth or beak, hand or claw, in the attainment of an immediate goal" (van Lawick-Goodall, 1970, p.195), following the observations of tool use in a wild population of chimpanzees (Goodall, 1964). Since then, the definition of tool use has evolved to encompass the physical properties of tools, specifying that a tool must alter "the form, position or condition of another object, another organism or the user itself" (Beck, 1980, p.10), and to include ways of manipulating and manufacture of tools (St Amant and Horton, 2008) in order to distinguish between those species that have the capability to create or use an external object to solve a problem via dynamic mechanical interactions, i.e., "flexible tool users," and those that are "stereotyped tool users," i.e., species that perform object-related mechanical actions that are not intended to have a goal-directed interaction with another object (Hunt et al., 2013). This is an important distinction because flexible tool use is not phylogenetically widespread and seems to require a certain level of cognitive processing (Baber, 2003). In this paper, our focus will primarily be on flexible tool use.

The definition of tool use specifies that the tool has to be detached from the substrate and directly held in the animal's hand or mouth. This definition has opened a debate in the field of tool-use since many animals, including birds, cannot hold tools in their hand or mouth but instead use their beak or foot and others may throw or drop objects to achieve their goals (St Amant and Horton, 2008). Given these differences in the way tools are used, a distinction was made between "true tool-users," i.e., those that follow the traditional definition of tool-use albeit broadening the criteria to include beaks and limbs [e.g., New Caledonian crows (*Corvus moneduloides*) using twigs as hooks to retrieve food from small holes (Hunt, 1996)], and borderline or "proto-tool" users, i.e., those that use objects to obtain food that would otherwise be out of reach but do not hold these objects in their limbs or mouths/beaks [e.g., American crows (*Corvus brachyrhynchos*) dropping nuts in roads to get them crashed by the passing of automobiles over them (Grobecker and Pietsch, 1978)]. From a cognitive perspective, it makes sense to make this distinction, since holding the tool in a part of the user's body might make the user include that object as part of their own body, while those animals that just throw or drop objects might not have the ability to include the object in their body-image. As Jacob Bronowski graciously expressed, "the hand is the cutting edge of the mind" (Bronowski, 1975, p.116), and thus we should not forget their importance for body awareness.

Recently, Fragaszy and Mangalam (2018) developed a theory of tool use which the authors termed "tooling." This theory is framed in biomechanical and spatial concepts of action in order to determine when an object is used as a tool. It aims to reconceptualize the phenomenon of tool use. The authors developed the concept of "tooling," which we adopt as a legitimate description of what we consider tool use: "Tooling is deliberately producing a mechanical effect upon a target object/surface by first grasping an object, thus transforming the body into the body-plus-object system, and then using the body-plus-object system to manage (at least one) spatial relation(s) between a grasped object and a target object/surface, creating a mechanical interface between the two" (Fragaszy and Mangalam, 2018, p.194).

Before Goodall's observations, it was widely believed that tool use was a uniquely human characteristic (Oakley, 1972), since the use and manufacture of tools has historically been linked to the emergence of technical intelligence in humans given the complex problem solving and planning needed to create and use composite tools (i.e., tools made of two or more joined parts) (Ambrose, 2001). Since then, many observations of both proto- and true-tool use have been reported, not only in primates but also in other mammals (Mann et al., 2008; Root-Bernstein et al., 2019), birds (Hunt, 1996), reptiles (Dinets et al., 2015), fish (Brown, 2012), and insects (Pierce, 1986). It is worth mentioning that flexible tool use is mostly found in birds and primates, while insects and fish mostly show stereotypical tool use (Hunt et al., 2013). Most significantly, Hunt made the remarkable discovery that New Caledonian crows manufacture and craft a variety of tools which they use to obtain food that cannot be reached in any other way (Hunt, 1996). Subsequent research by Hunt and other members of Gray and Taylor's research groups have revealed many fascinating findings about the complexity of physical cognition in these birds (Hunt and Gray, 2004; Taylor et al., 2009, 2010). These observations provide evidence that evolutionarily distant species are capable of similar complex motor skills that require a certain level of cognitive ability to perform them.

Hodos (1987) suggested the study of animal tool use as one of the specific intellectual abilities that can be used as a proxy to understanding the concept of animal intelligence proposed by Macphail (1987). Hodos (1987) argued that we would understand animal intelligence more rapidly if we focus our efforts in the study of specific intellectual abilities rather than in the search of general intelligence. However, our understanding of the neural processes underlying tool use in non-human animals remains scant, even though descriptive reports and ecological literature related to animal tool use has grown dramatically.

Having a proper understanding of the neural mechanisms underlying tool use is pivotal to comprehend the evolutionary processes that enabled evolutionarily distant animals to achieve similar cognitive capabilities because the comparison of the brain structures that are needed for this specific intellectual ability will shed light on the evolutionary paths that give rise to animal intelligence. This review compiles information regarding brain areas active during tool use in humans and macaques, and will suggest possible areas in the bird brain that could be a focus of study in the future.

TOOL USE IN HUMANS

The neural basis of tool use in humans was first studied in patients with brain lesions that impaired their ability to use tools, a consequence of a disorder known as apraxia (Johnson-Frey, 2004; Lewis, 2006; Higuchi et al., 2007; Frey, 2008). Patients with apraxia do not show difficulties in linguistic, sensory or lower level motor functions. However, they do exhibit an impaired ability to carry out acquired skills, including, although not specific to, the use of tools. There are two types of apraxia that affect tool use: ideomotor apraxia and ideational, or conceptual, apraxia. In ideomotor apraxia, although patients know what to do with a tool and can grasp and manipulate it, they seem to be unable to represent the associated motor actions needed to properly use the tools, failing to pantomime how the tools are used. These patients suffer from damage to the left posterior parietal and/or premotor cortex, or damage to the corpus callosum that results in isolation of the left hemisphere from the right (Frey, 2008). On the other hand, in ideational or conceptual apraxia, patients know how to handle the tools, but can not follow the order of a sequence of movements to achieve a goal that is the product of a multistep action. Ideational apraxia patients commonly have lesions at the intersection of the temporal-parietal-occipital cortices of the left hemisphere (Frey, 2008). The studies of apraxia show, not only that the motor skills and conceptual knowledge about complex actions such as tool use are dissociable, but also that they are represented in dissociable neural systems within the left cerebral hemisphere (Johnson-Frey, 2004; Lewis, 2006; Frey, 2008).

Subsequently, fMRI and PET studies in healthy humans highlighted the areas of the cortex that are active during tool use or during mimicking and imaging tool use (reviewed in Lewis, 2006, see his Figure 5B). Lewis' figures (2006) show that most activity during tool use is in the left hemisphere of the human cortex. This is the case for right-handed people, while the right hemisphere might have higher activation during tool use in left-handed people (reviewed in Lewis, 2006). Further studies comparing right- vs. left-handed people are needed in order to disentangle the lateralization of tool use in humans.

Peeters et al. (2009, 2013) identified the anterior supramarginal gyrus (aSMG), a specific region of the human brain left inferior parietal lobule (IPL), as being involved in both the execution and observation of tool actions. They did not find a similar activation in the IPL of rhesus monkeys that were trained to use tools. In a subsequent study, Gallivan et al. (2013) found specific brain areas involved only in tool-related actions, in contrast with brain regions involved only in hand-related actions, and suggested a brain network for human tool use (**Figure 1**). This brain network was later expanded to include the connection between the aSMG, which is active during the observation of the tool being moved to achieve a goal, and the putative human homolog of anterior intraparietal (phAIP) in macaques, a region active during observation of tool grasping (**Figure 2**; Orban and Caruana, 2014). **Figure 2** showcases the cognitive processes involved in tool use, which are reasoning affordances (i.e., forming conclusions about the qualities of an object that defines its possible uses), mechanical problem solving (i.e., finding solutions to novel mechanical problems), and semantic knowledge (i.e., a type of long term memory consisting of concepts, facts, ideas, and beliefs). This suggested brain network also highlights the brain areas underpinning these cognitive processes, which would all provide input to the aSMG. For further reading, additional reviews on human tool use have been recently published (Osiurak and Badets, 2016; Reynaud et al., 2016).

Although the cerebellum is not mentioned in these brain networks, its function in tool-use has been under debate despite

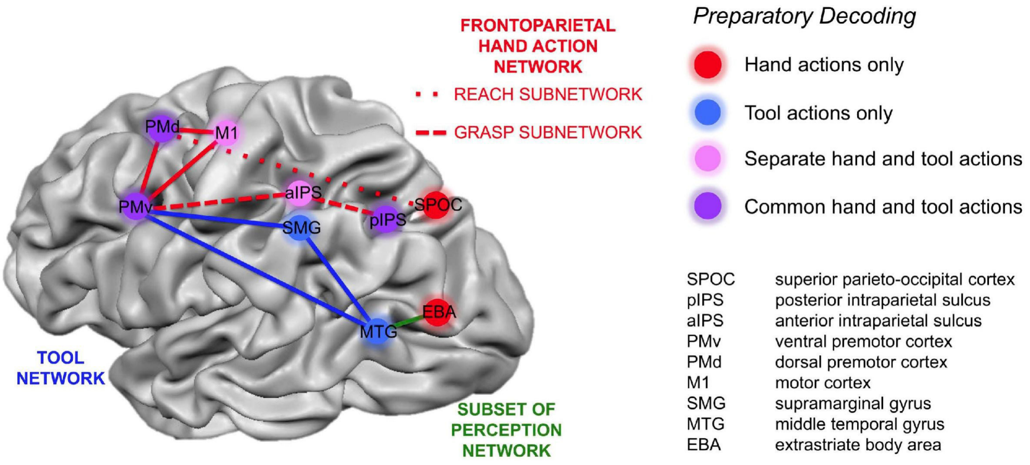


FIGURE 1 | Cortical areas that coded only hand actions (red), only tool actions (blue), both hand and tool actions but using different neural representations (pink), and areas coding an action independently of it being performed with the hand or a tool (purple). Reproduced from Gallivan et al. (2013).

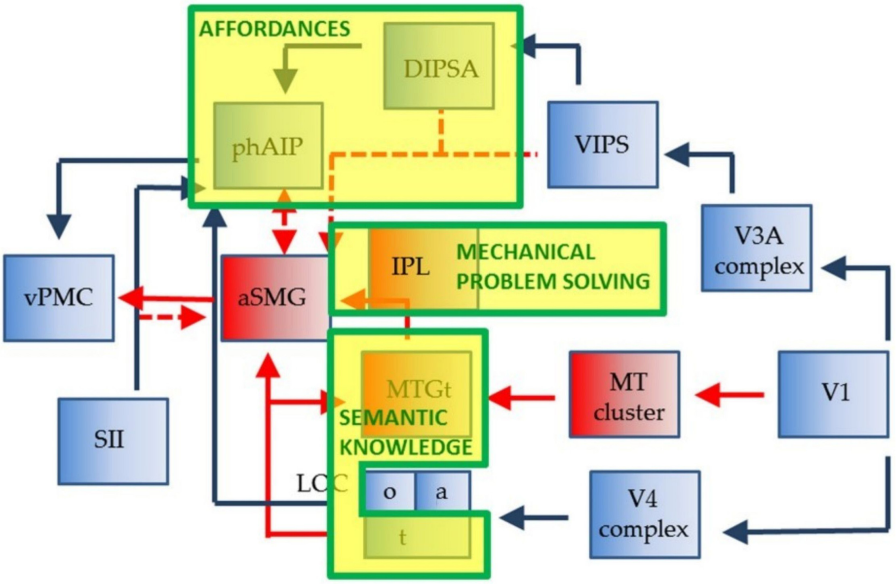


FIGURE 2 | Representation of the human brain network suggested to be involved in tool use, including biological actions in blue, tool actions in red and blue, and cognitive processes in yellow boxes with green outlines. Dashed lines represent postulated connections. Reproduced from Orban and Caruana (2014).

its role in sensory-motor control and learning of complex action sequences. For example, it is unclear why cerebellar lesions do not cause ideomotor apraxia if the cerebellum stores representations of tool-use skills (Johnson-Frey, 2004). Yet, its development during human evolution and its interactions with the neocortex have been related to greater computational efficiency for dealing with increasingly complex cultural and conceptual environments (Weaver, 2005). In great apes, cerebellar volume and lateral asymmetry of the cerebellum is related to species-specific differences in performance and hand preference for task that require precise motor skills, such as tool use and aimed throwing (Cantalupo and Hopkins, 2010). Additionally, the cerebellum plays an important role in the evolution of the capacity for planning, execution and understanding of complex behavioral sequences—including tool use and language (Barton, 2012; Barton and Venditti, 2014), and the neural mechanisms of tool use may be precursors for the neural basis of language and abstract thought (Hihara et al., 2003; Iriki and Taoka, 2012; Steele et al., 2012). These authors have suggested that the use of tools as equivalents of body parts would have triggered the development of more advanced problem solving skills: abstract thinking (i.e., the ability to think about things that are not physically present) was essential for the development of conceptual thinking (i.e., the ability of integrating a series of features that group together to form a class of ideas or objects) and hence, the development of language.

Despite the relevance of tool use to human evolution and the benefits that its study would bring to our understanding of brain evolution, the neural processes underlying tool use in other animals have not yet been closely studied, partly due to the methodological difficulty in reproducing the same sort of studies in animals. Given the hypothesis that tool use in humans would have led to the evolution of more complex cognitive abilities, comparisons of both the cognitive and neural mechanisms of tool use in humans and other vertebrates would increase our understanding of the evolution of physical cognition. An obvious starting point for such a comparison is to focus on non-human primates as there are many examples of tool-use in the wild, such as axing oysters (Gumert et al., 2009), nut-cracking, ant-dipping, and termite fishing, which have been previously reviewed (e.g., van Schaik et al., 1999; Whiten et al., 1999; Emery and Clayton, 2009), and in captivity, even in species that have not been observed using tools in their natural environment (van Lawick-Goodall, 1970; van Schaik et al., 1999). It is interesting to note that despite assumptions that non-human primates show flexible tool use, high cognitive abilities may not be necessary for the performance of tool use in many of the examples given (Penn and Povinelli, 2007). Instead, simpler forms of learning, such as affordance learning (i.e., learning about the use or purpose that an object can have, either directly or through social learning) may be responsible in some species for instrumental object manipulation (Whiten et al., 1999; Martin-Ordas et al., 2008; Gruber et al., 2011).

Keeping in consideration that the development of human high order cognition, including abstract thinking, might have been enhanced once humans used tools as equivalents of body parts (Iriki and Taoka, 2012), it is reasonable to assume that the comparison between humans and other vertebrates' cognitive and neural mechanisms of tool use would increase our understanding of the evolution of physical cognition. Despite the abundant number of tool-use observations and cognitive studies, the study of the neural mechanisms governing tool-use in non-human primates is challenging. PET scanning studies, for example, require primates to remain completely immobile, except for the limb that uses the tool, during data collection to prevent confounding motion artifacts. Accurate measurements can only be achieved by confining the limbs not involved in the studied actions to small spaces and by limiting movement of the subject's head using a custom-made chair. These constraints explain why there is so little non-human research on the matter, and also why most studies have very small sample sizes.

TOOL USE IN MACAQUES

Macaques are of particular importance because they have special neurons that become active both when they see another individual performing an action and when they do the action themselves (Rizzolatti and Craighero, 2004). These neurons, located in premotor cortex F5, are of two types: canonical neurons and mirror neurons. Canonical neurons respond to the presentation of a graspable object or are active when the macaques grasp that object, while mirror neurons respond when the macaque sees object-directed actions (Rizzolatti and Craighero, 2004; Iriki, 2006; although see Hickok, 2010 for a critical review on mirror neuron function). Similar neurons were subsequently found in both humans and birds (Rizzolatti and Craighero, 2004; Welberg, 2008), although it remains unclear whether they exists in birds outside of the context of song learning. The ability to use tools and the presence of mirror neurons in their brains make macaques an interesting model for the study of tool use in vertebrates since mirror neurons are active during object-directed actions. Although there are neurocognitive studies exploring tool-use in a number of other species of non-human primates (e.g., Hopkins et al., 2012, 2017; Phillips and Thompson, 2013; Mayer et al., 2019) we have focused the following section on macaques because our objective is to suggest a possible neural network for tool use in a non-human primate species. For this purpose, using a single genus instead of a combination of findings from multiple species prevents us from generating a misleading network, since different species might differ in many ways, including anatomically, mechanistically, behaviorally, and cognitively.

Macaque Active Brain Areas During Tool Use

Obayashi et al. (2001, 2002, 2004, 2007) performed a series of studies on two awaken-behaving male Japanese monkeys (*Macaca fuscata*). They explored the brain areas that are active during tool use by using PET scans during a task in which the subjects were previously trained to use tools to collect an unreachable food pellet. They used a control task in which the subjects experienced almost the same sensorimotor circumstances as in the experimental task, but without any learning involved (i.e., manipulation of the control apparatus did not result in the macaques learning how to reach the reward, but once the macaques had manipulated the control apparatus the experimenters moved the reward within their reach). In their 2001 study, the subjects had to reach the pellet with one rack. In their 2002 study, they had to poke a pellet with a rack out of a transparent tube and reach it with a second rack. In their 2004 and 2007 studies, the subjects had to obtain an unreachable pellet by manipulating a joystick or a pair of dials, respectively, which controlled the position of a shovel that moved in a two-dimensional space. They found the following active brain areas during tool use in macaques (Obayashi et al., 2001, 2002, 2004, 2007).

Prefrontal Cortex (PFC)

Specifically, area 9/46 seems to be involved in executive functions, since it is active during a sequence of tool combination tasks but not during single tasks (Obayashi et al., 2002). It is also active during abstract actions like remote operations using dials in a set of sequences (Obayashi et al., 2007). Together with the cerebellum, this area is involved in the automatization of learned motor sequences (Obayashi et al., 2004).

Intraparietal Sulcus (IPS)

This region is the area of the brain that creates, stores, and updates the body-image, i.e., the primate's awareness of where its limbs are in space, and what actions they are performing (Obayashi et al., 2001, 2002, 2004, 2007). It has an important role in tool use because it provides the individual with an updated spatial representation of the situation, which is vital for the successful completion of the goal.

Inferior Temporal Cortex

Including the posterior portion of inferior temporal cortices (area TEO). This region is involved in object recognition and memory. Its extensive connections with IPS suggests that it might help this other area in maintaining and manipulating the body-image (Obayashi et al., 2002).

Premotor Cortex

There are two areas of interest within this region: F5 and dorsal premotor cortex (PMD). F5, the area containing mirror neurons in macaques, is involved in the execution of goal-directed manual actions (Rizzolatti and Craighero, 2004). The PMD is involved in planning coordinated activation of muscles and joints to accurately perform desired movements (Obayashi et al., 2001, 2004). Thus, the combined activation of these two areas might be involved in the accurate execution of goal-directed actions performed with tools.

Pre-supplementary Motor Area (pre-SMA)

This area, which receives input from the premotor cortex (especially from F5), was suggested to be involved in the maintenance and updating of the body-image, which would be helpful for the execution of tool-based/use actions and sequential movements (Obayashi et al., 2001, 2004).

Basal Ganglia

It was suggested that the basal ganglia, as well as the IPS bimodal neurons, is involved in the creation and maintenance of the spatiotemporal representation of the hand during tool use (Obayashi et al., 2001).

Cerebellum

The cerebellum was suggested to be involved in the learning processes required for tool use and "reconstruction of the acquired body-image," and "may modulate higher cognitive functions of the executive process as a cerebro-cerebellar loop from an anatomical perspective" (Obayashi et al., 2007).

These brain areas are important for tool use but that is not their only role. They can also be involved in less functional or less goal-directed forms of object manipulation, such as object exploration or object play, often claimed to be associated in the development, evolution and daily expression of tool use (Smith, 1982; Kerney et al., 2017).

Based on the information collected in the aforementioned studies and the available data on macaque brain connections (Schaal, 1999; Hihara et al., 2006; Van Essen and Dierker, 2007; Borra et al., 2008; Smaers et al., 2013; Takada et al., 2013; Rizzolatti et al., 2014), we suggest the following brain network for tool use in macaques, represented in **Figure 3**. The visual input about the tool and the task or problem that needs to be solved using that tool is processed in the visual cortex, which sends this information to the inferior temporal cortex and the IPS. The inferior temporal cortex would process the information related to object recognition, and would then send this information to the IPS. All of this information would be processed in the IPS and a spatial representation of the situation would be created. This information would be sent to area F5 and PMD. These two areas would coordinate the muscles to accurately perform the goal-directed action, and would send this information to the Pre-SMA, which is involved in the execution of sequential movements. The Pre-SMA would update the basal ganglia about the motor action and the basal ganglia would update the Pre-SMA and the PMD about the hand movements during tool use. The information would be sent to a PFC-basal ganglia-cerebellum network, involved in novel motor sequences learning and automatization of learned motor sequences. Finally, the cerebellum would ensure a coordinated motor action. It will be interesting to know whether the same brain regions are involved in these aspects of tool use in other animals or whether they are specific to the macaque brain. Additionally, it will be interesting to study the correlation between the specific patterns of brain activity and the motor movements involved, including in the correlation the levels of cognitive and behavioral control, such as distinctions between flexible and stereotypic tool use, and true versus proto tool use. An exciting first step would be to evaluate these issues in corvids, since like macaques and some other primates, these birds have relatively large brains for their body size, are highly social, have relatively long life spans and are known to use tools for extractive foraging and other problem solving tasks (Emery and Clayton, 2004a) including species that only do so in captivity (Clayton and Emery, 2015). Indeed Clayton and her colleagues have argued for the convergence

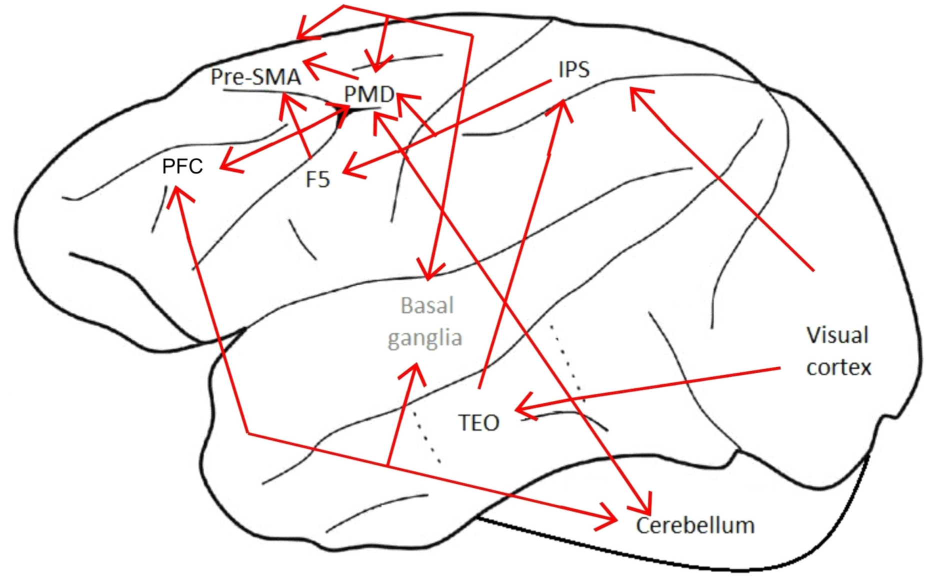


FIGURE 3 | Suggested macaque brain network for tool use. Suggested pathways in red. The names of cortical areas are in black, and subcortical areas in gray. The location of the areas is approximated. IPS, Intraparietal Sulcus; F5, Area F5 of the Premotor Cortex; PFC, Prefrontal Cortex; PMD, Dorsal Premotor Cortex; Pre-SMA, Pre-Supplementary Motor Area; TEO, Posterior Portion of the Inferior Temporal Cortex.

of cognition in primates and corvids (e.g., Emery and Clayton, 2004b; Seed et al., 2009; Taylor and Clayton, 2012; Legg et al., 2017; Baciadonna et al., 2020).

TOOL USE IN CORVIDS

The study of the neural processes of tool use in birds is of great interest, given that some species of birds have shown cognitive abilities and tool-use capability similar to those seen in primates (van Lawick-Goodall, 1970; Lefebvre et al., 2002; Emery and Clayton, 2004a,b, 2009; Clayton and Emery, 2015;, although see a comparison between chimpanzees and New Caledonian crows' tool use in McGrew, 2013). By studying the neural processes of tool use in birds we could identify the brain areas involved in meticulous motor skills and complex cognitive abilities such as problem solving and future planning, which would help us understanding the complexity of the avian brain and its analogies with the mammalian brain (Grodzinski and Clayton, 2010). This is crucial to increasing our understanding of the convergent evolution of physical cognition among vertebrates.

Among birds, corvids are the family with the highest known number of true tool-user species (Lefebvre et al., 2002). For instance, New Caledonian crows manufacture hook tools out of plants to collect food in the wild (Hunt, 1996), with different populations within New Caledonia showing different strategies to manufacture these tools, which is hypothesized to have evolved through a process of cumulative change (Hunt and Gray, 2003). New Caledonian crows also select tools of the right length to achieve their goal (Chappell and Kacelnik, 2002), shape unfamiliar materials to create usable tools for specific tasks (Weir et al., 2002), manufacture tools of the best diameter to achieve a reward (Chappell and Kacelnik, 2004), can infer weights from how objects move under a current of wind (Jelbert et al., 2019), and are capable of spontaneous meta-tool use (using a tool to obtain a tool in order to achieve a goal) (Taylor et al., 2007). Recently, it was found that Hawaiian crows (*Corvus hawaiiensis*), which were considered extinct in the early 2000s and currently live in captivity, are capable of tool use and tool manufacture (Rutz et al., 2016).

Other species of corvids that are not known to be tool-users in the wild are capable of readily using and making tools in captivity. For example, rooks (*Corvus frugilegus*) were able to spontaneously create hook tools from pieces of straight wire and use them to collect food that would otherwise have been out of beak reach (Bird and Emery, 2009a). Furthermore they could use stones to raise the water level to reach a floating reward (Bird and Emery, 2009b), a task also performed by other non-tool-using corvids such as Eurasian jays (*Garrulus glandarius*) (Cheke et al., 2011). Northern blue jays (*Cyanocitta cristata*) also made and used tools to obtain food out of reach (Jones and Kamil, 1973). Furthermore, ravens (*Corvus corax*) demonstrated problem-solving skills by string-pulling to collect food (Heinrich and Bugnyar, 2005), and are capable of planning for events using tools with delays of up to 17 h, including bartering tokens for food rewards (Kabadayi and Osvath, 2017). Many of these behaviors are under cognitive control, as tasks such as wire bending, meta-tool use, inferring the weight of objects, using objects to raise volumes of water, and bartering for tools, require some level of mental processing in order to conceive the problem, to understand the characteristics of the available objects, to plan a solution, and to successfully perform the task to solve the problem. Unfortunately, the neural mechanisms underlying these mental processes are still unknown.

It is for all these reasons that corvids make an excellent model for the study of the neural mechanisms of tool use (Clayton and Emery, 2015; Güntürkün, 2020). Furthermore, they belong to the order Passeriformes, in which the neural mechanisms of song learning in songbirds has been extensively studied, and so there is a wider amount of available information on neuroanatomy and neural networks for corvids compared to other families of tool-user birds (Jarvis et al., 2005). For example, some species of the order Psittaciformes (which includes parrots, lorikeets, and cockatoos) are capable of using and manufacturing tools (Auersperg et al., 2012, 2014; Lambert et al., 2015), and, like corvids, they are also social, big-brained, and vocal learners. However, psittaciformes are evolutionarily less closely related to songbirds, and so the study of their neural circuits has received less attention.

Lefebvre et al. (2002) found that the brain size of true tool-users was significantly larger than that of borderline tool-user species. However, the relatively larger brain of crows and ravens is not correlated to tool use, innovative feeding strategies, and dispersal success (Jønsson et al., 2012). Lefebvre et al. (2002) also found that, after examining the size of four different areas of the telencephalon, the neostriatum [later renamed as nidopallium (Jarvis et al., 2005)] was the only area significantly larger in true tool-using birds compared to borderline tool-users. However, this size difference does not mean that the nidopallium is the only area involved in tool use; rather, it means that the cognitive ability to use tools might depend on the relative size of the nidopallium.

Subsequent studies showed that the neuronal density of the brain of several bird species significantly exceeds the neural densities of many mammals, including primates, of similar brain mass (Olkowicz et al., 2016). This finding could explain why birds, which have relatively small brains compared to mammals, are capable of performing high cognitive tasks. When comparing studies of apes, corvids and pigeons, Güntürkün et al. (2017, p.39) suggest that a "neuronal surplus may translate into faster and more flexible learning, making the acquisition of certain abstract abilities a much easier task". However, a fundamental difference between avian and mammalian brains is that birds lack pyramidal neurons. The inability to create long extensions (i.e., association fibers) that arise from pyramidal neurons means that to exhibit similar behavior to mammals, birds would need more neurons to make the same computations. Additionally, the small size of avian neurons also allows them to have higher neuronal density. An additional hypothesis that we suggest is that the nuclei-organized bird forebrain, unlike the cerebral cortex of primates which is organized in layers, might be computationally more powerful to encode tool-use skills than the isolated neurons in monkeys. To explain the differences in tool use between humans and monkeys, Orban and Caruana (2014) suggested that humans are capable of using tools because we have grouped neurons that respond to tool action observation, unlike some species of monkeys which were unable to learn to use tools. This same hypothesis can therefore be applied to birds, as we have done here. However, further studies are needed to figure out whether tool-use birds own specific neurons that respond to tool action observation.

In mammals, the prefrontal cortex plays a crucial role in problem solving (Mushiake et al., 2009). Given the many examples of problem-solving skills of corvids (e.g., Hunt, 1996; Chappell and Kacelnik, 2002, 2004; Weir et al., 2002; Clayton, 2007; Taylor et al., 2007; Bird and Emery, 2009a,b; Tebbich et al., 2012), it is reasonable to assume that there must be a bird brain area capable of processing information to solve problems in a similar fashion as the mammalian prefrontal cortex does (Güntürkün, 2005; Güntürkün and Bugnyar, 2016). Identifying this possible area is crucial to understand how the cognitive abilities of corvids and apes have evolved via convergent evolution. A specific part of the nidopallium, called the caudolateral nidopallium (NCL) has been suggested as an analogous to the mammalian prefrontal cortex due to its high density of dopaminergic axons (Divac and Mogensen, 1985; Waldmann and Güntürkün, 1993; Güntürkün, 2005) and their function in reward processing (Koenen et al., 2013), prospective processing (reviewed in Colombo et al., 2017), reversal learning, response inhibition and working memory, obtained from studies in pigeons (reviewed in Striedter, 2013). The involvement of NCL in tool use is uncertain, as neurological studies in tool-using birds are rare. However, given the abundant connections of the NCL with other brain areas involved in tool use, it is thought that the NCL is "a critical integrative area for telencephalic sensorimotor pathways" (Striedter, 2013, p.63). Figure 1 in Striedter (2013) highlights the major avian brain areas and circuit diagrams that emphasize the role of NCL in tool use. Neurocognitive studies in crows have shown the involvement of the NCL in cognitive tasks that are important for tool use, such as visual working memory (Veit et al., 2014), associative learning (Veit et al., 2015), abstraction of general principles (Veit and Nieder, 2013), or relative numerosity discrimination (Ditz and Nieder, 2015, 2016). The NCL also shows properties such as flexible neuronal tuning depending on behaviorally relevant tasks (Veit et al., 2017) which is crucial to encode task relevant information.

New Caledonian crows, which are known for their exceptional ability to build and use tools, have an enlarged mesopallium, pallidostriatal complex, septum and tegmentum, compared to three other passeriformes (carrion crows, jays and sparrows) (Mehlhorn et al., 2010). Mehlhorn et al. (2010) suggest that the mesopallium might be required for enhancement of basic tool skills, while the nidopallium, which was also enlarged in this species although not significantly, might have a role in cognitive and motor skills required for basic tool use. They also suggest that the pallidostriatal complex might be important in these birds to learn to manufacture and use tools in novel and familiar situations, and the tegmentum would be involved in the fine motor skills needed for tool manufacture and use, while the septum would integrate several stimuli in order to modulate complex behaviors, which might not be directly involved with tool use. In summary, the study by Mehlhorn et al. (2010) corroborates the findings of previous studies, such as Timmermans et al. (2000), in that the size of the mesopallium is correlated to the feeding innovation rate in birds. However, it is difficult to determine what roles the septum and tegmentum could play in tool use without further investigations, particularly for the tegmentum, which is a large, multifunctional region.

Another region that seems to be important for tool use is the cerebellum. The cerebellum seems to be active during tool use in macaques and humans, as described previously. In birds, it was found that, although the total size of the cerebellum was not significantly different between tool-user and non-tool-user species, the number of folds in the cerebellar cortex was significantly larger in the former (Iwaniuk et al., 2009). It is possible that the increase in the number of folds might have been a way to supply the increased motor, sensory, and cognitive processing demands of the cerebellum of tool-user birds. Furthermore, a recent study in parrots describes a telencephalic-midbrain-cerebellar circuit that resembles the one in primates (Gutiérrez-Ibáñez et al., 2018), which is associated with the evolution of complex cognitive abilities, as described in previous sections. Particularly, the medial spiriform nucleus (SpM), which connects the pallial regions of the telencephalon with the cerebellum in birds, is greatly enlarged in parrots compared with other birds, suggesting that a stronger link between the pallium and cerebellum is associated with cognition (Gutiérrez-Ibáñez et al., 2018). Further studies in birds are needed to clarify the specific role of the cerebellum and the SpM during tool use.

The current literature differs in whether tool-using birds learn to use tools from conspecifics or not. A study on woodpecker finches (*Cactospiza pallida*) found that these birds probably learn to use tools by trial and error during their development (Tebbich et al., 2001). However, a study of Goffin cockatoos (*Cacatua goffiniana*) found social transmission of tool use and tool manufacture in the males of this species (Auersperg et al., 2014), and a study of New Caledonian crows found evidence for probable transmission of tool design between crows (Hunt and Gray, 2003). Thus, although true imitation of tool use has not been observed in birds and, hence, it is not possible to claim that tool-user birds must have mirror neurons, it is important to keep in mind that some of the tool-user species have developed the cognitive abilities to learn specific motor actions by observing others. This feature may imply the development of specific neural characteristics in tool-user birds, which might not be present in species that do not use tools and do not learn from others. In order to answer whether birds have developed neural mechanisms that are similar to humans and non-human primates when learning to use tools, it would be of great interest to study the activity in neurons of the above mentioned brain areas during observations of tool use. However, the size of these brain areas is so large that more research needs to be done before specific nuclei within these areas can be selected for study. The study of the neural processes of tool use in birds is therefore of pivotal interest in order to answer this and other questions, such as what mechanisms are behind the neural development that allows the generation of the same kind of complex behavioral patterns in unrelated species, so we can establish a framework to elucidate the neural mechanisms that supported the convergent evolution of tool use in birds and mammals.

CONCLUSION

The study of the neural processes underlying tool use in humans and non-human primates has received increased attention over the last decade. Early studies in patients suffering from apraxia showed that complex actions that require motor skills and conceptual knowledge, such as tool use, are represented in dissociable neural systems within the left cerebral hemisphere. Two interconnected regions are particularly important, the anterior supramarginal gyrus and the putative human homolog of the anterior intraparietal cortex. However, it is clear that these two regions do not work alone, and a complex neural network is required to use tools. Similarly, in macaques there are also two areas of particular interest: area F5 and the dorsal premotor cortex, both located within the premotor cortex. However, although these two regions seem to be of utmost importance during tool use in macaques, it is evident that, as in humans, tool use requires the activation of a complex network of brain activation, as we have suggested here. We hope that a similar brain network can be elucidated for tool-using corvids in the future, given the abundant evidence of tool use and manufacture in this family of birds. We know that New Caledonian crows and Hawaiian crows can use and manufacture tools, and New Caledonian crows can infer their physical properties (Hunt, 1996; Chappell and Kacelnik, 2002, 2004; Weir et al., 2002; Taylor et al., 2007; Rutz et al., 2016; Jelbert et al., 2019). Furthermore, corvid species that are not tool-users in the wild, such as rooks, Eurasian jays, and northern blue jays, can use and manufacture tools in captivity and use these skills to solve problems (Jones and Kamil, 1973; Seed et al., 2006; Tebbich et al., 2007; Bird and Emery, 2009a,b; Cheke et al., 2011). Additionally, ravens can use tools to obtain unreachable food and use tokens to barter (Heinrich and Bugnyar, 2005; Kabadayi and Osvath, 2017). Despite this plethora of examples, we have yet to discover the neural mechanisms underpinning these behaviors and their cognitive control in corvids. However, we do know about the neural mechanisms of song production and song learning in songbirds, which are birds of the same order as corvids. The knowledge on songbird neuroanatomy may serve as a basis to explore the neural mechanisms of tool use in corvids and to elucidate a neural network underpinning tool-using behavior. We have reviewed the current evidence of several bird brain regions that could be involved in tool use in corvids and that should be the focus of study in future research, such as the NCL, mesopallium, pallidostriatal complex, SpM, cerebellum, and areas of the tegmentum. Understanding the neural processes of tool use in animals other than primates would not only increase our understanding of the evolution of physical cognition in vertebrates, including a better understanding of animal intelligence, but also benefit our society by providing new models with which scientists can work to understand the origins of complex motor skills, and ultimately improve the lives of those affected by motor disabilities.

AUTHOR CONTRIBUTIONS

MC-Á wrote the manuscript, with critical reviews and important discussions and additions by NC. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank Prof. Sarah Woolley for useful comments on the first version of the manuscript, Ben Farrar for helpful comments and discussion of the manuscript, and Camilla Cenni, Prof. Jean-Baptiste Leca, Prof. Michael Colombo, and an undisclosed reviewer for their thorough and useful revisions of the manuscript.

REFERENCES

Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science* 291, 1748–1753. doi: 10.1126/science.1059487

Auersperg, A. M. I., Szabo, B., Von Bayern, A. M. P., and Kacelnik, A. (2012). Spontaneous innovation in tool manufacture and use in a Goffin's cockatoo. *Curr. Biol.* 22, R903–R904. doi: 10.1016/j.cub.2012.09.002

Auersperg, A. M. I., von Bayern, A. M. I., Weber, S., Szabadvari, A., Bugnyar, T., and Kacelnik, A. (2014). Social transmission of tool use and tool manufacture in Goffin cockatoos (*Cacatua goffini*). *Proc. Biol. Sci.* 281:20140972. doi: 10.1098/rspb.2014.0972

Baber, C. (2003). *Cognition and Tool Use: Forms of Engagement in Human and Animal Use of Tools*. London: Taylor & Francis.

Baciadonna, L., Cornero, F., Emery, N. J., and Clayton, N. S. (2020). Convergent evolution of complex cognition: Insights from the field of avian cognition into the study of self-awareness. *Learn. Behav.* doi: 10.3758/s13420-020-00434-5 [Epub ahead of print].

Barton, R. A. (2012). Embodied cognitive evolution and the cerebellum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2097–2107. doi: 10.1098/rstb.2012.0112

Barton, R. A., and Venditti, C. (2014). Report rapid evolution of the cerebellum in humans and other great apes. *Curr. Biol.* 24, 2440–2444. doi: 10.1016/j.cub.2014.08.056

Beck, B. B. (1980). *Animal Tool Behavior*. New York, NY: Garland STPM Press.

Bird, C. D., and Emery, N. J. (2009a). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10370–10375. doi: 10.1073/pnas.0901008106

Bird, C. D., and Emery, N. J. (2009b). Rooks use stones to raise the water level to reach a floating worm. *Curr. Biol.* 19, 1410–1414. doi: 10.1016/j.cub.2009.07.033

Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., et al. (2008). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb. Cortex* 18, 1094–1111. doi: 10.1093/cercor/bhm146

Bronowski, J. (1975). *The Ascent of Man*. Boston, MA: Little, Brown & Co.

Brown, C. (2012). Tool use in fishes. *Fish Fish.* 13, 105–115. doi: 10.1111/j.1467-2979.2011.00451.x

Cantalupo, C., and Hopkins, W. (2010). The cerebellum and its contribution to complex tasks in higher primates: a comparative perspective. *Cortex* 46, 821–830. doi: 10.1016/j.cortex.2009.10.004

Chappell, J., and Kacelnik, A. (2002). Tool selectivity in a non-primate, the New Caledonian crow (*Corvus moneduloides*). *Anim. Cogn.* 5, 71–78. doi: 10.1007/s10071-002-0130-2

Chappell, J., and Kacelnik, A. (2004). Selection of tool diameter by New Caledonian crows *Corvus moneduloides*. *Anim. Cogn.* 7, 121–127. doi: 10.1007/s10071-003-0202-y

Cheke, L. G., Bird, C. D., and Clayton, N. S. (2011). Tool-use and instrumental learning in the Eurasian jay (*Garrulus glandarius*). *Anim. Cogn.* 14, 441–455. doi: 10.1007/s10071-011-0379-4

Clayton, N. (2007). Animal cognition: crows spontaneously solve a metatool task. *Curr. Biol.* 17, 894–895. doi: 10.1016/j.cub.2007.08.028

Clayton, N. S., and Emery, N. J. (2015). Avian models for human cognitive neuroscience: a proposal. *Neuron* 86, 1330–1342. doi: 10.1016/j.neuron.2015.04.024

Colombo, M., Klarer, A., Johnston, M., and Rose, J. (2017). Prospective processing: behavioural and neural evidence. *Jpn. J. Anim. Psychol.* 67, 47–61. doi: 10.2502/janip.67.2.2

Dinets, V., Brueggen, J. C., and Brueggen, J. D. (2015). Crocodilians use tools for hunting. *Ethol. Ecol. Evol.* 27, 74–78. doi: 10.1080/03949370.2013.858276

Ditz, H. M., and Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proc. Natl. Acad. Sci. U.S.A.* 112, 7827–7832. doi: 10.1073/pnas.1504245112

Ditz, H. M., and Nieder, A. (2016). Sensory and working memory representations of small and large numerosities in the crow endbrain. *J. Neurosci.* 36, 12044–12052. doi: 10.1523/JNEUROSCI.1521-16.2016

Divac, I., and Mogensen, J. (1985). The prefrontal "cortex" in the pigeon catecholamine histofluorescence. *Neuroscience* 15, 677–682. doi: 10.1016/0306-4522(85)90069-7

Emery, N. J., and Clayton, N. S. (2004a). "Comparing the complex cognition of birds and primates," in *Comparative Vertebrate Cognition: Are Primates Superior to Non-Primates?* eds L. J. Rogers and G. Kaplan (New York, NY: Kluwer Academic), 3–55. doi: 10.1007/978-1-4419-8913-0_1

Emery, N. J., and Clayton, N. S. (2004b). The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306, 1903–1908. doi: 10.1126/science.1098410

Emery, N. J., and Clayton, N. S. (2009). Tool use and physical cognition in birds and mammals. *Curr. Opin. Neurobiol.* 19, 27–33. doi: 10.1016/j.conb.2009.02.003

Fragaszy, D. M., and Mangalam, M. (2018). Tooling. *Adv .Study Behav.* 50, 177–241. doi: 10.1016/bs.asb.2018.01.001

Frey, S. H. (2008). Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1951–1957. doi: 10.1098/rstb.2008.0008

Gallivan, J. P., McLean, D. A., Valyear, K. F., and Culham, J. C. (2013). Decoding the neural mechanisms of human tool use. *eLife* 2:e00425. doi: 10.7554/eLife.00425

Goodall, J. (1964). Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature* 201, 1264–1266. doi: 10.1038/2011264a0

Grobecker, D. B., and Pietsch, T. W. (1978). Crows use automobiles as nutcrackers. *Auk* 95, 760–761.

Grodzinski, U., and Clayton, N. S. (2010). Problems faced by food-caching corvids and the evolution of cognitive solutions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 977–987. doi: 10.1098/rstb.2009.0210

Gruber, T., Muller, M. N., Reynolds, V., Wrangham, R., and Zuberbu, K. (2011). Community-specific evaluation of tool affordances in wild chimpanzees. *Sci. Rep.* 1:128. doi: 10.1038/srep00128

Gumert, M. D., Kluck, M., and Malaivijitnond, S. (2009). The physical characteristics and usage patterns of stone axe and pounding hammers used by long-tailed macaques in the Andaman sea region of Thailand. *Am. J. Primatol.* 71, 594–608. doi: 10.1002/ajp.20694

Güntürkün, O. (2005). The avian 'prefrontal cortex' and cognition. *Curr. Opin. Behav. Sci.* 14, 686–693. doi: 10.1016/j.conb.2005.10.003

Güntürkün, O. (2020). The surprising power of the avian mind. *Sci. Am.* 322, 48–55.

Güntürkün, O., and Bugnyar, T. (2016). Cognition without Cortex. *Trends Cogn. Sci.* 20, 291–303. doi: 10.1016/j.tics.2016.02.001

Güntürkün, O., Ströckens, F., Scarf, D., and Colombo, M. (2017). Apes, feathered apes, and pigeons: differences and similarities. *Curr. Opin. Behav. Sci.* 16, 35–40. doi: 10.1016/j.cobeha.2017.03.003

Gutiérrez-Ibáñez, C., Iwaniuk, A. N., and Wylie, D. R. (2018). Parrots have evolved a primate-like telenceplalic-midbrain-cerebellar circuit. *Sci. Rep.* 8:9960.

Heinrich, B., and Bugnyar, T. (2005). Testing problem solving in ravens: string-pulling to reach food. *Ethology* 111, 962–976. doi: 10.1111/j.1439-0310.2005.01133.x

Hickok, G. (2010). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243. doi: 10.1162/jocn.2009.21189

Higuchi, S., Imamizu, H., and Kawato, M. (2007). Cerebellar activity evoked by common tool-use execution and imagery tasks: an fMRI study. *Cortex* 43, 350–358. doi: 10.1016/S0010-9452(08)70460-X

Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., Ojima, H., Obayashi, S., et al. (2006). Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. *Neuropsychologia* 44, 2636–2646. doi: 10.1016/j.neuropsychologia.2005.11.020

Hihara, S., Yamada, H., Iriki, A., and Okanoya, K. (2003). Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. *Neurosci. Res.* 45, 383–389. doi: 10.1016/s0168-0102(03)00011-7

Hodos, W. (1987). Animal general intelligence: an idea ahead of its time. *Behav. Brain Sci.* 10, 668–668. doi: 10.1017/S0140525X00055138

Hopkins, W. D., Meguerditchian, A., Coulon, O., Misiura, M., Pope, S., Mareno, M. C., et al. (2017). Motor skill for tool-use is associated with asymmetries in Broca ' s area and the motor hand area of the precentral gyrus in chimpanzees (*Pan troglodytes*). *Behav. Brain Res.* 318, 71–81. doi: 10.1016/j.bbr.2016.10.048

Hopkins, W. D., Russell, J. L., and Schaeffer, J. A. (2012). The neural and cognitive correlates of aimed throwing in chimpanzees: a magnetic resonance image and behavioural study on a unique form of social tool use. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 37–47. doi: 10.1098/rstb.2011.0195

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature* 379, 249–251. doi: 10.1038/379249a0

Hunt, G. R., and Gray, R. D. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proc. Biol. Sci.* 270, 867–874. doi: 10.1098/rspb.2002.2302

Hunt, G. R., and Gray, R. D. (2004). The crafting of hook tools by wild New Caledonian crows. *Biol. Lett.* 271, 88–90. doi: 10.1098/rsbl.2003.0085

Hunt, G. R., Gray, R. D., and Taylor, A. H. (2013). "Why is tool use rare in animals?" in *Tool Use in Animals: Cognition and Ecology*, eds C. M. Sanz, J. Call, and C. Boesch (Cambridge: Cambridge University Press), 89–118. doi: 10.1017/cbo9780511894800.007

Iriki, A. (2006). The neural origins and implications of imitation, mirror neurons and tool use. *Curr. Opin. Neurobiol.* 16, 660–667. doi: 10.1016/j.conb.2006.10.008

Iriki, A., and Taoka, M. (2012). Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 10–23. doi: 10.1098/rstb.2011.0190

Iwaniuk, A. N., Lefebvre, L., and Wylie, D. R. (2009). The comparative approach and brain–behaviour relationships: a tool for understanding tool use. *Can. J. Exp. Psychol.* 63, 150–159. doi: 10.1037/a0015678

Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6, 151–159. doi: 10.1038/nrn1606

Jelbert, S. A., Miller, R., Schiestl, M., Boeckle, M., Cheke, L. G., Gray, R. D., et al. (2019). New Caledonian crows infer the weight of objects from observing their movements in a breeze. *Proc. Biol. Sci.* 286:20182332. doi: 10.1098/rspb.2018.2332

Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. *Trends Cogn. Sci.* 8, 71–78. doi: 10.1016/j.tics.2003.12.002

Jones, T. B., and Kamil, A. C. (1973). Tool-making and tool-using in the Northern blue jay. *Science* 180, 1076–1078. doi: 10.1126/science.180.4090.1076

Jønsson, K. A., Fabre, P., and Irestedt, M. (2012). Brains, tools, innovation and biogeography in crows and ravens. *BMC Evol. Biol.* 12:72. doi: 10.1186/1471-2148-12-72

Kabadayi, C., and Osvath, M. (2017). Ravens parallel great apes in flexible planning for tool-use and bartering. *Science* 357, 202–204. doi: 10.1126/science.aam8138

Kerney, M., Smaers, J. B., Schoenemann, P. T., and Dunn, J. C. (2017). The coevolution of play and the cortico-cerebellar system in primates. *Primates* 58, 485–491. doi: 10.1007/s10329-017-0615-x

Koenen, C., Millar, J., and Colombo, M. (2013). How bad do you want it? Reward modulation in the Avian Nidopallium Caudolaterale. *Behav. Neurosci.* 127, 544–554. doi: 10.1037/a0033551

Lambert, M. L., Seed, A. M., and Slocombe, K. E. (2015). A novel form of spontaneous tool use displayed by several captive greater vasa parrots (*Coracopsis vasa*). *Biol. Lett.* 11:20150861. doi: 10.1098/rsbl.2015.0861

Lefebvre, L., Nicolakakis, N., and Boire, D. (2002). Tools and brains in birds. *Behaviour* 139, 939–973. doi: 10.1163/156853902320387918

Legg, E. W., Ostojic, L., and Clayton, N. S. (2017). "Routes to the convergent evolution of cognition," in *Routledge Handbook of Evolution and Philosophy*, ed. R. Joyce (Abingdon: Routledge), 237–251. doi: 10.4324/9781315764863-17

Lewis, J. W. (2006). Cortical networks related to human use of tools. *Neuroscientist* 12, 211–231. doi: 10.1177/1073858406288327

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–656. doi: 10.1017/S0140525X00054984

Mann, J., Sargeant, B. L., Watson-Capps, J. J., Gibson, Q. A., Heithaus, M. R., Connor, R. C., et al. (2008). Why do dolphins carry sponges? *PLoS One* 3:e3868. doi: 10.1371/journal.pone.0003868

Martin-Ordas, G., Call, J., and Colmenares, F. (2008). Tubes, tables and traps: great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Anim. Cogn.* 11, 423–430. doi: 10.1007/s10071-007-0132-1

Mayer, A., Lewenfus, G., Bittencourt-navarrete, R. E., Clasca, F., and Guedes, J. (2019). Thalamic inputs to posterior parietal cortical areas involved in skilled forelimb movement and tool use in the capuchin monkey. *Cereb. Cortex* 29, 5098–5115. doi: 10.1093/cercor/bhz

McGrew, W. C. (2013). Is primate tool use special? Chimpanzee and new caledonian crow compared. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120422. doi: 10.1098/rstb.2012.0422

Mehlhorn, J., Hunt, G. R., Gray, R. D., Rehkämper, G., and Güntürkün, O. (2010). Tool-making New Caledonian crows have large associative brain areas. *Brain Behav. Evol.* 75, 63–70. doi: 10.1159/000295151

Mushiake, H., Sakamoto, K., Saito, N., Inui, T., Aihara, K., and Tanji, J. (2009). Involvement of the prefrontal cortex in problem solving. *Int. Rev. Neurobiol.* 85, 1–11. doi: 10.1016/s0074-7742(09)85001-0

Oakley, K. P. (1972). *Man the Tool-Maker*, 6th Edn. London: British Museum (Natural History).

Obayashi, S., Matsumoto, R., Suhara, T., Nagai, Y., and Iriki, A. (2007). Functional organization of monkey brain for abstract operation. *Cortex* 43, 389–396. doi: 10.1016/s0010-9452(08)70464-7

Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., et al. (2001). Functional brain mapping of monkey tool use. *Neuroimage* 14, 853–861. doi: 10.1006/nimg.2001.0878

Obayashi, S., Suhara, T., Nagai, Y., Maeda, J., Hihara, S., and Iriki, A. (2002). Macaque prefrontal activity associated with extensive tool use. *Neuroreport* 13, 2349–2354. doi: 10.1097/01.wnr.0000044228.79663.ca

Obayashi, S., Suhara, T., Nagai, Y., Okauchi, T., Maeda, J., and Iriki, A. (2004). Monkey brain areas underlying remote-controlled operation. *Eur. J. Neurosci.* 19, 1397–1407. doi: 10.1111/j.1460-9568.2004.03200.x

Olkowicz, S., Kocourek, M., Luèan, R. K., Porteš, M., Fitch, W. T., Herculano-Houzel, S., et al. (2016). Birds have primate-like numbers of neurons in the forebrain. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7255–7260. doi: 10.1073/pnas.1517131113

Orban, G. A., and Caruana, F. (2014). The neural basis of human tool use. *Front. Psychol.* 5:310. doi: 10.3389/fpsyg.2014.00310

Osiurak, F., and Badets, A. (2016). Tool use and affordance: manipulation-based versus reasoning-based approaches. *Psychol. Rev.* 123, 534–568. doi: 10.1037/rev0000027

Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G., et al. (2009). The representation of tool use in humans and monkeys: common and uniquely human features. *J. Neurosci.* 29, 11523–11539. doi: 10.1523/JNEUROSCI.2040-09.2009

Peeters, R. R., Rizzolatti, G., and Orban, G. A. (2013). Functional properties of the left parietal tool use region. *Neuroimage* 78, 83–93. doi: 10.1016/j.neuroimage.2013.04.023

Penn, D. C., and Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: a comparative, critical review. *Annu. Rev. Psychol.* 58, 97–118. doi: 10.1146/annurev.psych.58.110405.085555

Phillips, K. A., and Thompson, C. R. (2013). Hand preference for tool-use in capuchin monkeys (*Cebus apella*) is associated with asymmetry of the primary motor cortex. *Am. J. Primatol.* 75, 435–440. doi: 10.1002/ajp.22079

Pierce, J. D. J. (1986). A review of tool use in insects. *Florida Entomol.* 69, 95–104. doi: 10.2307/3494748

Reynaud, E., Lesourd, M., Navarro, J., and Osiurak, F. (2016). On the neurocognitive origins of human tool use: a critical review of neuroimaging data. *Neurosci. Biobehav. Rev.* 64, 421–437. doi: 10.1016/j.neubiorev.2016.03.009

Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., and Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706. doi: 10.1152/physrev.00009.2013

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Root-Bernstein, M., Narayan, T., Cornier, L., and Bourgeois, A. (2019). Context-specific tool use by *Sus cebifrons*. *Mamm. Biol.* 98, 102–110. doi: 10.1016/j.mambio.2019.08.003

Rutz, C., Klump, B. C., Komarczyk, L., Leighton, R., Kramer, J., Wischnewski, S., et al. (2016). Discovery of species-wide tool use in the Hawaiian crow. *Nature* 537, 403–407. doi: 10.1038/nature19103

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends Cogn. Sci.* 3, 233–242. doi: 10.1016/s1364-6613(99)01327-3

Seed, A., Emery, N., and Clayton, N. (2009). Intelligence in corvids and apes: a case of convergent evolution? *Ethology* 115, 401–420. doi: 10.1111/j.1439-0310.2009.01644.x

Seed, A. M., Tebbich, S., Emery, N. J., and Clayton, N. S. (2006). Investigating physical cognition in rooks, *Corvus frugilegus*. *Curr. Biol.* 16, 697–701. doi: 10.1016/j.cub.2006.02.066

Smaers, J. B., Steele, J., Case, C. R., and Amunts, K. (2013). Laterality and the evolution of the prefronto-cerebellar system in anthropoids. *Ann. N.Y. Acad. Sci.* 1288, 59–69. doi: 10.1111/nyas.12047

Smith, P. K. (1982). Does play matter? Functional and evolutionary aspects of animal and human play. *Behav. Brain Sci.* 5, 139–155. doi: 10.1017/S0140525X0001092X

St Amant, R., and Horton, T. E. (2008). Revisiting the definition of animal tool use. *Anim. Behav.* 75, 1199–1208. doi: 10.1016/j.anbehav.2007.09.028

Steele, J., Ferrari, P. F., and Fogassi, L. (2012). From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 4–9. doi: 10.1098/rstb.2011.0295

Striedter, G. F. (2013). Bird brains and tool use: beyond instrumental conditioning. *Brain Behav. Evol.* 82, 55–67. doi: 10.1159/000352003

Takada, M., Hoshi, E., Saga, Y., Inoue, K., Miyachi, S., Hatanaka, N., et al. (2013). "Organization of two cortico-basal ganglia loop circuits that arise from distinct sectors of the monkey dorsal premotor cortex," in *Basal Ganglia - An Integrative View*, ed. F. A. Barrios (Rijeka: InTech), 103–116. doi: 10.5772/54822

Taylor, A. H., and Clayton, N. S. (2012). Evidence from convergent evolution and causal reasoning suggests that conclusions on human uniqueness may be premature. *Behav. Brain Sci.* 35, 241–242. doi: 10.1017/s0140525x11001956

Taylor, A. H., Elliffe, D., Hunt, G. R., and Gray, R. D. (2010). Complex cognition and behavioural innovation in New Caledonian crows. *Proc. Biol. Sci.* 277, 2637–2643. doi: 10.1098/rspb.2010.0285

Taylor, A. H., Hunt, G. R., Holzhaider, J. C., and Gray, R. D. (2007). Spontaneous metatool use by New Caledonian crows. *Curr. Biol.* 17, 1504–1507. doi: 10.1016/j.cub.2007.07.057

Taylor, A. H., Hunt, G. R., Medina, F. S., and Gray, R. D. (2009). Do New Caledonian crows solve physical problems through causal reasoning?? *Proc. Biol. Sci.* 276, 247–254. doi: 10.1098/rspb.2008.1107

Tebbich, S., Seed, A. M., Emery, N. J., and Clayton, N. S. (2007). Non-tool-using rooks, *Corvus frugilegus*, solve the trap-tube problem. *Anim. Cogn.* 10, 225–231. doi: 10.1007/s10071-006-0061-4

Tebbich, S., Stankewitz, S., and Teschke, I. (2012). The relationship between foraging, learning abilities and neophobia in two species of Darwin's finches. *Ethology* 118, 135–146. doi: 10.1111/j.1439-0310.2011.02001.x

Tebbich, S., Taborsky, M., Fessl, B., and Blomqvist, D. (2001). Do woodpecker finches acquire tool-use by social learning? *Proc. Biol. Sci.* 268, 2189–2193. doi: 10.1098/rspb.2001.1738

Timmermans, S., Lefebvre, L., Boire, D., and Basu, P. (2000). Relative size of the hyperstriatum ventrale is the best predictor of feeding innovation rate in birds. *Brain Behav. Evol.* 56, 196–203. doi: 10.1159/000047204

Van Essen, D. C., and Dierker, D. L. (2007). Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron* 56, 209–225. doi: 10.1016/j.neuron.2007.10.015

van Lawick-Goodall, J. (1970). Tool-using in primates and other vertebrates. *Adv. Study Behav.* 3, 195–249. doi: 10.1016/s0065-3454(08)60157-6

van Schaik, C. P., Deaner, R. O., and Merrill, M. Y. (1999). The conditions for tool use in primates: implications for the evolution of material culture. *J. Hum. Evol.* 36, 719–741. doi: 10.1006/jhev.1999.0304

Veit, L., Hartmann, K., and Nieder, A. (2014). Neuronal correlates of visual working memory in the corvid endbrain. *J. Neurosci.* 34, 7778–7786. doi: 10.1523/JNEUROSCI.0612-14.2014

Veit, L., Hartmann, K., and Nieder, A. (2017). Spatially tuned neurons in Corvid Nidopallium Caudolaterale signal target position during visual search. *Cereb. Cortex* 27, 1103–1112. doi: 10.1093/cercor/bhv299

Veit, L., and Nieder, A. (2013). Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nat. Commun.* 4:2878. doi: 10.1038/ncomms3878

Veit, L., Pidpruzhnykova, G., and Nieder, A. (2015). Associative learning rapidly establishes neuronal representations of upcoming behavioral choices in crows. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15208–15213. doi: 10.1073/pnas.1509760112

Waldmann, C., and Güntürkün, O. (1993). The dopaminergic innervation of the pigeon caudolateral forebrain: immunocytochemical evidence for a "prefrontal cortex" in birds? *Brain Res.* 600, 225–234. doi: 10.1016/0006-8993(93)91377-5

Weaver, A. H. (2005). Reciprocal evolution of the cerebellum and neocortex in fossil humans. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3576–3580. doi: 10.1073/pnas.0500692102

Weir, A. A. S., Chappell, J., and Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science* 297:981. doi: 10.1126/science.1073433

Welberg, L. (2008). Mirror neurons: singing in the brain. *Nat. Rev. Neurosci.* 9:163. doi: 10.1038/nrn2340

Whiten, A., Goodall, J., Mcgrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature* 399, 682–685.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 14 October 2020
doi: 10.3389/fpsyg.2020.568049

Arthropod Intelligence? The Case for *Portia*

Fiona R. Cross[1,2]*, Georgina E. Carvell[3†], Robert R. Jackson[1,2] and Randolph C. Grace[3]

[1]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand, [2]International Centre of Insect Physiology and Ecology, Mbita Point, Kenya, [3]School of Psychology, Speech and Hearing, University of Canterbury, Christchurch, New Zealand

Macphail's "null hypothesis," that there are no differences in intelligence, qualitative, or quantitative, between non-human vertebrates has been controversial. This controversy can be useful if it encourages interest in acquiring a detailed understanding of how non-human animals express flexible problem-solving capacity ("intelligence"), but limiting the discussion to vertebrates is too arbitrary. As an example, we focus here on *Portia*, a spider with an especially intricate predatory strategy and a preference for other spiders as prey. We review research on pre-planned detours, expectancy violation, and a capacity to solve confinement problems where, in each of these three contexts, there is experimental evidence of innate cognitive capacities and reliance on internal representation. These cognitive capacities are related to, but not identical to, intelligence. When discussing intelligence, as when discussing cognition, it is more useful to envisage a continuum instead of something that is simply present or not; in other words, a continuum pertaining to flexible problem-solving capacity for "intelligence" and a continuum pertaining to reliance on internal representation for "cognition." When envisaging a continuum pertaining to intelligence, Daniel Dennett's notion of four Creatures (Darwinian, Skinnerian, Popperian, and Gregorian) is of interest, with the distinction between Skinnerian and Popperian Creatures being especially relevant when considering *Portia*. When we consider these distinctions, a case can be made for *Portia* being a Popperian Creature. Like Skinnerian Creatures, Popperian Creatures express flexible problem solving capacity, but the manner in which this capacity is expressed by Popperian Creatures is more distinctively cognitive.

Keywords: arthropod, cognition, intelligence, problem solving, representation, spider

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Martin Giurfa,
UMR5169 Centre de Recherches sur la Cognition Animale (CRCA), France
Michael Colombo,
University of Otago, New Zealand

***Correspondence:**
Fiona R. Cross
fiona.r.cross@gmail.com

†Deceased

Specialty section:
This article was submitted to Comparative Psychology, a section of the journal Frontiers in Psychology

Received: *31 May 2020*
Accepted: *07 September 2020*
Published: *14 October 2020*

Citation:
Cross FR, Carvell GE, Jackson RR and Grace RC (2020) Arthropod Intelligence? The Case for Portia. Front. Psychol. 11:568049. doi: 10.3389/fpsyg.2020.568049

INTRODUCTION

For over a century, variation in human intelligence has been a topic of intensive study and debate (Wasserman, 2012) and, ever since Darwin, questions about the intelligence of non-human animals have also generated heated discussion and controversy. In an attempt to cast light on the evolution of intelligence, Macphail (1985, 1987; see also Macphail and Bolhuis, 2001) proposed what he called his "null hypothesis," that there are no differences in intelligence, qualitative, or quantitative, between non-human vertebrates. On the basis of the evidence he considered, Macphail argued there was no compelling reason to reject this hypothesis because, as he saw it, reported differences between species on intelligence-related tasks could be attributed to "contextual variables." In other words, he argued that if two species had been given the same problem to solve, but only one of these species succeeded or performed better than the other, this may have reflected a difference in intelligence, or it may have reflected an unrelated difference between the two species, such as in motivational factors, sensory systems, or other variables unrelated to intelligence (the "context").

As illustrated by the commentaries accompanying Macphail (1987), this null hypothesis has been heavily criticized, with further reservations coming from subsequent findings. For example, non-human primates, but not pigeons, more rapidly solve one-dimensional, rule-based visual categorization tasks in which selective attention provides an advantage, compared with the two-dimensional integration tasks in which it does not (Smith et al., 2012). It is hard to see how contextual variables could account for this distinct difference between two species because, in these experiments, the same stimuli were used with each species and the main difference was only in the required responses, with monkeys having to touch one of two boxes on a screen and with pigeons having to peck one of two keys (Smith et al., 2010, 2011). As a more recent example pertaining to a similar category learning task, the performance of rats was intermediate to the performance of pigeons and non-human primates (Broschard et al., 2019).

On the whole, numerous reservations and rebuttals pertaining to Macphail's null hypothesis seem valid, but there may be an indirect way in which this hypothesis has been useful because it encourages a comparative perspective and underscores the need to specify what "intelligence" means. Shortly before Macphail proposed his null hypothesis, Jensen (1980) had proposed a continuum of "intelligence," where, at the bottom of this continuum, we find single-cell protozoans, before moving up to invertebrates, then up to lower vertebrates, then mammals, and finally reaching humans at the pinnacle (see Macphail, 1985). Macphail left invertebrates out when proposing his null hypothesis and this seems consistent with a widespread intuition (conventional wisdom) about invertebrates being limited to behavior that barely, if at all, qualifies as intelligent. We tend to associate intelligence with brains and there are often major differences in size between vertebrate and invertebrate brains. Octopuses may be an exception, but most invertebrates are arthropods (insects, spiders and their relatives) and it can easily seem a foregone conclusion that insects and spiders are just too small-brained to be intelligent. However, recent research on insects (Dyer, 2012; Giurfa, 2013, 2015) and spiders (e.g., Jakob et al., 2011) challenges the convention of assuming severe constraints on the expression of cognition by small animals. As cognition tends to be associated with intelligence, including arthropods can serve as a step toward taking a broader view of Macphail's null hypothesis in terms of *scope* and *depth*.

Here, we will focus on *Portia*, a genus of jumping spiders (family Salticidae). These arthropods have unique, complex eyes and an exceptional ability for seeing detail in visual objects, making them especially suitable experimental subjects in research on behavior (Harland et al., 2012; Land and Nilsson, 2012), including behavior related to intelligence. With *Portia* in particular, we find flexibility and problem-solving capacities at a level that fits comfortably with the notion of what qualifies as "intelligence" when found in vertebrates. However, for a better focus, we might need at least a rough definition of what "intelligence" is.

For a definition, we can turn to Burkart et al. (2017, p. 2), who characterized non-human intelligence as an "individual's ability to acquire new knowledge from interactions with the physical or social environment, use this knowledge to organize effective behavior in both familiar and novel contexts, and engage with and solve novel problems." Their emphasis on flexibility and novelty highlights a key aspect of intelligence, this being that it applies to *domain-general* rather than *domain-specific* abilities.

When making comparisons in intelligence, Burkart et al. (2017) referred to between-species comparisons as differences in *G*, and within-species comparisons as differences in *g*. In this context, Macphail had mainly considered *G*, and an extension of his null hypothesis would predict that there are no differences in *g* as well as no differences in *G*. Yet, as Burkart et al. (2017) pointed out, there is considerable evidence from research on rodents (mice and rats) and non-human primates of differences in *g* and *G*. As an example of *g*, Matzel et al. (2003) tested 56 genetically-diverse mice in five different learning tasks (associative fear conditioning, operant avoidance, path integration, discrimination, and spatial navigation) and found that individual performances were positively correlated across tasks, with a single factor accounting for 38% of the total variance. As another example, Deaner et al. (2006) conducted a meta-analysis of non-human primate cognition studies using nine different experimental tasks, and found evidence for differences across genera, with great apes performing better than prosimians, New World monkeys, Old World monkeys, and lesser apes, suggesting differences in *G*.

Higher values of *G* have also been found to be correlated with larger brain size in non-human primates (Reader and Laland, 2002; Deaner et al., 2007) and, as Burkart et al. (2017) pointed out, this seems to present us with an evolutionary puzzle of general intelligence. It would seem that, when higher values of *G* evolve, we should find evidence of more domain-general intelligence compensating for the costs in resources and energy from growing bigger brains. Burkart et al. (2017, p. 20) refer to instances of domain-specific intelligence as "dedicated cognitive adaptations in response to recurrent fitness-relevant problems," which seems to correspond at least roughly to the notion Fodor (1983) had of modular minds. Returning to Macphail (1987), the null hypothesis seems to imply that we should demand especially strong evidence before accepting conclusions about animals relying on domain-specific intelligence and, when this strong evidence is not delivered, that we should accept a null hypothesis of domain-general intelligence. Yet, when we consider the "puzzle" related to trade-offs (Burkart et al., 2017), maybe the null hypothesis should pertain to domain-specific, not domain-general, intelligence.

Compared with most vertebrates, salticids, like most arthropods, have brains that would comfortably fit on pinheads (Harland and Jackson, 2000). Yet, despite their tiny brains, salticids often display behavior that normally qualifies as "intelligent" when displayed by vertebrates. This makes it all the more important to understand how these abilities might have evolved.

As a step toward this goal, we will first address what we mean by "cognition" and "intelligence." Next, we will review evidence for intelligent behavior in salticids, especially *Portia*, by focusing on experimental tasks involving pre-planned detours, expectancy violation, and novel problem solving. Based on the available evidence, we argue that *Portia* is an example of what Dennett (1996) called a Popperian Creature. Lastly, we consider the possible implications of research with arthropods for understanding the evolution of intelligence in non-human animals, and we discuss directions for future research.

INTELLIGENCE ON A CONTINUUM

It may be a forlorn hope that any strict formal definition of "intelligence" will ever be widely accepted (Wasserman, 2012), but we should say something about the way we think of "intelligence" because, otherwise, we risk talking past each other. We will also discuss the distinction between intelligence and cognition, but we will begin here with Dennett's (1995, 1996) informal notion of four Creatures (Darwinian, Skinnerian, Popperian, and Gregorian). When referring to these Creatures, we should acknowledge that we are doing a lot of simplifying because, with real organisms, we expect that the boundaries between Creature types will blur and that, when considering any one type, we can expect a continuum instead of a distinct category. As another simplification, we can envisage each of the four Creatures as having proficiency at responding to problems using solutions derived by trial-and-error.

A Darwinian Creature relies on a "hard-wired approach" (Geffner, 2013), with the animal's "innate" or "instinctive" (Lorenz, 1965) plans and solutions to problems being derived by natural selection, a trial-and-error process acting over evolutionary time (e.g., see Catania, 2010). A Darwinian Creature's solutions to problems may be "clever," but this is not the same as attributing to the individual Darwinian Creature the cleverness involved in deriving these solutions to problems. The situation is different with the Skinnerian Creature because, by trial-and-error learning of the relationship between responses and consequences in its own lifetime, the individual Skinnerian Creature derives its own individual solutions to problems (Domjan, 2010).

Popperian Creatures are distinctly different because, instead of solving problems by physically acting in the environment in real time, they derive solutions to problems ahead of time by formulating plans and then by acting on them (Dennett, 1996). As Geffner (2013, p. 341) put it, the Popperian animal is "thinking before acting." Gregorian Creatures go beyond this by making use of mind tools for solving problems, with this being most prominently seen with human verbal language (Dennett, 1995, 1996).

When we consider Skinnerian and Popperian Creatures as falling on continuums, we can indicate where intelligence and cognition become prominent. When looking for evidence of "intelligence," the relevant continuum pertains to an individual's proficiency at flexible problem solving and, following Grush (1997), we envisage "Popperian" as having crossed a threshold into the realm of genuine cognition because these are animals that rely on representations when deriving solutions to specific problems.

At the most basic level, a "representation" can be thought of as something that stands for something else (Webb, 2012) or, more accurately, something that is used to stand in for something else (Grush, 1997). Gallistel (1990a) and later Gallistel and King (2009) emphasized a functional equivalence between internal representations and relevant entities or events in the outside world, with representations serving as theoretical constructs that have a role in cognitive science analogous to the way homomorphism works in mathematics. The emphasis on representation as being critical to cognition is important because this is a step toward understanding how a Popperian Creature can interface with the outside world in a way that goes beyond stimulus-and-response. This allows for foresight, predicting outcomes of plans and acting in ways that flexibly anticipate what is likely to be beneficial rather than relying more strictly on stamped in solutions to problems.

This perspective might make it easier to break free from an intuition that there must be a tight relationship between brain size and intelligence. Where invertebrates fit on a continuum of intelligence is an empirical question and, as Chittka and Niven (2009) illustrated with examples from social insects, often the answer may be considerably different from what is expected. Bees and ants defy the common sense notion that being a mammal or a bird with a large brain is a prerequisite for a substantial level of intelligence. For spiders, we find comparable defiance of common sense among the species in the salticid genus *Portia*.

THE SALTICID BRAIN

As body size gets smaller, it is inevitable that the maximum number of neurons that can be housed in a brain will also get smaller because there is a limit on how small neurons can be and still remain functional (Faisal et al., 2005; Niven and Farris, 2012; Niven and Chittka, 2016). Of course, nobody has ever literally counted the number of neurons in a salticid's brain, but the estimated number even for much larger spiders is in only the tens of thousands (Babu, 1975; Babu and Barth, 1984). Saying "only" is relevant when we compare this to the brains of large vertebrates. For example, elephant brains are estimated to have 100,000,000,000 neurons and human brains 85,000,000,000 (Eberhard and Wcislo, 2011), but the possibility of spider-sized brains having such a large number is ruled out due to limitations on the extent to which neurons can be miniaturized. This also leads to vast differences in the possible numbers of dendritic connections between neurons. In the human brain, for example, individual neurons often have thousands or tens of thousands of dendritic connections to other neurons (Edelman, 1998), with these being numbers that rival the total number of neurons in a spider's entire brain.

Interesting possibilities arise when brains are large and, although their focus was not specifically on animal intelligence, Eberhard and Wcislo (2011) suggested that qualitatively different brain processes might be found depending on whether the brain is that of a spider or a primate. In primates, for example, brain functioning can be based on recurrent pathways involving huge populations of neurons and their dendritic connections on a scale that has no parallel in spider-sized brains (Eberhard and Wcislo, 2011). It seems inevitable that, for spider-sized brains, the level at which intelligence-related processes take place will be more at a neuron-to-neuron level instead of at the level of recurrent pathways in large populations of neurons.

Understanding precisely how this and other size-related consequences might influence the expression of animal intelligence seems particularly important when discussing Macphail's null hypothesis. This hypothesis challenges us to find distinctive instances of different animals using qualitatively different intelligence-related processes. It is in this context that research on *Portia* may become especially relevant.

THE SALTICID SPIDER *PORTIA*

Found in Africa, Asia, and Australia, 17 species from this genus currently have names and formal taxonomic description (Platnick, 2020). Most of what we know from using *Portia* in research pertaining to intelligence has come from five of these: *Portia africana* and *Portia schultzi* from East Africa, *Portia labiata* and *Portia occidentalis* from Asia, and *Portia fimbriata* from Australia. There are over 6,000 salticid species (Maddison, 2015), with little known about the behavior of most of them, but it seems likely that most salticid species prey primarily on insects, which they capture without the assistance of a web (Jackson and Pollard, 1996). *Portia* cannot be characterized so simply because, besides capturing prey away from webs, *Portia* also builds prey-capture webs and also invades the webs of other spiders where it uses many different prey-specific prey-capture tactics (Harland and Jackson, 2004). The tactics used while in other spiders' webs include *Portia* using its appendages to move and tense web silk, thereby making signals with which to control the resident spiders' behavior (Jackson and Cross, 2013).

In and out of webs, *Portia* has an active preference for spiders instead of insects as prey. Besides being potential prey, another spider is, for *Portia*, a potential predator and the risk of the hunter becoming the hunted may have favored reliance on especially flexible prey-capture methods that can be finely tuned to the particular spider being pursued (Jackson and Cross, 2013). Flexibility and fine tuning includes more than *Portia* making web signals and, of particular interest here, it extends to making strategic prey-capture plans ahead of time (Jackson and Cross, 2011).

Something else needs to be emphasized. Learning is typically emphasized when animal intelligence is discussed, often almost as though, by definition, intelligence and learning have to go together (e.g., see Burkart et al., 2017). Yet very little of the research on *Portia* pertains specifically to practice and having prior experience with the problems to be solved. Experiments have repeatedly demonstrated that *Portia* expresses the behavior we envisage as being intelligent without needing to rely on prior personal experience with particular prey or with particular environmental situations for acquiring the information critical to solving problems. On this basis, we conclude that *Portia*'s behavior in experiments is innate (see O'Neill, 2015), but being innate does not mean inflexible or non-intelligent. This is something we will illustrate by reviewing research based on using three particular experimental approaches in which training and learning are not part of the procedure.

PRE-PLANNED DETOURS

Part of *Portia*'s strategy when preying on other spiders is often to adopt an indirect path (i.e., a detour) leading to an optimal location from which to launch an attack (Jackson and Wilcox, 1993), and findings from laboratory experiments imply that *Portia* can make strategic detouring decisions ahead of time. This includes decisions related to the risk of being attacked by the prey spider (Jackson et al., 2002), decisions related to whether a more direct path is available (Cross and Jackson, 2019) and choosing between two indirect paths, with only one leading to prey (Tarsitano and Jackson, 1997; Cross and Jackson, 2016). In each of these studies, the apparatus and the testing protocol were designed with respect to a specific objective of looking for evidence of planning.

We will focus on Cross and Jackson (2016) here. In this study (**Figure 1**), each trial began with *Portia* (the "test spider") on the top of a tower from which it could view two displays and two pathways, with one pathway leading to a display where there were lures made from prey spiders and the other pathway leading to a control display (dead leaves that were similar in size to the lures). The displays were out of reach from the tower. Moreover, the tower and pathways were on a platform which, in turn, sat in a shallow pan of water, and *Portia* is averse to getting wet. This is important because it meant that the only way *Portia* could reach the lure display without getting wet was to first walk down from the top of the tower to the platform, walk directly away from the location of the lures to arrive at the pole where the correct pathway began (i.e., the pathway that led to the lures) and then continue along this pathway to the display. *Portia* needed to plan ahead because, once it left the tower, the lures and control leaves were removed from the displays, meaning that the test spider could no longer navigate on the basis of seeing the location of the lures. Yet, in this study (Cross and Jackson, 2016), 251 test spiders chose the correct pathway and only 15 test spiders chose the incorrect pathway.

Training and learning were not part of this experimental design; each test spider was used in a single trial and test spiders had no prior experience with the apparatus or testing protocol, so they could not use trial and error. It is a common fallacy to assume that "innate" must imply "inflexible." However, these experiments, having been specifically designed as ways to look for innate capacity to plan detours ahead of time in

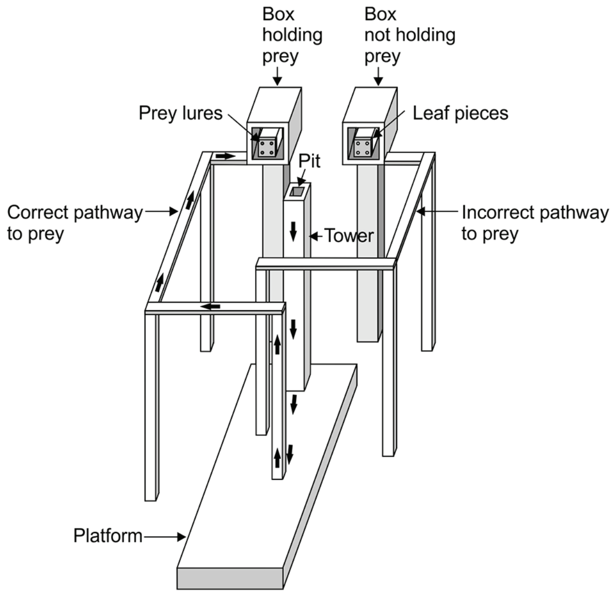


FIGURE 1 | Example of apparatus used in detour-choice experiments. Trial began with a test spider walking out of the pit and on to top of the tower where it could view two boxes. One box contained four lures made from *Oecobius amboseli* and other box contained lures made from four green-leaf pieces. Which box contained prey was determined at random. After the test spider left the pit and walked down from the top of the tower, all of the lures were removed from the apparatus. To complete a successful trial, the test spider chose a walkway after it left the tower and then walked across the platform. The thick arrows indicate the path the test spider took from the tower to the beginning of the correct walkway and then to the end of that walkway. The apparatus sat in a shallow pan filled with water (not shown). Drawing modified from Cross and Jackson (2016).

a single trial, are a striking illustration of a capacity related to intelligence that is highly flexible and also innate.

Octopus, turtles, dogs, and cats – animals much larger than, and only distantly related to, *Portia* – have also been the test subjects in detouring experiments (Kabadayi et al., 2018), but the objectives and methods were substantially different. Typically, these were experiments in which a test subject viewed a target of interest (e.g., food) behind a see-through barrier (e.g., a glass sheet or a wire fence) and the target could not be directly accessed (e.g., Smith and Litchfield, 2010). The test subject's typical response was to make repeated unsuccessful attempts to access this target directly and, when the test subject finally succeeded, it is only by moving around the barrier. It may be easy to envisage this as the test subject having a "eureka moment" in which it suddenly accepted that its efforts to go directly to the target were futile and that, in this eureka moment, it recognized a detour was a workable alternative (Jones, 2003; Chronicle et al., 2004).

This is almost the exact opposite to the way *Portia* behaved (Cross and Jackson, 2016), and these experiments were designed very differently. For instance, repeated efforts to leap directly toward the lures was absent from these experiments. There was nothing suggestive of a eureka moment and, instead, a more accurate characterization may be that *Portia* first assessed the situation and then acted on a plan from the beginning, with this being a spontaneous plan requiring no prior training with the experimental apparatus or protocol. The detouring experiments reviewed by Kabadayi et al. (2018) appear to be especially good for finding evidence of impulse control, but there was little to suggest *Portia* having an impulse-control problem to solve. Impulse control seems to be more aligned with operating as a Skinnerian Creature, but the *Portia* detouring experiments were designed instead as a way of looking for evidence of a kind of intelligence that Popperian Creatures express. These are the Creatures that spontaneously find solutions to problems by internal processing instead of having to first try out potential solutions by actually acting in the physical environment.

EXPECTANCY VIOLATION

Macphail (1985) argued that comparisons should be made between animals that occupy contrasting ecological niches because of how different animals adapt to the specific demands of the particular environments in which they live. We can consider this argument in the design of expectancy violation experiments, in which pre-verbal infants (e.g., Wynn, 1992), non-human primates (e.g., Hauser et al., 1996) and even parrots (e.g., Pepperberg and Kozak, 1986) have been the typical subjects. However, Pepperberg (2002) argued that, as long as the methodological details are tailored to the biology of the particular species being investigated, expectancy-violation methods should be applicable to a wide taxonomic range of animals. To date, very little has been done to investigate expectancy violation by an arthropod but, consistent with Pepperberg's argument, *Portia*-specific expectancy violation methods were used successfully in research on *P. africana*.

In expectancy-violation experiments, it is customary to let a test subject preview a scene that disappears and then, at a later time, comes into view again (Shettleworth, 2010). For example, a screen might be put between the scene and the test subject and then, during the time when the test subject's view is blocked, a scientist can alter the items in the scene. Data relevant to expectancy violation come from comparing how test subjects respond to altered scenes with how they respond to scenes that stay the same. Instances of subjects gazing at an altered scene for longer than they gaze at an unaltered scene (i.e., instances of longer "looking time": see Winters et al., 2015) have been typically regarded as evidence that the subject has detected a mismatch between the current scene and a representation of a scene it had previously loaded into working memory (i.e., this has been a basis for concluding that the individual has experienced expectancy violation).

The expectancy-violation experiments using *P. africana* have been designed to take advantage of how these spiders have exceptional eyesight for animals of their size (Harland et al., 2012), as well as how they respond to lures in similar ways to how they respond to living prey (Jackson and Cross, 2011). These experiments also took advantage of how, in their natural habitat, these spiders pay attention to the different features of their prey (Harland et al., 2012), they routinely take detours to reach prey (Tarsitano and Jackson, 1997; Cross and Jackson, 2016), and they encounter various numbers of other conspecific individuals (Nelson and Jackson, 2012). However, the data of interest when using *P. africana* differ from the standard "looking time" used in experiments on bigger animals.

In the first study (Cross and Jackson, 2014), experiments were designed to determine whether *P. africana* holds representations of specific prey types in working memory during predatory sequences (**Figure 2**). After seeing a particular prey item at the beginning of a trial (**Figure 3**), *Portia* positioned itself for initiating an attack, but then, before *Portia* could act, the prey item was hidden behind a shutter for 90 s. During this time, *Portia* waited while facing the shutter and then, once the shutter was lifted after this delay, *Portia* could see prey that was either identical to or different from the type of prey it had seen earlier.

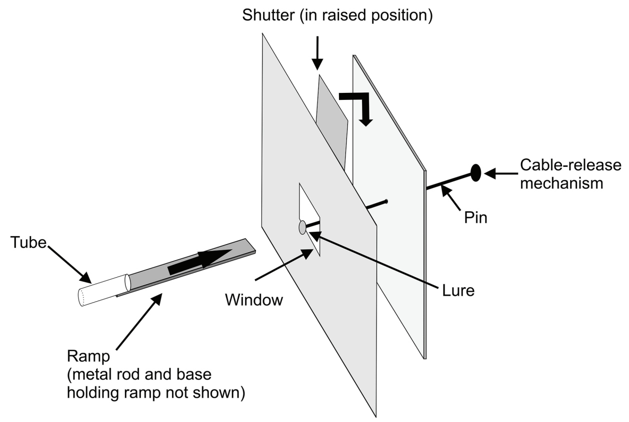


FIGURE 2 | Apparatus used in expectancy-violation experiments for changes in prey type. A trial began with a test spider leaving the glass tube and walking across the ramp (thick arrow) toward a lure. Once the spider had faced the lure for 30 s, the lure was pulled back from the window and the shutter was lowered for 90 s, blocking the spider's view of the lure. The lure was removed from the pin during the 90 s. In experimental trials, a different lure was then attached to the pin and, in control trials, the same lure was re-attached to the pin. After the 90 s, the shutter was raised, and it was recorded whether the test spider leapt at the lure. Reprinted by permission from Springer (Cross and Jackson, 2014).

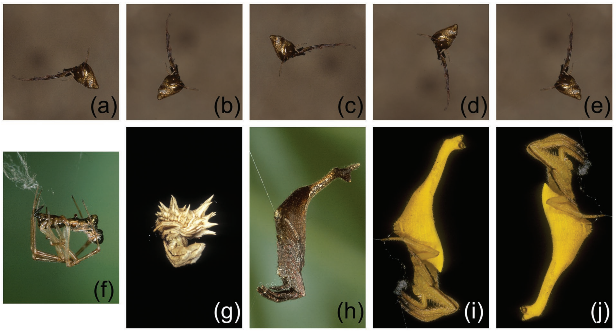


FIGURE 3 | Prey spiders used for making lures in expectancy-violation experiments for changes in prey type. **(A–E)** *Argyrodes* sp. 1 positioned in different orientations; **(F)** *Argyrodes* sp. 2; **(G)** *Pycnacantha tribulis*; **(H)** *Arachnura scorpionoides* (brown morph); **(I,J)** *Arachnura scorpionoides* (yellow morph) positioned in different orientations. Reprinted by permission from Springer (Cross and Jackson, 2014).

In these experiments (Cross and Jackson, 2014), the data of interest were the number of test spiders (*P. africana*) that attempted to attack this lure, instead of looking time (i.e., instead of the length of time the test spider spent gazing at the lure). There was no evidence that *Portia* was more or less likely to attack if only a lure's orientation had changed during a trial (**Figures 3A–E,I,J**, **4A**). However, when the prey species (**Figures 3A,F–I**) or prey color (**Figures 3H,I**) had changed during the trial, significantly fewer *Portia* individuals followed through with an attack (**Figure 4B**). All of these experiments were counterbalanced, and there was no evidence to suggest *Portia*'s responses were influenced by the order in which prey were presented. This suggests that *Portia* experienced expectancy violation when the representation of the prey type it had loaded into working memory at the beginning of a trial did not match with the prey it saw later.

In the second study (Cross and Jackson, 2017), experiments were designed to determine whether *P. africana* represents the specific number of prey in a scene (**Figure 5**), with the methods required for this being substantially different from the methods in the earlier study (Cross and Jackson, 2014). In these experiments, *Portia* had to complete a detouring task, and the data of interest pertained to whether *Portia* became less inclined to complete the detour when presented with an unexpected number of prey at the end of the detour. These experiments began with *Portia* leaving a pit and standing on top of a starting tower from which it could view a scene consisting of a particular number of prey items. In order to reach this scene, *Portia* walked down from the starting tower before it walked across a walkway, up a viewing tower and then across an access ramp. However, when walking up the viewing tower near the end of the detour, the scene went out of *Portia*'s view because the walls of this tower were opaque. The scene was either changed or it remained the same during the time that *Portia* walked up the viewing tower. It was only after reaching the top of the viewing tower when *Portia* could view the scene again, but now the number of prey items might have changed.

Compared with control trials in which the number of prey seen was the same as before, *Portia* became disinclined to complete the detour when the following changes in number were made: 1 vs. 2, 1 vs. 3, 1 vs. 4, 2 vs. 3, 2 vs. 4, or 2 vs. 6 (**Figure 6**). These effects were independent of whether the larger number of prey was seen at the start or at the end of the trial. Moreover, when the number remained the same during a trial, there was no evidence that changing the size or arrangement of the prey influenced *Portia*'s inclination to complete the detour (see Cross and Jackson, 2017). There were also no significant effects when the number of prey changed between 3 vs. 4 and 3 vs. 6 (**Figure 6**), which suggests that *Portia* may characterize three or more prey as a single category ("many prey"). However, *Portia* seems to represent 1 and 2 as discrete number categories.

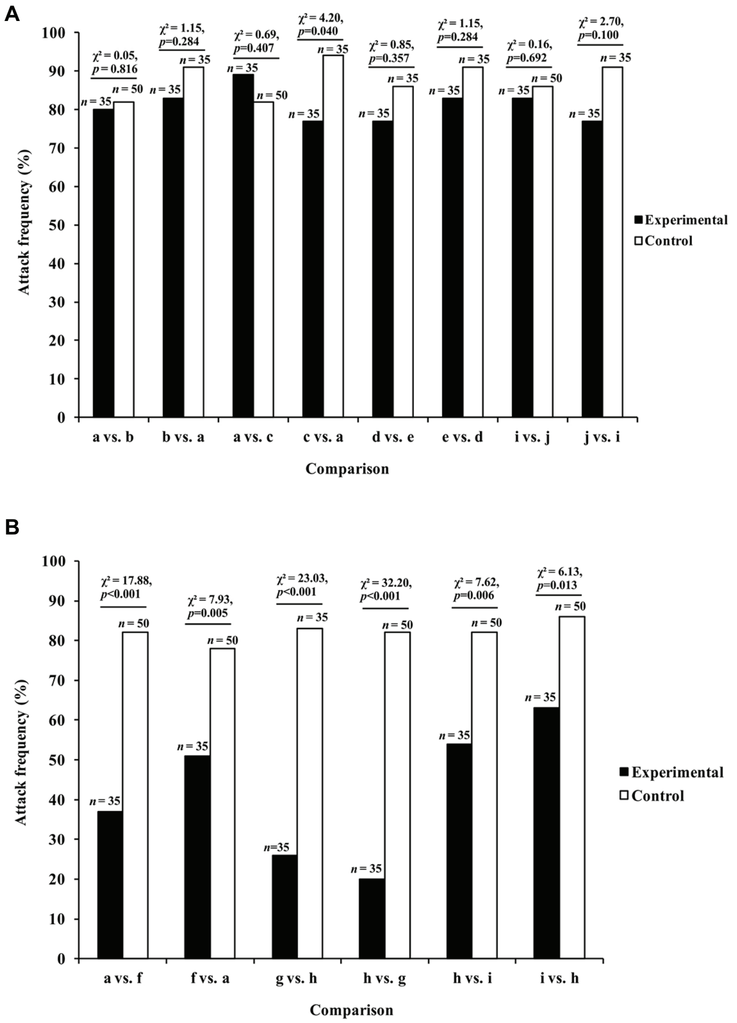


FIGURE 4 | Results from expectancy violation experiments in which *Portia africana* was tested with **(A)** different prey orientations (prey type remained constant) and **(B)** different prey types (prey orientation remained constant). See **Figure 3** for the different prey orientations and types shown during trials. Experimental trials: first prey orientation or prey type replaced by second prey orientation or prey type. Control trials: first prey orientation or prey type did not change during trial. Data analyzed using χ^2 tests of independence. Attack frequency: percentage of test spiders that leapt at the prey. Total number of test spiders (*n*) shown above bars. Reprinted by permission from Springer (Cross and Jackson, 2014).

The range over which *Portia* represents prey number appears to be consistent with the range over which vertebrates have been shown to practice subitizing, this being the rapid and accurate estimation of small numbers of individuated objects (Davis and Pérusse, 1988). Yet the way *Portia* responded in experiments is inconsistent with how subitizing is usually characterized. For instance, primates (Hauser et al., 1996) normally respond no more than a few seconds after viewing a stimulus, but *Portia* normally responded after a minute or longer. *Portia* typically engages in a slow, methodical visual-inspection routine before responding (Harland et al., 2012), which is also inconsistent with how subitizing is normally characterized as being automatic and pre-attentive. We propose that, instead of subitizing, *Portia* slowly individuated objects and then held separate representations of these objects in working memory. More specifically, we propose that, while on top of the first tower, *Portia* loaded a representation of the number of prey individuals in the scene into working memory and then, while on top of the second tower up to 21 min later, *Portia* compared the number of prey in the scene now in view with the number of prey represented while on the first tower.

SOLVING A NOVEL CONFINEMENT PROBLEM

When discussing Macphail's null hypothesis of no differences in intelligence, we need to specify the kind of difference being considered. Much of the time, it seems implicit that

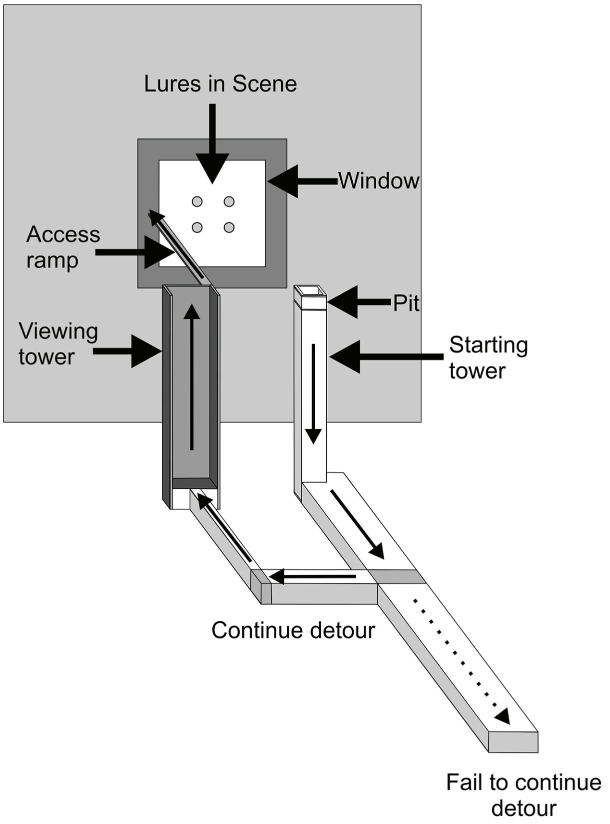


FIGURE 5 | Apparatus used in expectancy-violation experiments for changes in prey number (apparatus sat in a shallow pan filled with water; not shown). A trial began with a test spider walking out of the pit and on top of the starting tower. The thin arrows indicate the path that test spiders then took to reach the lures without getting wet; the dotted arrow indicates the path that test spiders took to opt out of completing the detour. When the test spider arrived at the bottom of the viewing tower, which was opaque, the scene was removed and then replaced by a different scene or else the previous scene was returned. It was then recorded whether the test spider crossed the access ramp after reaching the top of the viewing tower. Drawing modified from Cross and Jackson (2017).

the issue of interest is the level to which intelligence is expressed in a domain-general manner. However, the extent to which cognitive processes used by animals are domain-specific instead of domain-general remains poorly understood (Chiappe and MacDonald, 2005; Anselme, 2012), and arguments that improving capacity for domain-general intelligence requires a higher investment in mass of expensive brain tissue (Burkart et al., 2017) suggests that especially small animals, including spiders, will be skewed more toward domain-specific intelligence than is the case for larger animals such as birds and mammals (Logan et al., 2018). Research on *Portia* may be especially interesting in this context.

Part of *Portia*'s strategy for targeting web-inhabiting spiders as prey involves using signals to gain dynamic fine control over the resident spider's behavior ("aggressive mimicry"; Jackson and Cross, 2013). This is achieved by *Portia* using any one or a combination of its 10 appendages (eight legs and two palps) to generate web signals (i.e., vibration and tension patterns on the silk lines in the web). Sometimes *Portia*'s signals may be readily understood as mimicking the movements of a small insect trapped on the web; in these instances, *Portia* lures its victim over for the kill. The variety of signals at *Portia*'s disposal seems unlimited; the way any one appendage moves can vary and *Portia* can move each appendage independent of how other appendages are moving (Jackson and Blest, 1982). By repeating signals that elicit an appropriate response from its intended prey and by trying new signals when an appropriate response is not forthcoming (Jackson and Wilcox, 1993; Jackson and Nelson, 2011), *Portia* achieves a high level of proficiency at adjusting its predatory strategy to the particular prey spiders it encounters. Using this trial-and-error strategy (a "generate-and-test algorithm"; Simon, 1969), *Portia* preys on a vast array of different kinds of spiders (Jackson and Pollard, 1996), including spiders that can prey on *Portia*. It has been proposed that *Portia*'s capacity for flexibly deriving signals by trial-and-error is an important adaptation for successfully targeting prey that are also predators (Jackson, 1992).

Whether *Portia*'s proficiency at using trial-and-error is restricted to this predatory strategy (domain specific) or whether it is applicable to novel problems (see Beecher, 1988) has been considered in experiments where individuals were faced with a problem of how to escape from an island in a water-filled tray (i.e., a confinement problem; **Figure 7**). The island, in the middle of the tray, was surrounded by an atoll; water filled the space between the island and the atoll, and also filled the space between the atoll and the edge of the tray. The basis for calling this problem "novel" includes how there is no evidence that *Portia* routinely crosses water in nature. Moreover, adding to the novelty of the problem, test spiders were helped forward across the water to the atoll or forced back to the island during the experimental trials. There were only two ways *Portia* could leave the island, either by stepping into the water and then swimming the whole way across to the atoll or by first leaping into the water and then by swimming. When leaving the island by swimming, test spiders slowly placed their forelegs on the water, pushed off from the island with their rear legs, moved completely out into the water in a spread-eagle posture and then propelled their bodies across the water surface by moving their legs in a stepwise fashion (see Suter, 2013). When leaving the island by leaping, spiders landed on the water at a point about halfway across, and then swam the rest of the way to the atoll.

Portia individuals were assigned at random to two groups, with these two groups differing with respect to the method of leaving the island that would be successful. When *Portia* used the escape method pre-determined for its group to succeed, a small plastic scoop was used to make waves behind *Portia* to help it across to the surrounding atoll. When *Portia* used the other escape method, the scoop was used to make waves to move *Portia* back to where it had started from. Once on the atoll or back on the island, test spiders could then try again. In the first experiments based on this confinement problem, the test spiders were *P. fimbriata* (Jackson et al., 2001). The test spiders that had succeeded at arriving on the atoll usually repeated the same escape method to then reach the

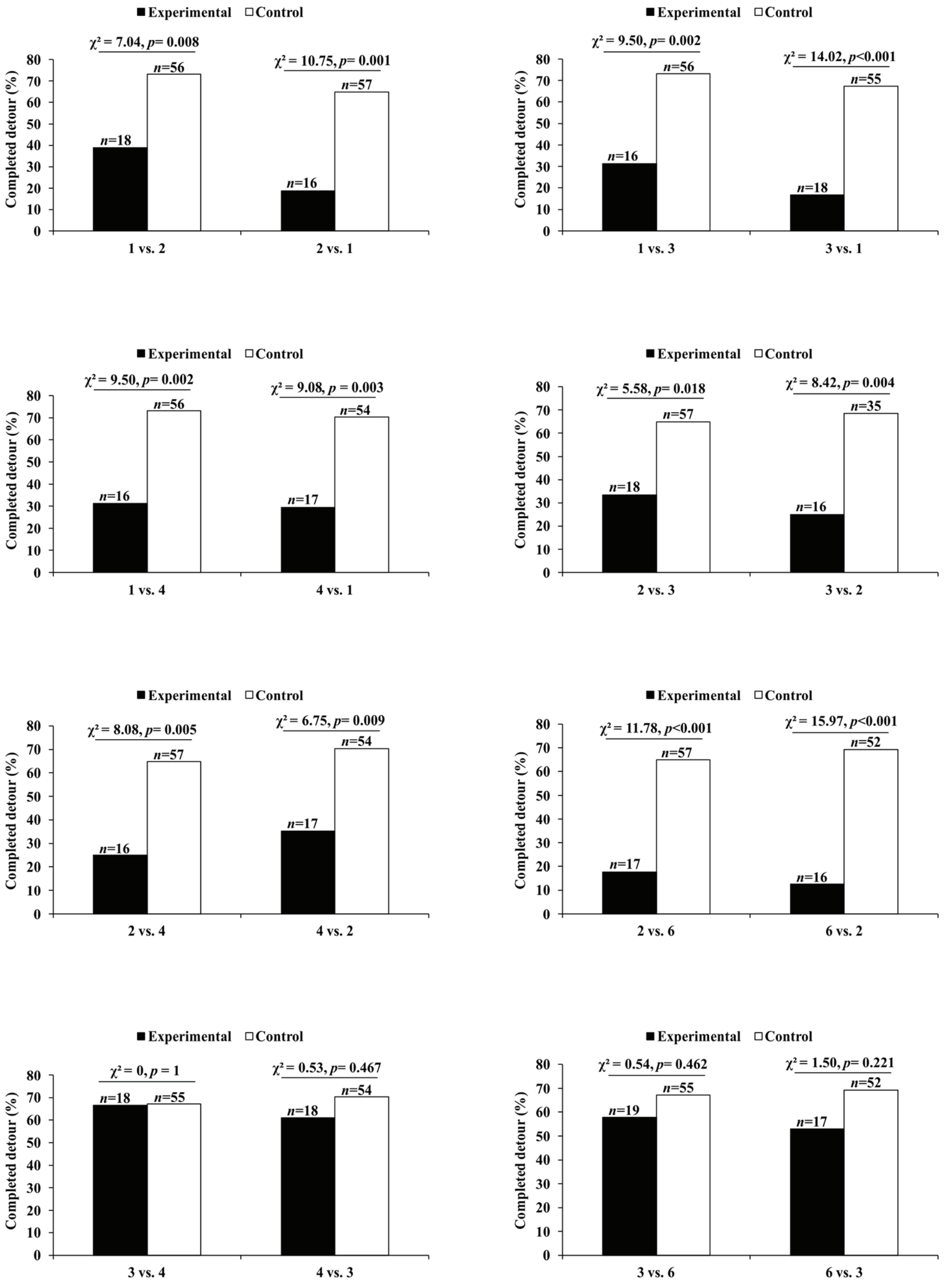


FIGURE 6 | Results from expectancy violation experiments in which *P. africana* was tested with different numbers of prey. Experimental trials: first prey number replaced by second prey number. Control trials: first prey number did not change during trial. Data analyzed using χ^2 tests of independence. Completed detour: percentage of test spiders that crossed the Access Ramp to reach the location of the prey. Total number of test spiders (n) shown above bars. Figure modified from Cross and Jackson (2017).

edge of the tray and those that failed usually switched to using the other escape method.

In a more recent study using the confinement problem (Cross and Jackson, 2015), two other *Portia* species (*P. africana* and *P. schultzi*) were used as test spiders, and there were also seven other test-spider species from different genera, but with all of these genera being from the same salticid subfamily (*Spartaeinae*) as *Portia*. Five of the non-*Portia* species (*Brettus adonis*, *Brettus albolimbatus*, *Cyrba algerina*, *Cyrba ocellata*, and *Cyrba simoni*), along with the two *Portia* species, are known to practice aggressive mimicry and derive signals by trial and error, whereas the other two non-*Portia* species (*Cocalus gibbosus*

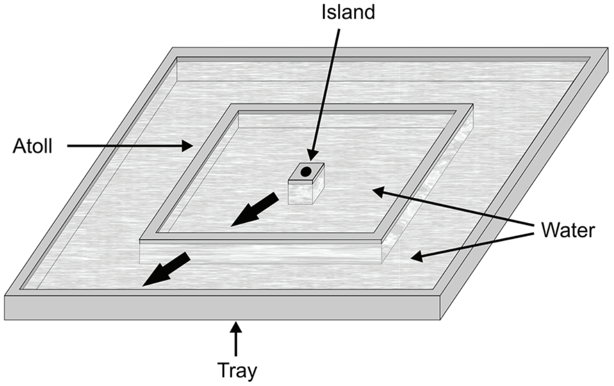


FIGURE 7 | Apparatus used for ascertaining whether spiders can solve a novel confinement problem by using trial-and-error. In a water-filled tray, there was an island surrounded by an atoll. The test spider emerged from a hole in the island and then either chose to leap or swim away from the island. Before testing began, it was determined at random which of these two potential choices would succeed. When the test spider made the successful choice, it was helped across to the atoll, and a record was made of whether the test spider repeated its choice to reach the edge of the tray (thick arrows). When the test spider made the unsuccessful choice, it was forced back to the island, and a record was made of whether the test spider switched its choice when attempting to reach the atoll again. Reprinted by permission from Springer (Cross and Jackson, 2015).

and *Paracyrba wanlessi*) are not known to practice aggressive mimicry. The findings from experiments showed that the seven aggressive-mimic species were proficient at solving the novel confinement problem by repeating "correct" choices (i.e., the choices that delivered them to the atoll) and by switching when they made "incorrect" choices (i.e., the choices that sent them back to the island), but there was no evidence of the two non-aggressive-mimic species solving the same novel problem (**Figure 8**). These findings suggest that species which use trial and error to solve aggressive mimicry problems are predisposed to be proficient at using trial and error in a novel context.

Local adaptation is another relevant factor because different populations of a single *Portia* species are known to adopt different predatory strategies (Jackson and Pollard, 1996). For example, two populations of *P. occidentalis* from the Philippines have been investigated (Los Baños and Sagada). Los Baños is a low-elevation rainforest habitat where the range of prey-spider species is much wider than in Sagada, a high-elevation pine-forest habitat, and it was found that individuals from Los Baños were significantly more inclined than individuals from Sagada to derive web signals by trial-and-error (Jackson and Carter, 2001). In the context of domain-generality, there was another interesting difference. Individuals from Los Baños were significantly more inclined than individuals from Sagada to solve the novel confinement problem by trial-and-error (Jackson et al., 2006).

These findings from different species, and from different populations of a single species, appear to be salticid examples of a transition from domain-specific to domain-general problem-solving capacity, this being a transition also suggested as happening sometimes with other animals (Johnston, 1982; Papaj, 1986; Dukas, 1998), but we should be wary of envisaging domain-specific and domain-general as two distinct categories. A more useful alternative is to envisage "domain specific" and "domain general" as being terms pertaining to different ends of a continuum (e.g., see Jackson and Cross, 2011). The findings from the confinement problem experiments suggest that the domain-general region of this continuum is particularly relevant for understanding the behavior of aggressive-mimic spartaeines. In other words, being proficient at solving a novel confinement problem by trial-and-error may be a spin-off from these spiders having evolved proficiency at deploying highly plastic aggressive-mimicry strategies in the context of predation.

ARE SPIDERS INTELLIGENT?

When discussing his null hypothesis, Macphail (1985, 1987; Macphail and Bolhuis, 2001) focused almost entirely on vertebrates and he had little to say about arthropods (insects, spiders, crustaceans and their kin). This drastically limited the scope of his hypothesis since only a small minority of animal species are vertebrates; most animal species are arthropods. Yet, while this omission is jarring to some of us (e.g., see Kupfermann, 1987), it may be hardly noticed or else assumed to be obviously justified by others.

To decide how serious this omission might be, it may help first to ask how the null hypothesis might actually be useful. How we answer this question is similar to how Zentall (2018) answered a comparable question about Morgan's canon. Rather than being a call to reduce intelligence to its lowest common denominator, the null hypothesis will be more useful when seen as a way to challenge investigators to develop procedures for identifying differences in intelligence. When we refer to "intelligence," the core topic of interest is flexible problem-solving capacity, and especially the distinction on a continuum between domain generality and domain specificity. The null hypothesis can then be useful in challenging us to develop procedures for objectively specifying levels of flexibility and domain-generality.

This may work as a rationale for the null hypothesis, but not with arthropods relegated to a footnote. Even if not explicitly stated, a decision not to include arthropods in the conversation often seems to be based on accepting as a foregone conclusion that, at best, arthropod intelligence is distinctly inferior to vertebrate intelligence or, at worst, arthropods are not intelligent at all (i.e., arthropods are automatons).

For examples of arthropod intelligence, we have focused especially on *Portia*, but we should point out that cognition, instead of intelligence, was the context in which the research we reviewed was originally discussed. This is important because we envisage cognition, which pertains to representation, and intelligence, which pertains to flexibility, as being overlapping, but not identical, topics. *Portia* gets our attention because we are especially interested in instances in which well-developed cognitive capacities are deployed in flexible problem solving.

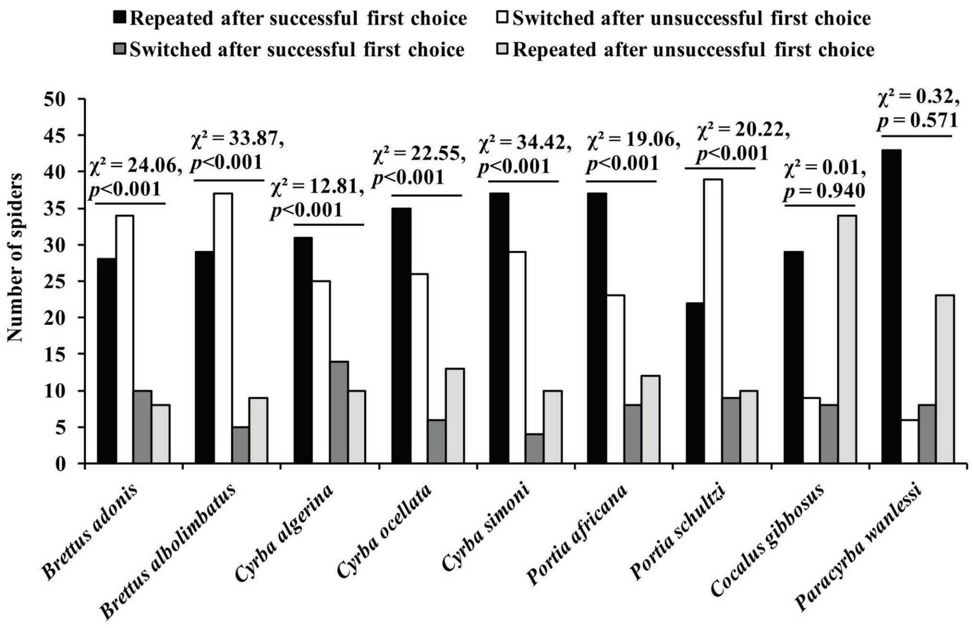


FIGURE 8 | Results from confinement experiments. Spider began trial on an island surrounded by an atoll in a pan of water and given two opportunities to choose its method of crossing the water (i.e., by leaping or by swimming). Successful first choice: plastic scoop made waves to help spider across to the atoll. Unsuccessful first choice: plastic scoop made waves to move spider back to the island. After making its first choice, it was recorded whether the spider repeated that choice or switched. Data analyzed using χ^2 tests of independence. Reprinted by permission from Springer (Cross and Jackson, 2015).

Flexible problem solving is not necessarily cognitive in any substantial way. To suggest otherwise would be to forget radical behaviorism's explicitly non-cognitive interpretation of operant conditioning. This behaviorist interpretation may have faded with time but, at the very least, it shows that non-cognitive intelligence, or flexible problem solving, by animals is conceivable. In Dennett's scheme, these animals are Skinnerian Creatures. When we consider Darwinian Creatures, which might be aptly called automatons, non-cognitive intelligence is also relevant, but in a different way. These animals use solutions derived by natural selection, a non-cognitive flexible problem-solving algorithm with formal similarities to operant conditioning (Skinner, 1981; Watson and Szathmáry, 2016). It is with Popperian Creatures that the expression of flexible problem solving becomes distinctively cognitive in character.

When defining intelligence, Burkart et al. (2017) emphasized individuals showing proficiency at acquiring new knowledge from interacting with the physical or social environment. This might seem more characteristic of a Skinnerian Creature rather than a Popperian Creature, but the extent to which Burkart et al. (2017) allude to knowledge, understanding and representation would probably go well beyond anything a radical Skinnerian would accept. Burkart et al. (2017, p. 2) also refer to using "this knowledge to organize effective behavior in both familiar and novel contexts" and, by saying this, they imply that learning-based intelligence is cognitive in character, this being aligned with post-Skinnerian representation-based theory of learning (see Gallistel and King, 2009) and, as such, more related to Popperian than Skinnerian Creatures.

Intelligence and learning are often discussed together, but making learning a necessary part of the definition of "intelligence" would artificially remove the research we reviewed on *Portia* from the conversation. For this research, each individual test spider was a subject in a single trial and all of the test spiders had been reared under standardized conditions in the laboratory with no prior experience of the procedures and apparatus. The rationale for these procedures was to ensure that test spiders were not solving the experimental problems as Skinnerian Creatures. In the detour-choice experiments, for example, the test spider solved the problem by choosing a particular path without having had any prior experience of the consequences of taking that path, which is not compatible with being a Skinnerian Creature. The findings from the detour-choice experiments are also incompatible with test spiders being Darwinian Creatures because the particular path that spiders took to solve the detouring problem was set at random before each trial began.

It might be disconcerting that we say "innate" because this word is often associated with the idea of animal being an inflexible automaton, as though being innate is the antithesis of being intelligent or cognitive. To understand why this is not the case, there is an important distinction to make between having solutions to problems and having the capacity to find these solutions. When presented with detour-choice problems, for instance, *Portia* uses an innate and flexible problem-solving capacity. In other words, the capacity to solve detour-choice problems is innate, but the specific solutions to these detour-choice problems are not innate and they are also not memorized solutions derived from prior personal experience.

The setting in which *Portia* encounters prey is typically accompanied by extreme unpredictability and mortal risk, and this may be a major component of the adaptive context in which these flexible capacities evolved. When invading other spiders' webs, *Portia* enters the prey-capture arena of another predator and, when making web signals, intimately interfaces with that predator's sensory system. From *Portia*'s perspective, encounters that end with *Portia* killing the resident spider are successful and encounters that end with the resident spider killing *Portia* are unsuccessful. Success often depends on *Portia* gaining dynamic control of the resident spider's behavior by deploying especially intricate and flexible behavior that is cognitive and intelligent in character (Harland and Jackson, 2004; Jackson and Cross, 2011, 2013). Burkart et al. (2017) emphasized the role of social unpredictability in the evolution of general intelligence, which is interesting because social unpredictability occurs when groups of conspecific individuals are actively engaged in complex interactions. This seems similar to the unpredictability *Portia* contends with while engaged in intricate and intimate interactions with other predators.

Burkart et al. (2017) also envisaged general intelligence as being closely related to three core executive functions: working memory (see Baddeley, 2012), cognitive flexibility, and inhibitory control, with inhibitory control including selective attention as well as behavioral inhibition and cognitive inhibition. Interest in all of these executive functions has been integral to research on *Portia* (Jackson et al., 2002), as well as other salticids (Jackson and Cross, 2011). Research on *Evarcha culicivora* (Cross and Jackson, 2009, 2010a,b) has been especially relevant. This salticid specializes at preying on mosquitoes and, for this salticid, specialization includes intricate, innate systems of deploying selective attention to specific types of prey. This includes specific odors priming selective olfactory attention and specific optical cues priming selective visual attention, and also cross-modality priming in both directions (selective visual attention being primed by specific odors and selective olfactory attention being primed by specific visual stimuli).

To more fully address whether salticid species differ with respect to domain generality, it is important to determine whether there are correlations in how individuals perform in certain tasks. It would be interesting to investigate, for instance, whether superior proficiency at deploying selective attention is correlated with superior performance on other tasks, including solving novel problems and planning detours. Computational complexity may also be especially relevant when characterizing animal intelligence and potentially measuring how intelligence varies among species and populations within species. This may be especially tractable when focusing on converging topics, such as expectancy violation with respect to a change in number.

REPRESENTATION OF NUMBER

Numbers are related to mathematics and, being abstract, mathematics is often experienced by people as a hard subject. Perhaps this is why we tend to admire people who are good at mathematics and why even a hint of mathematical aptitude by a spider can seem sensational. However, relying on impressions alone will not take us very far toward a goal of understanding number-related capacities in the context of animal intelligence. Conceptual clarity is especially important whenever intelligence and cognition are discussed, but it is also especially difficult when discussing numerical cognition in particular. Returning to *Portia*'s performance in the expectancy-violation experiments, we can begin by focusing on two of the most basic questions that arise when we propose that this is an example of cognition-based intelligence. If this is cognitive, then what is represented? If the answer to that question is something related to numbers, then what kind of number do we mean?

In these experiments (Cross and Jackson, 2017), test spiders could see prey at the beginning and at the end of a detour path, but not while taking the detour. The number of prey in view at the end was either the same as or different from the number in view at the beginning, and the primary finding was that test spiders hesitated at the end of a detour when the number was different. This basic finding held even when there were control trials for considering the possibility of non-numerical variables related to prey size and prey arrangement being alternative explanations.

Based on the experimental findings, we can offer an answer to the question of what is represented by saying it is the number of prey individuals in a scene. We can also specify the kind of number we mean, but we have to do this carefully because, all too often, we have been forced to guess how number-related expressions are used in the vast scientific literature. The first step is to be mindful of the word "of" when we propose that *Portia* represents the "number of prey." The prey are individual physical things, but numbers are abstract. The type of number we mean when saying "the number of prey" is important, but whether *Portia* literally represents numbers as strictly abstract constructs is not the specific question the expectancy-violation experiments addressed.

We can begin with the casual expressions "counting numbers" and "measuring numbers" before moving on to kinds of numbers as formally-defined in mathematics. For doing this, we can envisage the expectancy-violation experiments as presenting test spiders with a problem of determining how many prey were seen at the end of a detour and whether this is the same as or different from how many prey were seen before. We say "how many" because the kind of answer we expect is a counting number (1, 2, and so forth), implying discrete countable things. For *Portia*, the countable things were individuated objects or, more specifically, prey individuals, but saying "countable" is not the same as saying "counted." "How much" implies a different kind of problem, with the answer being expressed using measuring numbers (i.e., the kind of number that applies to a continuum). "How much" pertains to stuff that is measured, not counted. Our hypothesis is that, instead of representing "how much" prey stuff is in a scene, *Portia* individuates prey items and represents "how many" prey individuals are in a scene.

As Gregorian Creatures, we may rarely think about how remarkable it is that we express both "how many" and "how much" using numbers. In mathematics, the abstract analogues of the counting numbers are the positive integers (1, 2, …). All numbers are abstract, but the rationale for the abstracting that leads to the positive integers comes from focusing on the concrete action of individuating objects. Owing to this focus on correspondence, "natural numbers" is an appropriate name for the positive integers (Rucker, 1987).

These are the numbers used for doing basic arithmetic (addition, subtraction, multiplication, and division), and it is with further abstraction that we derive the rational numbers and then the real numbers, which we use for expressing magnitudes on a continuum. The progression leading by abstracting to the real numbers began with the natural numbers during the history of mathematics and this also appears to be the typical progression in the development of numerical comprehension and proficiency by children (Carey and Barner, 2019). As Gregorian Creatures, mathematicians and children do this abstract work using an arsenal of mind tools inclusive of verbal language, numerals, decimal places, equations, formal logic, and so forth (Rucker, 1987).

It is here that we need to be especially careful when discussing non-human animals, including *Portia*. Macphail's null hypothesis challenges us to look for qualitatively distinct kinds of intelligence, and the abstract, flexible problem-solving capacity supported by mathematics as mind tools seems to be a prime candidate. Animals expressing this kind of intelligence are Gregorian Creatures, but the findings from using *Portia* in expectancy-violation experiments are not evidence of *Portia* engaging in abstract numerical reasoning as a Gregorian Creature.

However, *Portia*'s performance in these experiments is comparable to the performance of pre-verbal human infants in similar experiments (Carey, 2004). For *Portia*, as for a 1-year-old infant, this capacity does not seem to be applied beyond three individuated objects. Yet, as Carey (2004) points out, this is a non-trivial cognitive capacity and it seems to be an innate cognitive precursor to the abstract derivation of integers and then the other numbers. The way the expression "exact" corresponds to integers as abstract constructs is similar to the way "exact" corresponds to individuated objects. Owing to the experimental methods, the findings for *Portia* corresponded with "exact" in this context related to individuating. *Portia* displayed evidence of expectancy violation when the scene in view at the end of a detour, compared with the scene at the beginning of a trial, had one more prey individual and also when it had one fewer prey individuals. However, these experiments using *Portia* seem to differ considerably from much of the literature pertaining to non-human animals displaying number-related capacities.

Conclusions from more familiar research on animals displaying number-related cognitive capacities tend to be based on correspondence to the Weber-Fechner law and referred to as instances of animals using an "approximate number system" (e.g., Nieder, 2019). Although the Weber-Fechner law, and expressions such as "quantity" and "amount," can be relevant when comparing scenes populated by discontinuous objects, the Weber-Fechner law is not about individuating objects as directly as is the case when *Portia* was investigated using expectancy-violation methods.

The Weber-Fechner law pertains to finding that the discriminability of two magnitude values is a function of their ratio (Nieder, 2019). The magnitudes relevant to this law include, for example, brightness, loudness, duration, length, and area, all of which are normally envisaged as continuous variables. Real numbers, as highly abstract constructs, can be applied to continuous variables, but it is apparent that this is not the kind of number intended when a system used by an animal is called the "approximate number system." In better designed experiments, the animal compares scenes populated by objects, and considerable effort is made to rule out the influence of continuous variables on experimental findings. This leads to conclusions pertaining to number, but now from a perspective different from expectancy-violation experiments using *Portia* and preverbal infants.

The perspective we have when considering the approximate number system is relevant to intelligence, cognition, and numbers, but with the sense in which it pertains to numbers seeming less direct and less specific. Reference to ratios might suggest that, when applied to scenes populated by objects, the cognitive capacity revealed by correspondence to Weber-Fechner law is a precursor to understanding fractions and the rational numbers expressed to decimal places. Saying "approximate" would seem appropriate because, although all rational numbers are discontinuous, there is no conceptual end to how small they can be, which in turn means rational numbers correspond at least roughly to answering "how much" questions with measuring numbers. Another logical possibility is that, when using the approximate number system, the animal renders a representation corresponding to a specific natural number, but with an accompanying representation of a level of uncertainty. However, trying to answer questions about the intended kind of number might be misguided because the major distinction seems to be between individuating as primary versus ratios as primary, with this distinction being more fundamentally important than is widely acknowledged (Gebuis et al., 2016). It may be only Gregorian Creatures that can achieve the level of abstract reasoning needed to bring about a convergence of the different concepts of number implicit in the distinction between individuating as primary and ratios as primary.

SPATIAL NAVIGATION

Spatial navigation may be a more rewarding context in which to investigate the intelligence-related capacities animals display specifically with respect to "measuring numbers." This could be especially interesting with respect to abstract intelligence because, when based on path integration, spatial navigation implies computationally complex behavior, by which we mean behavior that appears to require the equivalent of mathematical calculation by the animal (Gallistel, 1990b; Grace et al., 2020).

Path integration by arthropods has been investigated especially often in the context of homing behavior, with some of the most striking examples coming from research on desert ants. When foraging in featureless environments, these ants may meander and wind about while moving away from their nests, but they are very proficient at returning directly to the nest without retracing the path that they took on the outward journey (Bühlmann et al., 2011). As path integration pertains to vector algebra, concluding that the desert ant relied on path integration suggests that the ant represented the outward journey from the nest as a series of vectors and then estimated its current location with respect to the nest by summing these vectors (Collett and Collett, 2000). The direct path back is then the inverse of the vector sum.

Finding examples of animal behavior that can be described mathematically is not, by itself, a basis for concluding that the mathematical description corresponds to the internal processing carried out by the animal. However, it is hard to escape this implication with path integration because there is no known way of implementing path integration without also implementing the vector-based computations (Gallistel, 2017).

Among spiders, there is experimental evidence of homing behavior based on path integration from research on an assortment of non-salticid species (e.g., Ortega-Escobar and Ruiz, 2017). Homing by salticids has been demonstrated experimentally (Hoefler and Jakob, 2006), but this has been in the context of relying on landmarks instead of path integration. There is evidence of salticids relying on path integration (Hill, 1979) but, instead of being in the context of homing, this has been in the context of taking detours while pursuing prey.

In the detouring experiments we discussed earlier, the objective was not to look for evidence of path integration, but rather to look for evidence of *Portia* making a plan to access prey that is no longer visible while the plan is being implemented. *Portia*'s behavior in these experiments can be characterized as "navigating," but with *Portia*'s primary navigational decision being to reach the beginning of the correct path. This might entail *Portia* moving directly away from the location of the prey, and it might mean walking past the beginning of the incorrect path before reaching the beginning of the correct path. However, there was no need for *Portia*'s plan to be inclusive of every twist and turn along the correct path. Once on the path, *Portia* only had to reach the end of that path, with the prey remaining out of view.

In the field, *Portia* often negotiates more complex detouring paths that include multiple branches (Jackson and Wilcox, 1993) and require repeated directional decisions. Although observations from the field might suggest ways in which *Portia* could be used in research more directly related to navigation, including navigation by path integration, these more complex settings for detouring have not been simulated using *Portia*. For this, we can turn to research on salticids that inhabit vegetation and normally target active insects as prey in complex three-dimensional habitats.

In a series of elegant experiments, Hill (1979) demonstrated how salticids from the genus *Phidippus* navigate along paths with side branches. *Phidippus* normally adopts a reconnaissance position on a plant and, after sighting an insect that is inaccessible by a direct path, *Phidippus* takes multiple short detours to reach successive vantage points in the vegetation before arriving close enough to complete the prey-capture sequence. In experiments using artificial plants, Hill demonstrated that *Phidippus* identifies an accessible part of the artificial plant closer to the prey (the "secondary goal") and then makes a detour to the secondary goal, during which time the prey was moved out of view. Upon arriving at the secondary goal, *Phidippus* then re-orients accurately toward the location of where the prey would have been had it not been moved. The re-orientation data are evidence of *Phidippus* having implemented path integration, based on summing vectors in three dimensions, with respect to the prey's location as seen from the reconnaissance position on the plant.

Although *Phidippus*'s detours are short compared with *Portia*'s, *Phidippus* takes detours in a setting where additional directional decisions are needed. In other words, after completing a short detour, *Phidippus* can quickly identify the location of the prey from its new vantage point and then take another detour to get closer to the prey's location. By taking successive short detours based on successive use of path integration, and then re-orienting to the prey's location, *Phidippus* navigates through the vegetation, a complex physical habitat, to arrive at the primary goal, the prey. This differs from path integration in the context of homing because, in these experiments, path integration was implemented by a test spider with respect to a distant prey individual's location instead of with respect to the test spider's own earlier location. *Phidippus* using path integration in the context of navigating to distant prey seems to depend critically on the exceptional capacity for spatial vision supported by salticid eyes.

G AND BRAIN SIZE

When the notion of larger values for *G* or *g* requiring larger brains approaches the status of an axiom, it becomes unsurprising that vertebrate examples dominate the literature on animal intelligence (Chittka and Niven, 2009; Logan et al., 2018). All the while, there is the inconvenient fact that most animal species are arthropods, and the differences in brain size are enormous when we compare most vertebrates to arthropods. There is also a tendency for vertebrates to have a much slower pace of development and a much longer lifespan than is typical for arthropods. The way vertebrate intelligence is typically discussed may make it seem that, *ceteris paribus*, the expression of intelligence by arthropods can only be negligible. Yet, when we look at examples from spiders and *Portia* in particular, we find capacities that are routinely discussed as examples of intelligence when they are expressed by vertebrates. Moreover, *Portia* is not an isolated aberration. It does not take long to find many comparable examples from research on other arthropods, especially bees and ants (Chittka and Niven, 2009).

When the focus is on vertebrates, the discussion tends to be directed more toward looking for potential advantages gained by having larger brains, but including arthropods in the discussion may shift the discussion more toward looking for potential handicaps or limitations imposed by having minute brains (Eberhard and Wcislo, 2012). Common sense leads us to expect severe limitations more widely than just in the context of intelligence, and some of these other contexts might be more amenable to objective measurement than intelligence currently is for spiders. For example, as an orb-weaving spider's web is a detailed record of the numerous intricate decisions made when building the web, data acquired from examining webs can be used for comparing the precision with which smaller and larger spiders build their webs. Yet, when detailed comparisons were made, no evidence was found of smaller orb weavers building less precise webs. With this being the case despite orb-weaving spiders varying in body mass by 400,000 times (Eberhard, 2011; Eberhard and Wcislo, 2011), these spiders give us a rather emphatic warning that intuition related to the consequences of small brain size can be misleading (Eberhard and Wcislo, 2012).

We are not proposing that brain size is irrelevant. Envisaging a ceiling on what is possible with respect to intelligent behavior still seems valid (Harland and Jackson, 2004) and it still seems to be a matter of common sense that this ceiling will be lower for arthropods with their minute nervous systems and higher for vertebrates with their enormously larger nervous systems. However, if addressing Macphail's null hypothesis is of interest, then arthropod-vertebrate comparisons might be a good place to start. Discussing the null hypothesis only in the context of vertebrate-to-vertebrate comparisons seems excessively arbitrary. It seems to us that, when the goal is to identify qualitative and quantitative differences in intelligence, the context should be inclusive of all animals that express capacities pertaining to intelligence, irrespective of whether they are vertebrates.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version.

FUNDING

This research is supported by grants from the Royal Society of New Zealand [Marsden Fund (UOC305, UOC0507, UOC1301, and UOC1603) and James Cook Fellowship (02/05)] and the National Geographic Society (#WW-146R-17).

ACKNOWLEDGMENTS

This manuscript is dedicated to GC, who unfortunately passed away on 8 November 2019. Her intelligence, kindness, and dedication for research and teaching have been an inspiration for us all, and she is sorely missed. Address correspondence to FC or RG, Department of Psychology, University of Canterbury, Christchurch 8041, New Zealand. Email: fiona.r.cross@gmail.com; randolph.grace@canterbury.ac.nz.

REFERENCES

Anselme, P. (2012). Modularity of mind and the role of incentive motivation in representing novelty. *Anim. Cogn.* 15, 443–459. doi: 10.1007/s10071-012-0499-5

Babu, K. S. (1975). Post embryonic development of the central nervous system of the spider *Argiope aurantia* (Lucas). *J. Morphol.* 146, 325–342. doi: 10.1002/jmor.1051460303

Babu, K. S., and Barth, F. G. (1984). Neuroanatomy of the central nervous system of the wandering spider, *Cupiennius salei* (Arachnida, Araneida). *Zoomorphology* 104, 344–359. doi: 10.1007/BF00312185

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Beecher, M. D. (1988). "Some comments on the adaptationist approach to learning" in *Evolution and learning*. eds. R. C. Bolles and M. D. Beecher (London: Lawrence Erlbaum Assoc. Pubs), 239–248.

Broschard, M. B., Kim, J., Love, B. C., Wasserman, E. A., and Freeman, J. H. (2019). Selective attention in rat visual category learning. *Learn. Mem.* 26, 84–92. doi: 10.1101/lm.048942.118

Bühlmann, C., Cheng, K., and Wehner, R. (2011). Vector-based and landmark-guided navigation in desert ants inhabiting landmark-free and landmark-rich environments. *J. Exp. Biol.* 214, 2845–2853. doi: 10.1242/jeb.054601

Burkart, J. M., Schubiger, M. N., and van Schaik, C. P. (2017). The evolution of general intelligence. *Behav. Brain Sci.* 40:e195. doi: 10.1017/S0140525X16000959

Carey, S. (2004). Bootstrapping & the origin of concepts. *Daedalus* 133, 59–68. doi: 10.1162/001152604772746701

Carey, S., and Barner, D. (2019). Ontogenetic origins of human integer representations. *Trends Cogn. Sci.* 23, 823–835. doi: 10.1016/j.tics.2019.07.004

Catania, K. C. (2010). Born knowing: tentacled snakes innately predict future prey behavior. *PLoS One* 5:e10953. doi: 10.1371/journal.pone.0010953

Chiappe, D., and MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *J. Gen. Psychol.* 132, 5–40. doi: 10.3200/GENP.132.1.5-40

Chittka, L., and Niven, J. (2009). Are bigger brains better? *Curr. Biol.* 19, R995–R1008. doi: 10.1016/j.cub.2009.08.023

Chronicle, E. P., MacGregor, J. N., and Ormerod, T. C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 14–27. doi: 10.1037/0278-7393.30.1.14

Collett, M., and Collett, T. S. (2000). How do insects use path integration for their navigation? *Biol. Cybern.* 83, 245–259. doi: 10.1007/s004220000168

Cross, F. R., and Jackson, R. R. (2009). Cross-modality priming of visual and olfactory selective attention by a spider that feeds indirectly on vertebrate blood. *J. Exp. Biol.* 212, 1869–1875. doi: 10.1242/jeb.028126

Cross, F. R., and Jackson, R. R. (2010a). The attentive spider: search-image use by a mosquito-eating predator. *Ethology* 116, 240–247. doi: 10.1111/j.1439-0310.2009.01731.x

Cross, F. R., and Jackson, R. R. (2010b). Olfactory search-image use by a mosquito-eating predator. *Proc. Biol. Sci.* 277, 3173–3178. doi: 10.1098/rspb.2010.0596

Cross, F. R., and Jackson, R. R. (2014). Specialised use of working memory by *Portia africana*, a spider-eating salticid. *Anim. Cogn.* 17, 435–444. doi: 10.1007/s10071-013-0675-2

Cross, F. R., and Jackson, R. R. (2015). Solving a novel confinement problem by spartaeine salticids that are predisposed to solve problems in the context of predation. *Anim. Cogn.* 18, 509–515. doi: 10.1007/s10071-014-0819-z

Cross, F. R., and Jackson, R. R. (2016). The execution of planned detours by spider-eating predators. *J. Exp. Anal. Behav.* 105, 194–210. doi: 10.1002/jeab.189

Cross, F. R., and Jackson, R. R. (2017). Representation of different exact numbers of prey by a spider-eating predator. *Interface Focus* 7:20160035. doi: 10.1098/rsfs.2016.0035

Cross, F. R., and Jackson, R. R. (2019). *Portia's* capacity to decide whether a detour is necessary. *J. Exp. Biol.* 222:jeb203463. doi: 10.1242/jeb.203463

Davis, H., and Pérusse, R. (1988). Numerical competence in animals: definitional issues, current evidence, and a new research agenda. *Behav. Brain Sci.* 11, 561–579. doi: 10.1017/S0140525X00053437

Deaner, R. O., Isler, K., Burkart, J., and van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain Behav. Evol.* 70, 115–124. doi: 10.1159/000102973

Deaner, R. O., van Schaik, C. P., and Johnson, V. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evol. Psychol.* 4, 149–196. doi: 10.1177/147470490600400114

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York: Simon & Schuster.

Dennett, D. C. (1996). *Kinds of minds: Towards an understanding of consciousness*. New York: Simon & Schuster.

Domjan, M. (2010). *The principles of learning and behavior*. *6th Edn*. Belmont, CA: Wadsworth.

Dukas, R. (ed.) (1998). "Evolutionary ecology of learning" in *Cognitive ecology: The evolutionary ecology of information processing and decision making*. (Chicago: University of Chicago Press), 129–174.

Dyer, A. G. (2012). The mysterious cognitive abilities of bees: why models of visual processing need to consider experience and individual differences in animal performance. *J. Exp. Biol.* 215, 387–395. doi: 10.1242/jeb.038190

Eberhard, W. G. (2011). Are smaller animals behaviourally limited? Lack of clear constraints in miniature spiders. *Anim. Behav.* 81, 813–823. doi: 10.1016/j.anbehav.2011.01.016

Eberhard, W. G., and Wcislo, W. T. (2011). Grade changes in brain–body allometry: morphological and behavioural correlates of brain size in miniature spiders, insects and other invertebrates. *Adv. Insect Physiol.* 40, 155–214. doi: 10.1016/B978-0-12-387668-3.00004-0

Eberhard, W. G., and Wcislo, W. T. (2012). Plenty of room at the bottom? *Am. Sci.* 100, 226–233. doi: 10.1511/2012.96.226

Edelman, G. M. (1998). Building a picture of the brain. *Daedalus* 127, 37–69.

Faisal, A. A., White, J. A., and Laughlin, S. B. (2005). Ion-channel noise places limits on the miniaturization of the brain's wiring. *Curr. Biol.* 15, 1143–1149. doi: 10.1016/j.cub.2005.05.056

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gallistel, C. R. (1990a). Representations in animal cognition: an introduction. *Cognition* 37, 1–22. doi: 10.1016/0010-0277(90)90016-d

Gallistel, C. R. (1990b). *The organization of learning*. Cambridge, Massachusetts: MIT Press.

Gallistel, C. R. (2017). The coding question. *Trends Cogn. Sci.* 21, 498–508. doi: 10.1016/j.tics.2017.04.012

Gallistel, C. R., and King, A. P. (2009). *Memory and the computational brain: Why cognitive science will transform neuroscience*. Chichester: Wiley-Blackwell.

Gebuis, T., Kadosh, R. C., and Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: a critical review. *Acta Psychol.* 171, 17–35. doi: 10.1016/j.actpsy.2016.09.003

Geffner, H. (2013). Computational models of planning. *Wiley Interdiscip. Rev. Cogn. Sci.* 4, 341–356. doi: 10.1002/wcs.1233

Giurfa, M. (2013). Cognition with few neurons: higher-order learning in insects. *Trends Neurosci.* 36, 285–294. doi: 10.1016/j.tins.2012.12.011

Giurfa, M. (2015). Learning and cognition in insects. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 383–395. doi: 10.1002/wcs.1348

Grace, R. C., Carvell, G. E., Morton, N. J., Grice, M., Wilson, A. J., and Kemp, S. (2020). On the origins of computationally complex behavior. *J. Exp. Psychol. Anim. Learn. Cogn.* 46, 1–15. doi: 10.1037/xan0000227

Grush, R. (1997). The architecture of representation. *Philos. Psychol.* 10, 5–23. doi: 10.1080/09515089708573201

Harland, D. P., and Jackson, R. R. (2000). 'Eight-legged cats' and how they see-a review of recent research on jumping spiders (Araneae: Salticidae). *Cimbebasia* 16, 231–240.

Harland, D. P., and Jackson, R. R. (2004). "*Portia* perceptions: the *Umwelt* of an araneophagic jumping spider" in *Complex worlds from simpler nervous systems*. ed. F. R. Prete (Cambridge, Massachusetts: MIT Press), 5–40.

Harland, D. P., Li, D., and Jackson, R. R. (2012). "How jumping spiders see the world" in *How animals see the world: Comparative behavior, biology, and evolution of vision*. eds. O. Lazareva, T. Shimizu and E. A. Wasserman (New York: Oxford University Press), 133–164.

Hauser, M. D., MacNeilage, P., and Ware, M. (1996). Numerical representations in primates. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1514–1517. doi: 10.1073/pnas.93.4.1514

Hill, D. E. (1979). Orientation by jumping spiders of the genus *Phidippus* (Araneae: Salticidae) during the pursuit of prey. *Behav. Ecol. Sociobiol.* 5, 301–322. doi: 10.1007/BF00293678

Hoefler, C. D., and Jakob, E. M. (2006). Jumping spiders in space: movement patterns, nest site fidelity and the use of beacons. *Anim. Behav.* 71, 109–116. doi: 10.1016/j.anbehav.2005.03.033

Jackson, R. R. (1992). Eight-legged tricksters: spiders that specialize at catching other spiders. *Bioscience* 42, 590–598. doi: 10.2307/1311924

Jackson, R. R., and Blest, A. D. (1982). The biology of *Portia fimbriata*, a web-building jumping spider (Araneae, Salticidae) from Queensland: utilization of webs and predatory versatility. *J. Zool.* 196, 255–293. doi: 10.1111/j.1469-7998.1982.tb03504.x

Jackson, R. R., and Carter, C. M. (2001). Geographic variation in reliance on trial-and-error signal derivation by *Portia labiata*, an araneophagic jumping spider from the Philippines. *J. Insect Behav.* 14, 799–827. doi: 10.1023/A:1013041501584

Jackson, R. R., Carter, C. M., and Tarsitano, M. S. (2001). Trial-and-error solving of a confinement problem by a jumping spider, *Portia fimbriata*. *Behaviour* 138, 1215–1234. doi: 10.1163/15685390152822184

Jackson, R. R., Clark, R. J., and Harland, D. P. (2002). Behavioural and cognitive influences of kairomones on an araneophagic jumping spider. *Behaviour* 139, 749–775. doi: 10.1163/156853902320262808

Jackson, R. R., and Cross, F. R. (2011). Spider cognition. *Adv. Insect Physiol.* 41, 115–174. doi: 10.1016/B978-0-12-415919-8.00003-3

Jackson, R. R., and Cross, F. R. (2013). A cognitive perspective on aggressive mimicry. *J. Zool.* 290, 161–171. doi: 10.1111/jzo.12036

Jackson, R. R., Cross, F. R., and Carter, C. M. (2006). Geographic variation in a spider's ability to solve a confinement problem by trial and error. *Int. J. Comp. Psychol.* 19, 282–296.

Jackson, R. R., and Nelson, X. J. (2011). Reliance on trial and error signal derivation by *Portia africana*, an araneophagic jumping spider from East Africa. *J. Ethol.* 29, 301–307. doi: 10.1007/s10164-010-0258-5

Jackson, R. R., and Pollard, S. D. (1996). Predatory behavior of jumping spiders. *Annu. Rev. Entomol.* 41, 287–308. doi: 10.1146/annurev.en.41.010196.001443

Jackson, R. R., Pollard, S. D., Li, D., and Fijn, N. (2002). Interpopulation variation in the risk-related decisions of *Portia labiata*, an araneophagic jumping spider (Araneae, Salticidae), during predatory sequences with spitting spiders. *Anim. Cogn.* 5, 215–223. doi: 10.1007/s10071-002-0150-y

Jackson, R. R., and Wilcox, R. S. (1993). Observations in nature of detouring behaviour by *Portia fimbriata*, a web-invading aggressive mimic jumping spider from Queensland. *J. Zool.* 230, 135–139. doi: 10.1111/j.1469-7998.1993.tb02677.x

Jakob, E. M., Skow, C. D., and Long, S. (2011). "Plasticity, learning and cognition" in *Spider behaviour: Flexibility and versatility*. ed. M. E. Herberstein (Cambridge: Cambridge University Press), 307–347.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Johnston, T. D. (1982). The selective costs and benefits of learning: an evolutionary analysis. *Adv. Study Behav.* 12, 65–106. doi: 10.1016/S0065-3454(08)60046-7

Jones, G. (2003). Testing two cognitive theories of insight. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1017–1027. doi: 10.1037/0278-7393.29.5.1017

Kabadayi, C., Bobrowicz, K., and Osvath, M. (2018). The detour paradigm in animal cognition. *Anim. Cogn.* 21, 21–35. doi: 10.1007/s10071-017-1152-0

Kupfermann, I. (1987). Bony argument. *Behav. Brain Sci.* 10:673. doi: 10.1017/S0140525X00055199

Land, M. F., and Nilsson, D. -E. (2012). *Animal eyes*. *2nd Edn*. Oxford: Oxford University Press.

Logan, C. J., Avin, S., Boogert, N., Buskell, A., Cross, F. R., Currie, A., et al. (2018). Beyond brain size: uncovering the neural correlates of behavioral and cognitive specialization. *Comp. Cogn. Behav. Rev.* 13, 55–90. doi: 10.3819/CCBR.2018.130008

Lorenz, K. (1965). *Evolution and modification of behavior*. Chicago: University of Chicago Press.

Macphail, E. M. (1985). Vertebrate intelligence: the null hypothesis. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 308, 37–51. doi: 10.1098/rstb.1985.0008

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–656. doi: 10.1017/S0140525X00054984

Macphail, E. M., and Bolhuis, J. J. (2001). The evolution of intelligence: adaptive specializations *versus* general process. *Biol. Rev. Camb. Philos. Soc.* 76, 341–364. doi: 10.1017/S146479310100570X

Maddison, W. P. (2015). A phylogenetic classification of jumping spiders (Araneae: Salticidae). *J. Arachnol.* 43, 231–292. doi: 10.1636/arac-43-03-231-292

Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., et al. (2003). Individual differences in the expression of a "general" learning ability in mice. *J. Neurosci.* 23, 6423–6433. doi: 10.1523/JNEUROSCI.23-16-06423.2003

Nelson, X. J., and Jackson, R. R. (2012). The role of numerical competence in a specialized predatory strategy of an araneophagic spider. *Anim. Cogn.* 15, 699–710. doi: 10.1007/s10071-012-0498-6

Nieder, A. (2019). *A brain for numbers: The biology of the number instinct*. Cambridge, Massachusetts: MIT Press.

Niven, J. E., and Chittka, L. (2016). Evolving understanding of nervous system evolution. *Curr. Biol.* 26, R937–R941. doi: 10.1016/j.cub.2016.09.003

Niven, J. E., and Farris, S. M. (2012). Miniaturization of nervous systems and neurons. *Curr. Biol.* 22, R323–R329. doi: 10.1016/j.cub.2012.04.002

O'Neill, E. (2015). Relativizing innateness: innateness as the insensitivity of the appearance of a trait with respect to specified environmental variation. *Biol. Philos.* 30, 211–225. doi: 10.1007/s10539-014-9465-8

Ortega-Escobar, J., and Ruiz, M. A. (2017). Role of the different eyes in the visual odometry in the wolf spider *Lycosa tarantula* (Araneae, Lycosidae). *J. Exp. Biol.* 220, 259–265. doi: 10.1242/jeb.145763

Papaj, D. R. (1986). Interpopulation differences in host preferences and the evolution of learning in the butterfly, *Battus philenor*. *Evolution* 40, 518–530. doi: 10.1111/j.1558-5646.1986.tb00504.x

Pepperberg, I. M. (2002). The value of the Piagetian framework for comparative cognitive studies. *Anim. Cogn.* 5, 177–182. doi: 10.1007/s10071-002-0148-5

Pepperberg, I. M., and Kozak, F. A. (1986). Object permanence in the African grey parrot (*Psittacus erithacus*). *Anim. Learn. Behav.* 14, 322–330. doi: 10.3758/BF03200074

Platnick, N. I. (2020). World spider catalog. Natural history museum Bern, Version 21.0. Available at: http://www.wsc.nmbe.ch/ (Accessed September 3, 2020).

Reader, S. M., and Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proc. Natl. Acad. Sci.* 99, 4436–4441. doi: 10.1073/pnas.062041299

Rucker, R. (1987). *Mind tools: The five levels of mathematical reality*. Boston: Houghton Mifflin.

Shettleworth, S. J. (2010). *Cognition, evolution, and behavior*. *2nd Edn*. New York: Oxford University Press.

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, Massachusetts: MIT Press.

Skinner, B. F. (1981). Selection by consequences. *Science* 213, 501–504. doi: 10.1126/science.7244649

Smith, J. D., Ashby, F. G., Berg, M. E., Murphy, M. S., Spiering, B., Cook, R. G., et al. (2011). Pigeons' categorization may be exclusively nonanalytic. *Psychon. Bull. Rev.* 18, 414–421. doi: 10.3758/s13423-010-0047-8

Smith, J. D., Beran, M. J., Crossley, M. J., Boomer, J., and Ashby, F. G. (2010). Implicit and explicit category learning by macaques (*Macaca mulatta*) and humans (*Homo sapiens*). *J. Exp. Psychol. Anim. Behav. Process.* 36, 54–65. doi: 10.1037/a0015892

Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., et al. (2012). Implicit and explicit categorization: a tale of four species. *Neurosci. Biobehav. Rev.* 36, 2355–2369. doi: 10.1016/j.neubiorev.2012.09.003

Smith, B. P., and Litchfield, C. A. (2010). How well do dingoes, *Canis dingo*, perform on the detour task? *Anim. Behav.* 80, 155–162. doi: 10.1016/j.anbehav.2010.04.017

Suter, R. B. (2013). Spider locomotion on the water surface: biomechanics and diversity. *J. Arachnol.* 41, 93–101. doi: 10.1636/M13-14

Tarsitano, M. S., and Jackson, R. R. (1997). Araneophagic jumping spiders discriminate between detour routes that do and do not lead to prey. *Anim. Behav.* 53, 257–266. doi: 10.1006/anbe.1996.0372

Wasserman, J. D. (2012). "A history of intelligence assessment: the unfinished tapestry" in *Contemporary intellectual assessment: Theories, tests, and issues*. eds. D. P. Flanagan and P. L. Harrison (New York: The Guilford Press), 3–55.

Watson, R. A., and Szathmáry, E. (2016). How can evolution learn? *Trends Ecol. Evol.* 31, 147–157. doi: 10.1016/j.tree.2015.11.009

Webb, B. (2012). Cognition in insects. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 367, 2715–2722. doi: 10.1098/rstb.2012.0218

Winters, S., Dubuc, C., and Higham, J. P. (2015). Perspectives: the looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology* 121, 625–640. doi: 10.1111/eth.12378

Wynn, K. (1992). Addition and subtraction by human infants. *Nature* 358, 749–750. doi: 10.1038/358749a0

Zentall, T. R. (2018). Morgan's canon: is it still a useful rule of thumb? *Ethology* 124, 449–457. doi: 10.1111/eth.12750

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Psychology

OPINION
published: 28 October 2020
doi: 10.3389/fpsyg.2020.581899

Simplicity From Complexity in Vertebrate Behavior: Macphail (1987) Revisited

Stephen B. Fountain[1]*, **Katherine H. Dyer**[2] **and Claire C. Jackman**[2]

[1] *Department of Psychological Sciences and Brain Health Research Institute, Kent State University, Kent, OH, United States,*
[2] *Department of Psychological Sciences, Kent State University, Kent, OH, United States*

Keywords: associations, hierarchical plan, rule learning, multiple cognitive processes, multiple-item memory, stimulus-response learning, timing, internal clock

INTRODUCTION

"Causality is a constraint common to all ecological niches."
Macphail, 1987

Macphail (1987) claimed that all vertebrate nervous systems rely on detecting, encoding, and acting upon causality, and that there are no differences in intelligence between vertebrate species. The latter claim constitutes what is widely described as the "null hypothesis." We examine the null hypothesis from the perspective of how vertebrates learn based on the order of events—that is, we will examine the ubiquity and foundations of sequential learning and memory in vertebrate behavior. The claim will be that several neurobehavioral systems subserve vertebrate sequence learning and that these and perhaps other systems together simplify encoding environmental complexity during learning and provide the foundation for performing complex but highly organized behavior.

What constitutes "causality" as coded by vertebrate brains? From an animal's perspective, behavior is inherently sequential and relevant events in the environment occur in probabilistic relationships with behavior. In the laboratory, these relationships may be highly constrained, as in Pavlovian conditioning, instrumental conditioning, and even more complex cognitive or neuroscience-oriented paradigms employing animals as complex as humans. This fact does not imply that the foundations of vertebrate behavior depend on a single underlying mechanism, though it should be no surprise that associative theory is a powerful approach to understanding and describing such behavior. On the contrary, much evidence suggests that vertebrate behavior is the result of multiple complementary systems that converge, interact, and often compete. These systems produce the remarkably adaptive and complex behavior befitting the remarkably diverse and complex environments in which vertebrates live. Yet, despite the diversity of scenarios in which behavior is played out, causality is universally available for organisms to exploit to survive and perpetuate the species. We propose that non-human vertebrates, like humans, abstract representations of simple causal relationships between events from complex environments, that is, they encode "simplicity from complexity." Furthermore, vertebrates may share separate interacting systems for different types of sequential information.

Note that the critical and most challenging test of MacPhail's null hypothesis claim is *not* that the simplest processes are conserved "upward" to the most complex vertebrates, but that the most complex processes can be observed when we look "downward" toward the simplest vertebrates. To be explicit, we ask whether vertebrates in general extract "simplicity from complexity" through common learning mechanisms and neural substrates. As a start, this is the question we examine directly comparing humans, rats, mice, and pigeons. We conclude that additional evidence is needed to confirm our speculations regarding the generality of learning consisting of extracting

Edited by:
Damian Scarf,
University of Otago, New Zealand

Reviewed by:
Gema Martin-Ordas,
University of Stirling, United Kingdom

****Correspondence:***
Stephen B. Fountain
sfountai@kent.edu

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *10 July 2020*
Accepted: *18 September 2020*
Published: *28 October 2020*

Citation:
Fountain SB, Dyer KH and
Jackman CC (2020) Simplicity From
Complexity in Vertebrate Behavior:
Macphail (1987) Revisited.
Front. Psychol. 11:581899.
doi: 10.3389/fpsyg.2020.581899

"simplicity from complexity" in all vertebrates, but that Macphail (1987) was not far wrong in proposing that common processes may underlie vertebrate cognitive abilities though not necessarily resulting in equivalent capacities.

WHAT COGNITIVE MECHANISMS ARE COMMON TO VERTEBRATE SPECIES?

Lashley (1951) rejected the notion that sequential behavior was accounted for by simple reflex chaining and argued instead for cognitive encoding of hierarchical plans. This notion contributed to the development of cognitive research in both human and non-human animals which continues to this day. Recently, Rosenbaum et al. (2007) shifted the theoretical concern from Lashley's focus on the nature of encoded sequence structures to identifying and describing the processes that contribute to the emergence of sequence structures in behavior. According to their view, even individual component movements are controlled by hierarchically organized plans (Rosenbaum et al., 2007).

Our work on non-human animal sequential rule abstraction, learning, and memory (retention) has also been motivated by Lashley's insights regarding the human capacity to use mental representations of pattern structures to plan ahead. As Reber (1993) and others have shown, plans need not be conscious or verbal. Early on, non-verbal pattern learning methods were developed to study sequence learning in rats (Hunter, 1920; Capaldi et al., 1966; Hulse, 1978), in monkeys (Straub et al., 1979; Straub and Terrace, 1981; Terrace, 1987), and in humans (Restle, 1970, 1973; Restle and Brown, 1970a,b). We developed a more general Serial Multiple Choice (SMC) task to analyze how rats, mice, pigeons, and humans learn complex sequential patterns of responses. A typical method allows animals to respond to circular arrays of 8 items, for example, a circular array of 8 manipulanda, one each on the walls of an octagonal chamber, or a circular array of 8 nosepoke locations on a touchscreen. In both cases, rats learn long but highly-organized patterns of responses on the manipulanda (e.g., Fountain et al., 2012; Garlick et al., 2017). Evidence from a variety of studies using this task indicate that all these species employ multiple concurrent cognitive processes to encode and produce complex sequential behavior (Fountain, 2008; Fountain et al., 2012). In the SMC task, the animal learns to make responses in a circular array (8-walled chamber or a circular touchscreen array). The required sequence of responses is typically a highly organized serial pattern of responses that may be characterized by multi-level hierarchical organization and "exceptions-to-the-rule." Such patterns recruit multiple concurrent cognitive processes, namely, processes for encoding stimulus-response associations, timing/counting of events, and pattern structure including simple and higher-order rules (e.g., Muller and Fountain, 2010, 2016).

Furthermore, the learning and memory systems involved depend on different behavioral and neural systems as shown by dissociations observed in adulthood long after chronic adolescent drug exposure (Pickens et al., 2013; Renaud et al., 2015; Rowan et al., 2015; Renaud and Fountain, 2016; Sharp et al., 2019). Similar dissociations of cognitive behavioral systems can be seen in normal rats administered acute muscarinic cholinergic blockade (Pickens et al., 2013; Renaud et al., 2015; Rowan et al., 2015; Renaud and Fountain, 2016; Sharp et al., 2019). The SMC method in mice and rats produces data comparable to data in humans in an analogous task (Fountain and Rowan, 1995) and rats use multiple cognitive processes concurrently: rule-learning, stimulus-response (S-R) learning, discrimination learning, and multiple-item memory (Muller and Fountain, 2010, 2016). Fountain and Benson (2006) demonstrated chunking, rule-learning, and multiple-item memory when rats learned to anticipate the elements of two interleaved serial patterns. Mice show more limited abilities, but do show evidence of multiple concurrent learning processes (Fountain et al., 1999). Finally, pigeons in a comparable touchscreen task were able to abstract sequence structure (Garlick et al., 2017). Rats and, to a lesser degree, mice concurrently encode stimulus-response associations, time and count events, and abstract rules describing pattern structure (Muller and Fountain, 2010, 2016).

Taken together, these results strongly support the view that pigeons, mice, rats, and humans likely share multiple dissociable serial pattern learning and memory systems that encode multiple types of sequential information (Fountain et al., 2012). With reference to the underlying processes we have discussed in humans, rats, mice, and pigeons—namely, processes for encoding stimulus-response associations, timing/counting events, and rule-learning—are these common to all vertebrates? A broader survey of more species and new species-specific methods would be required to answer this question.

CAN A SINGLE ASSOCIATIVE PROCESS ACCOUNT FOR ALL NON-HUMAN "INTELLIGENCE?"

In our attempts to characterize how rats learn to anticipate items in a sequence, in one approach we sought to use mathematical models to determine whether a single mechanism might account for all sequence learning in rats. One early success in this line of research that bears on this question was a mathematical model we used to determine whether a simple mathematical model based on simple associative principles could account for rat serial pattern learning for sequences of food quantities presented in a runway (Wallace and Fountain, 2002, 2003). We used a modified version of Metcalfe's Composite Holographic Associative Recall Model (CHARM; Metcalfe, 1990). In CHARM, items to be remembered are represented by vectors of random numbers, where each vector represents an item to be remembered. Our model based on CHARM is named the Sequential Pairwise Associative Memory (SPAM) model. SPAM used the same system of creating vectors of random numbers to represent food quantities, but vectors for different quantities varied in similarity to represent a range of food quantities from small to large. When the appropriate vectors for food quantities were stored for different sequential patterns, SPAM was able to simulate a full range of effects previously reported in the rat sequential learning literature of the time (Wallace and Fountain, 2002, 2003). On the other hand, that model and variations of it have so far failed to account for the variety of differences in learning phenomena revealed in how rats learn highly-structured sequences (Muller and Fountain, 2010, 2016). Nevertheless, the failure of one model, no matter how successful within a single domain yet failing in another, does not preclude the possibility that it might be possible to develop a successful single-process model that would be consistent with Macphail's claim.

Some aspects of cross-species behavioral comparisons of rats, mice, and pigeons (Fountain et al., 1999; Kelley and Rowan, 2004; Garlick et al., 2017) do not easily fit within the SPAM framework. SPAM does not account for several very robust aspects of serial pattern learning in rats in the SMC task. For example, a large body of our work indicates that independent processes mediate different types of learning via dissociable systems that operate concurrently for encoding simple associations, serial position, and lower- and higher-order rule structure (Muller and Fountain, 2010, 2016). These observations suggest that much research is needed to determine whether the same patterns of results would be observed in species other than the species we have already examined, namely, humans, rats, mice, and pigeons.

THE ROLE AND POWER OF "SIMPLE ASSOCIATIONS," "HIERARCHICAL PLANS," AND "IMPLICIT KNOWLEDGE" IN VERTEBRATE COGNITIVE CAPACITIES

One conclusion we draw from the foregoing is that Lashley (1951) was correct to reject simple associative chaining accounts of sequential behavior. Lashley argued instead for a more cognitive account proposing that humans encoded and used hierarchical plans based on sequence structure and grouping which now seem fundamental to an analysis of animal sequence learning (Fountain et al., 2000; Muller and Fountain, 2010, 2016). Hierarchical plans, whether implicit or explicit, may be more fundamental than one might suppose given that even 8- to 10-month-old infants "exhibited sensitivity to the difference between hierarchical and non-hierarchical structure" and that the ability "to perceive, learn, and generalize recursive, hierarchical, pattern rules emerges in infancy" (Lewkowicz et al., 2018). The foregoing suggests that rule-learning in infancy must be implicit, and perhaps non-human vertebrates in general likewise can learn highly-organized implicit structures like those we have observed in pigeons, mice, and rats. This notion of hierarchical organization in behavior also unites our conceptions of behaviors as diverse as foraging (Feeney et al., 2011), bird song production and perception (e.g., Cazala et al., 2019), and sequential behavior (Swartz et al., 1991, 2000; Terrace and Chen, 1991a,b; Swartz and Himmanen, 2002; Suge and Okanoya, 2010; Spierings et al., 2015; Ramkumar et al., 2016). We go further to claim that vertebrate behavioral systems in their diversity encode different responses or different types of information, including complex associations, number (Brannon and Terrace, 2002), and time via internal clock processes (Tucci et al., 2014).

Macphail (1987) argued that all vertebrate nervous systems rely on detecting, encoding, and acting upon causality, and that there are no differences in intelligence between vertebrate species. We have described how a range of studies across paradigms and a variety of species support the view that complex learning processes may very well be broadly or even universally available to vertebrates. A challenge for the field is to develop experimental paradigms for assessing potentially common mechanisms in diverse species. At a foundational level, Macphail's claim continues to challenge all empiricists and theorists to consider the power of even simple neural systems to account for animals' ability to encode simplicity in terms of neural representation from the complexity of the surrounding environmental milieu.

AUTHOR CONTRIBUTIONS

All authors contributed significantly to background reading, organization of the paper, and writing.

REFERENCES

Brannon, E. M., and Terrace, H. S. (2002). "*The evolution and ontogeny of ordinal numerical ability*," in *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, eds M. Bekoff, C. Allen, and G. M. Burghardt (Cambridge, MA: MIT Press) 197–204.

Capaldi, E. J., Veatch, R. L., and Stefaniak, D. E. (1966). Stimulus control of patterning behavior. *J. Comp. Physiol. Psychol.* 61, 161–164. doi: 10.1037/h0022871

Cazala, A., Giret, N., Edeline, J. M., and Del Negro, C. (2019). Neuronal encoding in a high-level auditory area: from sequential order of elements to grammatical structure. *J. Neurosci.* 39, 6150–6161. doi: 10.1523/JNEUROSCI.2767-18.2019

Feeney, M. C., Roberts, W. A., and Sherry, D. F. (2011). Black-capped chickadees (*Poecile atricapillus*) anticipate future outcomes of foraging choices. *J. Exp. Psychol. Anim. Behav. Process.* 37, 30–40. doi: 10.1037/a0019908

Fountain, S. B. (2008). Pattern structure and rule induction in sequential learning. *Comp. Cogn. Behav. Rev.* 3, 66–85. doi: 10.3819/ccbr.2008.30004

Fountain, S. B., Benson, A. M., and Wallace, D. G. (2000). Number, but not rhythmicity, of temporal cues determines phrasing effects in rat serial-pattern learning. *Learn. Motiv.* 31, 301–322. doi: 10.1006/lmot.2000.1057

Fountain, S. B., and Benson, D. M. Jr. (2006). Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. *Learn. Motiv.* 37, 95–112. doi: 10.1016/j.lmot.2005.09.002

Fountain, S. B., Krauchunas, S. M., and Rowan, J. D. (1999). Serial-pattern learning in mice: pattern structure and phrasing. *Psychol. Rec.* 49, 173–192. doi: 10.1007/BF03395315

Fountain, S. B., and Rowan, J. D. (1995). Coding of hierarchical versus linear pattern structure in rats and humans. *J. Exp. Psychol. Anim. Behav. Process.* 21, 187–202. doi: 10.1037/0097-7403.21.3.187

Fountain, S. B., Rowan, J. D., Muller, M. D., Kundey, S. M. A., Pickens, L. R. G., and Doyle, K. E. (2012). "The organization of sequential behavior: conditioning, memory, and abstraction," in *Handbook of Comparative Cognition*, eds E. A. Wasserman and T. R. Zentall (Oxford: Oxford University Press), 594–614.

Garlick, D., Fountain, S. B., and Blaisdell, A. P. (2017). Serial pattern learning in pigeons: rule-based or associative? *J. Exp. Psychol. Anim. Learn. Cogn.* 43, 30–47. doi: 10.1037/xan0000109

Hulse, S. H. (1978). "Cognitive structure and serial pattern learning by animals," in *Cognitive Processes in Animal Behavior*, eds S. H. Hulse, H. Fowler, and W. K. Honig (Hillsdale, NJ: Erlbaum), 311–340.

Hunter, W. S. (1920). The temporal maze and kinaesthetic sensory processes in the white rat. *Psychobiology* 2, 1–17. doi: 10.1037/h0073855

Kelley, B. M., and Rowan, J. D. (2004). Long-term, low-level adolescent nicotine exposure produces dose-dependent changes in cocaine sensitivity and reward in adult mice. *Int. J. Dev. Neurosci.* 22, 339–348. doi: 10.1016/j.ijdevneu.2004.04.002

Lashley, K. S. (1951). "The problem of serial order in behavior," in *Cerebral Mechanisms in Behavior*, ed L. A. Jeffress (New York, NY: Wiley), 112–146.

Lewkowicz, D. J., Schmuckler, M. A., and Mangalindan, D. M. J. (2018). Learning of hierarchical serial patterns emerges in infancy. *Dev. Psychobiol.* 60, 243–255. doi: 10.1002/dev.21614

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695. doi: 10.1017/S0140525X00054984

Metcalfe, J. (1990). Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *J. Exp. Psychol. Genrl.* 119, 145–160. doi: 10.1037/0096-3445.119.2.145

Muller, M. D., and Fountain, S. B. (2010). Concurrent cognitive processes in rat serial pattern learning: item memory, serial position, and pattern structure. *Learn. Motiv.* 41, 252–272. doi: 10.1016/j.lmot.2010.08.003

Muller, M. D., and Fountain, S. B. (2016). Concurrent cognitive processes in rat serial pattern learning: II. Discrimination learning, rule learning, chunk length, and multiple-item memories. *J. Exp. Anal. Behav.* 105, 155–175. doi: 10.1002/jeab.186

Pickens, L. R. G., Rowan, J. D., Bevins, R. A., and Fountain, S. B. (2013). Sex differences in adult cognitive deficits after adolescent nicotine exposure in rats. *Neurotoxicol. Teratol.* 38, 72–78. doi: 10.1016/j.ntt.2013.05.001

Ramkumar, P., Acuna, D. E., Berniker, M., Grafton, S. T., Turner, R. S., and Kording, K. P. (2016). Chunking as the result of an efficiency computation trade-off. *Nat. Commun.* 7:12176. doi: 10.1038/ncomms12176

Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge.* New York, NY: Oxford University Press.

Renaud, S. M., and Fountain, S. B. (2016). Transgenerational effects of adolescent nicotine exposure in rats: evidence for cognitive deficits in adult female offspring. *Neurotoxicol. Teratol.* 56, 47–54. doi: 10.1016/j.ntt.2016.06.002

Renaud, S. M., Pickens, L. R., and Fountain, S. B. (2015). Paradoxical effects of injection stress and nicotine exposure experienced during adolescence on learning in a serial multiple choice (SMC) task in adult female rats. *Neurotoxicol. Teratol.* 48, 40–48. doi: 10.1016/j.ntt.2014.12.003

Restle, F. (1970). Theory of serial pattern learning: structural trees. *Psychol. Rev.* 77, 481–495. doi: 10.1037/h0029964

Restle, F. (1973). Serial pattern learning: higher order transitions. *J. Exp. Psychol.* 99, 61–69. doi: 10.1037/h0034751

Restle, F., and Brown, E. R. (1970a). "Organization of serial pattern learning," in *Psychology of Learning and Motivation*, ed G. H. Bower (New York, NY: Academic Press), 249–331.

Restle, F., and Brown, E. R. (1970b). Serial pattern learning. *J. Exp. Psychol.* 83, 120–125. doi: 10.1037/h0028530

Rosenbaum, D. A., Cohen, R. G., Jax, S. A., Weiss, D. J., and van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Hum. Mov. Sci.* 26, 525–554. doi: 10.1016/j.humov.2007.04.001

Rowan, J. D., Mccarty, M. K., Kundey, S. M. A., Osburn, C. D., Renaud, S. M., Kelley, B. M., et al. (2015). Adolescent exposure to methylphenidate impairs serial pattern learning in the serial multiple choice (SMC) task in adult rats. *Neurotoxicol. Teratol.* 51, 21–26. doi: 10.1016/j.ntt.2015.07.007

Sharp, J. L., Miller-cahill, M. E., Renaud, S. M., Kundey, S. M. A., Kelley, B. M., Matoushek, A. W., et al. (2019). Adolescent exposure to fluoxetine impairs serial pattern learning in the serial multiple choice (SMC) task in adult rats. *Neurobiol. Learn. Mem.* 164:107048. doi: 10.1016/j.nlm.2019.107048

Spierings, M., de Weger, A., and Ten Cate, C. (2015). Pauses enhance chunk recognition in song element strings by zebra finches. *Anim. Cogn.* 18, 867–874. doi: 10.1007/s10071-015-0855-3

Straub, R., Seidenberg, M., Bever, T. G., and Terrace, H. S. (1979). Serial learning in the pigeon. *J. Exp. Anal. Behav.* 32, 137–148. doi: 10.1901/jeab.1979.32-137

Straub, R., and Terrace, H. S. (1981). Generalization of serial learning in the pigeon. *Anim. Learn. Behav.* 9, 454–468. doi: 10.3758/BF03209775

Suge, R., and Okanoya, K. (2010). Perceptual chunking in the self-produced songs of Bengalese finches (*Lonchura striata* var. domestica). *Anim. Cogn.* 13, 515–523. doi: 10.1007/s10071-009-0302-4

Swartz, K. B., Chen, S., and Terrace, H. S. (1991). Serial learning by rhesus monkeys: I. acquisition and retention of multiple four-item lists. *J. Exp. Psychol. Anim. Behav. Process.* 17, 396–410. doi: 10.1037/0097-7403.17.4.396

Swartz, K. B., Chen, S., and Terrace, H. S. (2000). Serial learning by rhesus monkeys: II. learning four-item lists by trial and error. *J. Exp. Psychol. Anim. Behav. Process.* 26, 274–285. doi: 10.1037/0097-7403.26.3.274

Swartz, K. B., and Himmanen, S. A. (2002). "Serial learning by rhesus monkeys: Strategies for list learning and execution," in *Animal Cognition and Sequential Behavior: Behavioral, Biological, and Computational Perspectives*, eds S. B. Fountain, M. Bunsey, J. H. Danks, and M. K. McBeath (Boston, MA: Kluwer Academic), 58–67.

Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature* 325, 149–151. doi: 10.1038/325149a0

Terrace, H. S., and Chen, S. (1991a). Chunking during serial learning by a pigeon: II. integrity of a chunk on a new list. *J. Exp. Psychol. Anim. Behav. Process.* 17, 94–106. doi: 10.1037/0097-7403.17.1.94

Terrace, H. S., and Chen, S. (1991b). Chunking during serial learning by a pigeon: III. What are the necessary conditions for establishing a chunk? *J. Exp. Psychol. Anim. Behav. Process.* 17, 107–118. doi: 10.1037/0097-7403.17.1.107

Tucci, V., Buhusi, C. V., Gallistel, R., and Meck, W. H. (2014). Towards an integrated understanding of the biology of timing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20120470. doi: 10.1098/rstb.2012.0470

Wallace, D. G., and Fountain, S. B. (2002). What is learned in sequential learning? an associative model of reward magnitude serial-pattern learning. *J. Exp. Psychol. Anim. Behav. Process.* 28, 43–63. doi: 10.1037/0097-7403.28.1.43

Wallace, D. G., and Fountain, S. B. (2003). An associative model of rat serial pattern learning in 3-element sequences. *Q. J. Exp. Psychol. Comp. Physiol. Psychol.* 56B, 301–320. doi: 10.1080/02724990244000304

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

REVIEW
published: 21 January 2021
doi: 10.3389/fpsyg.2020.618636

Macphail (1987) Revisited: Pigeons Have Much Cognitive Behavior in Common With Humans

*Thomas R. Zentall**

Department of Psychology, University of Kentucky, Lexington, KY, United States

The hypothesis proposed by Macphail (1987) is that differences in intelligent behavior thought to distinguish different species were likely attributed to differences in the *context* of the tasks being used. Once one corrects for differences in sensory input, motor output, and incentive, it is likely that all vertebrate animals have comparable intellectual abilities. In the present article I suggest a number of tests of this hypothesis with pigeons. In each case, the evidence suggests that either there is evidence for the cognitive behavior, or the pigeons suffer from biases similar to those of humans. Thus, Macphail's hypothesis offers a challenge to researchers to find the appropriate conditions to bring out in the animal the cognitive ability being tested.

Keywords: Macphail, comparative cognition, cognitive biases, animal intelligence, pigeons

Edited by:
Giorgio Vallortigara,
University of Trento, Italy

Reviewed by:
Roland Pusch,
Ruhr University Bochum, Germany
Rosa Rugani,
University of Padua, Italy

Correspondence:
Thomas R. Zentall
zentall@uky.edu

Specialty section:
This article was submitted to
Comparative Psychology,
a section of the journal
Frontiers in Psychology

Received: *17 October 2020*
Accepted: *20 November 2020*
Published: *21 January 2021*

Citation:
Zentall TR (2021) Macphail (1987) Revisited: Pigeons Have Much Cognitive Behavior in Common With Humans. Front. Psychol. 11:618636. doi: 10.3389/fpsyg.2020.618636

INTRODUCTION

In a classic article, Macphail (1987) made the remarkable claim that differences among vertebrate species in the acquisition of tasks thought to be a measure of intelligence, can be attributed largely to differences in contextual variables. In particular, those contextual differences are likely attributable to differences in the animal's perception of the task, the motor skills required, or to the animal's motivation for the rewards involved, rather than to differences in intellect. In comparisons between species, differences in those factors may give the impression of differences in intellectual ability.

An example of how differences in task performance between species can be misleading can be seen readily in research on learning set. Learning set, sometimes referred to as learning-to-learn, is defined as the improvement in discrimination learning that comes with experience with successive discriminations. For example, Harlow (1949) found that when monkeys were given simultaneous discrimination training between pairs of three-dimensional objects, the rate of acquisition improved with as the number of discriminations increased. Specifically, the accuracy of the monkeys on the second trial of a discrimination increased from about 65% correct on early discriminations, to about 98% correct after about 60 discriminations. Thus, after considerable training, based on the outcome of Trial 1, the monkeys appeared to develop a win-stay/lose-shift strategy that they could apply on Trial 2 and thereafter. This strategy has been interpreted as a higher cognitive ability.

When Kay and Oldfield-Box (1965) trained rats on a similar set of discriminations involving three-dimensional objects, the rats improved to only about 75% correct on Trials 2–10 after 78 discrimination problems. Based on this difference in findings, and consistent with one's intuitive belief about the natural order of animal intelligence, one might conclude that monkeys are more intelligent than rats.

Not long after, however, Slotnick and Katz (1974) reasoned that the visual system of the rat may not be ideal for learning visual discriminations. They reasoned that rats might do better with such learning if the discriminations were better suited to their sensory abilities. They tested this hypothesis by giving the rats a series of olfactory discriminations and found that the rate at which the rats learned the successive discriminations improved much faster than with visual discriminations and rivaled that of the monkeys.

The task for psychologists who study comparative cognition is to find the methodology best suited for the species studied. In a sense, one needs to find the best input, output, and motivational conditions appropriate to the animal. The problem, of course, is how to know when one has found the ideal set of variables for the species being studied. How can experimenters take the perspective of the animal? How does one decide that the species does not have a particular capacity?

One approach is to view Macphail's hypothesis as a challenge. Macphail's hypothesis can serve as a useful model for an approach to the study of comparative cognition. A good rule of thumb is, when designing an experiment to test for an animal's cognitive capacity, one should attempt to consider the task from the stimulus, response, and motivational perspective of the animal.

Most of my research has been done with pigeons. I have chosen pigeons, in part, because they are highly visual animals and it is relatively easy to manipulate colors and shapes that are quite easy for them to discriminate. Also, pigeons naturally peck for their food, so pecking at the stimuli is relatively easy for them to learn. Finally, as they are granivors it is relatively easy to motivate them with grain as a reinforcer.

In the remainder of this paper I will describe several of the presumed cognitive abilities attributable to humans (and sometimes to non-human primates) and describe how we have attempted to ask if pigeons too have at least some of this ability. The set of abilities described in this article is not meant to be comprehensive. It is merely a sample of the cognitive abilities that I have studied. Furthermore, it is not meant to examine the comparable ability of other species. The purpose of this enumeration of cognitive abilities is merely to show some of the breadth of competencies that can be found in one particular species, the pigeon. Most of this research was conducted in an operant box with stimuli projected on pecking keys and reinforcement provided from a mixed grain feeder. The conclusion that I have come to in conducting these lines of research is that Macphail's hypothesis has a lot to be said for it. Furthermore, I am pleased to admit that pursuing this approach to comparative cognition research has been a very rewarding experience.

COMPARATIVE COGNITION

The Sameness Concept

The typical method to assess concept learning in animals is to train them with one set of stimuli and ask if they can apply that conceptual rule they have learned to new stimuli. For example, pigeons can easily learn a task called matching-to-sample with colored stimuli, a task that has the potential to develop a sameness rule. This research generally involves an operant box with three pecking keys. The stimulus is projected on middle key is the sample and the stimuli projected on the two side keys are the comparison stimuli. Specifically, for example, if the sample is red, choice of the red comparison stimulus is reinforced, if the sample is green choice of the green comparison stimulus is reinforced. To test for a sameness rule, one should transfer the pigeons to novel stimuli. We have found that when pigeons are transferred to novel blue and yellow stimuli, there is some evidence of positive transfer (Zentall and Hogan, 1974). However, it is possible that stimulus generalization between the training colors and the testing colors played a role in the transfer found. More convincing evidence was found when the training was with shapes and the transfer task involved colors (Zentall and Hogan, 1976). But there is an inherent problem with transfer designs that involve novel stimuli. Pigeons tend to be neophobic and there is generally a substantial initial disruption of matching accuracy that could be attributed to the novelty of the transfer stimuli.

An alternative approach was attempted by Zentall et al. (2018). They trained pigeons on either a matching or mismatching task with four colors. In training, although each color served as a sample and as the matching comparison, with each sample only one color served as the mismatching stimulus (see **Figure 1** for the design of this experiment). This meant that all four colors had served as sample, correct, and incorrect stimulus in one of the four matching problems. Following training, on test trials, either the matching or the mismatching comparison color was replaced with a familiar color but one that was never before seen with that sample. Results for the matching task were as one might expect. Replacing the matching stimulus resulted in a sharp drop in accuracy, whereas replacing the mismatching stimulus resulted in only a small drop in accuracy. The results with the mismatching task, however, were surprising (see **Figure 2**). Replacing the mismatching stimulus (the correct stimulus from training) resulted in only a small drop in matching accuracy, whereas replacing the matching stimulus (the incorrect stimulus from training) resulted in a large drop in matching accuracy. These results were not only unexpected but are inconsistent with Skinner's (1950) prediction that all conditional discriminations (including matching and mismatching) involve the learning of simple sample-correct-comparison stimulus-response chains. The results of Zentall et al. (2018; see also Zentall et al., 1981) suggest that the pigeons use the matching stimulus as the basis of choice in both the matching and mismatching tasks. In matching, they locate the matching stimulus and choose it. In mismatching, they locate the matching stimulus and avoid it. Thus, the matching relation between stimuli determines how pigeons learn both of these conditional discriminations and thus, the sameness relation is important for the pigeon.

Prospective Coding

In Pavlovian conditioning animals are able to anticipate the arrival of biologically important events (e.g., food or shock). Humans, however, have the ability to anticipate the arrival of events and use those anticipations as the basis for making

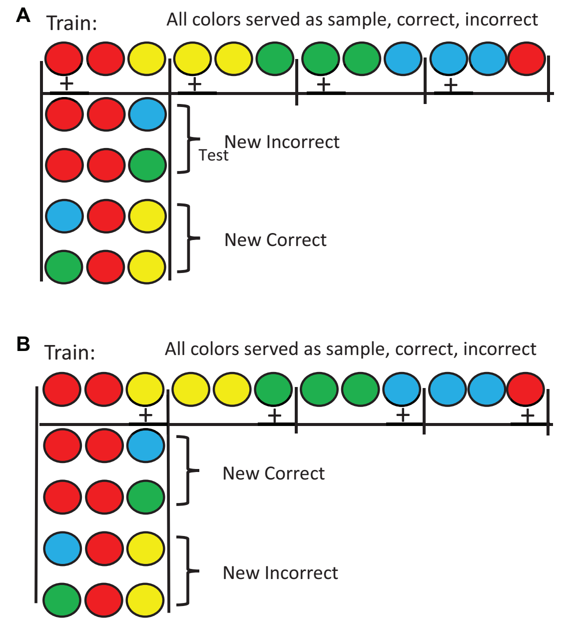


FIGURE 1 | (A) Half of the pigeons were trained on matching with the stimuli on the top line (counterbalanced for position correct). Pigeons pecked the center of the three stimuli 10 times to produce the two comparison stimuli. A single peck to either comparision stimulus terminated the trial. Reinforcement is indicated by a +. Testing was done with New Incorrect stimuli and with New Correct stimuli (as shown). The figure shows the red sample test trials. There were also similar test trials with the other three colors (not shown). **(B)** The remaining pigeons were trained on mismatching with the stimuli on the top line (counterbalanced for position correct). Testing was done with New Incorrect stimuli and with New Correct stimuli (as shown). The figure shows the red sample test trials. There were also similar test trials with the other three colors (not shown). After Zentall et al. (2018).

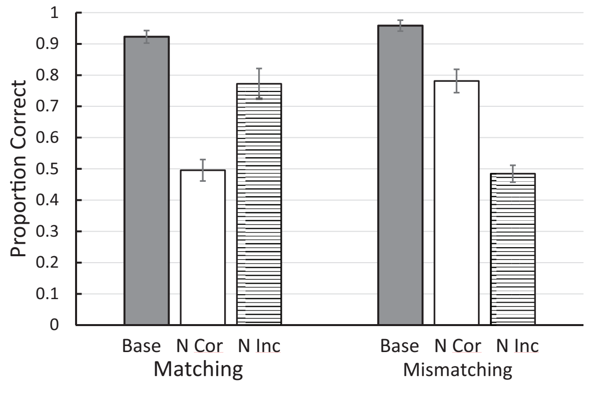


FIGURE 2 | Results of the Zentall et al. (2018) experiment (see **Figures 1**, **2** for the design). Base = baseline matching and mismatching accuracy. N Cor = accuracy on new correct test trials. N Inc = accuracy on new incorrect test trials. Error bars = ± 1 standard error of the mean.

decisions. Humans have the ability to imagine the outcome that they expect to experience. What about other animals?

Trapold (1970) found that in a conditional discrimination, if each sample-correct-comparison chain is followed by a distinctive outcome (e.g., food or water) the anticipation of that outcome can serve as a stimulus to facilitate comparison choice. This phenomenon is known as the differential outcomes effect. Differential outcomes have also been found to improve memory in a delayed matching task. For example, if a delay is inserted between the offset of the sample and the onset of the comparison stimuli, pigeons appear to be able to use the expected outcome as the basis for comparison choice, even when the sample itself is forgotten (Peterson, 1984).

Even more impressive, one can train pigeons on two matching tasks with similar differential outcomes on each (e.g., in each discrimination corn follows correct choice of one comparison, while wheat follows correct choice of the other; see design in **Figure 3**). If on transfer tests, the sample stimuli are exchanged between the two tasks, it can be shown that outcome associations provide the sole basis for choice of the comparison stimulus (see e.g., Edwards et al., 1982).

Further evidence for anticipatory memory comes from research with the radial maze. In the radial maze, the animal is placed on a central platform and there is food in each of 8, 12, or more arms of the maze. Rats should be motivated to enter each arm once to eat the food there and not repeat arm entries and they generally do so. In fact, to produce errors one must insert a delay at some point in the trial. But how do they keep track of the arms with few repeat entries (errors) as they proceed through the trial? There is evidence that rats start by remembering the arms already entered, but once they have entered half of the arms, they switch to anticipate the arms not yet taken (Cook et al., 1985). If the rats were remembering only the arms already taken, one would expect the probability of an error to increase as a function of the number of arms already taken because of the increasing memory load. Although the probability of an error does initially

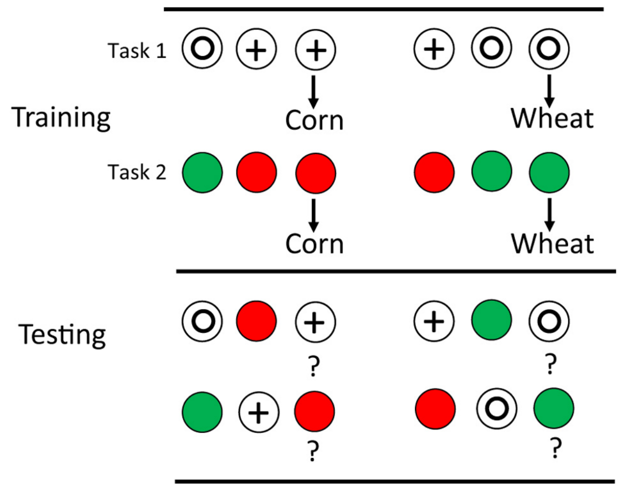


FIGURE 3 | Pigeons were trained on two matching tasks: red green and circle plus with differential outcomes (corn for one trial type, wheat for the other). They were then tested with the samples from one task and the comparisons from the other task. Significant transfer indicated that outcome anticipation based on the samples could serve as discriminative stimuli for comparison choice (after Edwards et al., 1982).

increase as more arms are visited, it then decreases as the number of arms not yet visited decreases. These data demonstrate that the rats use an efficient strategy for visiting the arms by minimizing the memory load as they proceed through the trial.

Interestingly, pigeons show a similar effect in an operant analog of the radial maze involving pecking keys on a panel (Zentall et al., 1990). In this task, the pigeons must peck each key for reinforcement but on any trial, reinforcement is not provided if the pigeon returns to an already pecked key. Once again, if a delay is inserted either early in the trial or late in the trial the error rate is quite low, but if the delay is inserted toward the middle of the trial the error rate is considerably greater.

Acquired Equivalence

In an operant box, pigeons can learn a conditional discrimination in which there is an arbitrary relation between the sample and the correct comparison stimulus. When two sample stimuli (e.g., a red light and a vertical line) are each associated with a common comparison stimulus (e.g., a circle), there is a many-to-one mapping of samples on to the same comparison stimulus. Under these conditions one can ask if an equivalence relation develops between the two samples. That is, do those two samples come to *mean* the same thing.

There are several ways to test for equivalence. In one design, the red light is now associated with a new comparison stimulus (e.g., a blue light). The test for equivalence is to ask if, without further training, the vertical line is also associated with the blue light, in spite of the fact that the vertical line and the blue light had never been presented together before (see design in **Figure 4**). Using this design, Urcuioli et al. (1989) found that pigeons showed positive transfer to those stimuli never presented together before. This finding suggests that for the pigeon, the red light and the vertical line have come to be similarly represented.

Later research attempted to determine the nature of the common representation by inserting a variable duration delay between the offset of the sample and the onset of the comparison stimuli (Friedrich et al., 2004). This research took advantage of the fact that earlier research had found that colored samples were remembered better than line orientations. That is, the forgetting function for line orientation samples was steeper than for colored samples. When a red light and a vertical line were both associated with the same comparison stimulus, however, the slopes of the resulting retention functions suggested that the two samples were commonly represented during the delay. Furthermore, other research suggested that the representation was likely the sample that was easiest to remember (Zentall et al., 1995). So presumably, the pigeons represented the vertical line sample as a red sample, a stimulus that earlier research had indicated was easier to remember.

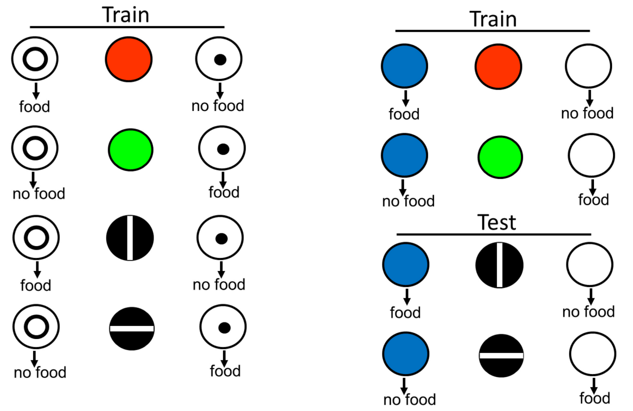


FIGURE 4 | Pigeons were originally trained to choose the circle when the sample was red or a vertical line, and to choose the dot when the sample was green or a horizontal line. To determine if red and vertical were commonly coded and green and horizontal were commonly coded, the pigeons were then trained to choose blue when the sample was red and to choose white when the sample was green. They were then tested with vertical and horizontal line samples and blue and white comparison stimuli. Evidence for functional stimulus equivalence was choice of blue when the sample was a vertical line and white when the sample was a horizontal line (after Urcuioli et al., 1989).

Directed Forgetting

When humans are shown a list of words and are told that they will have to remember some of them but not others, they don't remember as well the words they were told they could forget, as the words they were told they would have to remember (see e.g., Golding and MacLeod, 1998). The implication of this finding is that there is an active rehearsal process triggered by the instruction to remember and the rehearsal process is not triggered by the instruction to forget.

It is often assumed that animals do not have active control over their memory. It is thought that events are remembered and forgotten automatically as a function of the passage of time or intervening events. The challenge in assessing directed forgetting in animals is how to give them *instructions* to remember or forget.

The first presumed evidence of directed forgetting in pigeons was reported by Maki (1981), who used a delayed matching task. Once pigeons had learned to match with delays, on some trials, a stimulus was presented during the delay and on those trials, the comparison stimuli were omitted. Thus, one can think of the delay stimulus as a cue to forget because on those trials, there would not be a test of memory for the sample. As with humans, the test of directed forgetting occurred when, on infrequent probe trials with the forget cue, comparison stimuli *were* presented. In several experiments, pigeons performed very poorly on those probe trials, suggesting that their memory was impaired. Thus, the results suggested that memory for the sample was not automatic.

An important problem with that design, however, is that the forget cue signaled not only the absence of a comparison stimulus test, but also the absence of the possibility of reinforcement on that trial. Thus, because of its association with the absence of reinforcement, the forget cue likely became an aversive stimulus, with all of the accompanying inhibitory affects associated with such a stimulus.

There are several ways to avoid that problem. For example, Roper and Zentall (1994) trained pigeons on a delayed matching task with red and green stimuli and when inserting a forget cue in the delay, followed the forget cue with a simultaneous discrimination involving stimuli different from the matching task (e.g., vertical and horizontal line orientation stimuli in which the vertical lines were always correct). Thus, the forget cue still signaled that the sample could be forgotten but it also indicated that reinforcement (in the form of the simple simultaneous discrimination) would follow. Then, on probe trials, once again, the forget cue was presented followed by the comparison stimuli from the matching task. Results indicated, however, that with this procedure there was little evidence of directed forgetting. That is, the pigeons matched with no loss of accuracy on the probe trials.

Roper et al. (1995) reasoned that perhaps when the forget cue signaled that a simple simultaneous discrimination would follow, the *instruction* to forget the sample may have been ineffective because the memory load was insufficient to produce forgetting. In human directed forgetting research, being told to forget a word allows the subject time to rehearse other words that they were told to remember. Using this idea, Roper et al. created an analogous task for pigeons in which the forget cue(s) actually served as the sample for another matching task (see **Figure 5**). With this procedure, the presence of the forget cue instructed the pigeon to forget the sample but remember the forget cue because memory for the forget cue would be required for reinforcement. Thus, the appearance of the forget cue should cause the pigeon to *reallocate* its memory from the sample to the forget cue itself. On probe trials, in which the forget cue was followed by the comparisons appropriate to the sample, the pigeons showed significant forgetting of the sample. Thus, the pigeons showed significant directed forgetting, evidence that under appropriate conditions, they have at least some direct control over what they remember.

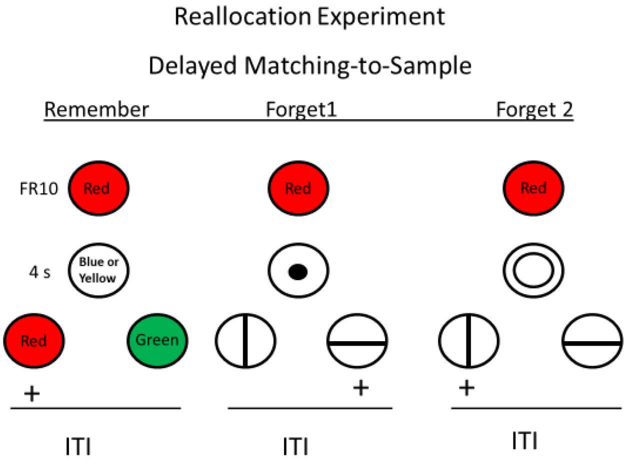


FIGURE 5 | Directed forgetting training with pigeons on red sample trials (green sample trials are not shown). Blue or yellow stimulus presented during the delay signaled sample memory would be tested (remember cue). Vertical or horizontal line presented during the delay signaled signaled sample memory would not be tested (forget cue) but memory for the fortget cue would be tested. Probe trials involved a sample, followed by a forget cue, followed by test (red and green comparison stimuli). Pigeons were significantly less accurate on probe trials than on remember trials (after Roper et al., 1995).

Factors Affecting Oddity Learning

In a mismatching task, reinforcement is provided for choice of the comparison stimulus that does not match the sample. A related task is oddity, in which three stimuli are presented and reinforcement is provided for choice of the stimulus that does not match the other two. The two tasks differ in important ways. In mismatching the sample always appears on the center key, and the pigeons must peck the sample several times before the two comparison stimuli are presented. In oddity, there is no sample (thus no sample pecking) and the odd stimulus can appear on any of the three response keys.

Zentall et al. (1974) compared the pigeon's acquisition of a two-color mismatching task, with and without required responding to the sample, and oddity in which the odd stimulus could appear on the center key. They found that mismatching was acquired quickly with sample responding required but only slowly without responding to the sample. In the same experiment, they also found little learning of the oddity task. Correct responding by chance on the three-key oddity task is 33% correct and the pigeons generally improved to 50% correct by developing a color preference. Yet, after that, they showed little evidence of learning to select the odd stimulus.

Zentall et al. (1980a) asked if increasing the number of matching stimuli from two to four would affect pigeons' acquisition of the oddity task. Although increasing the number of matching stimuli decreased the probability being correct by chance to 20%, surprisingly, they found that when the odd stimulus was part of a five-stimulus array, the pigeons acquired the task rapidly. It appears that with four matching stimuli, the odd stimulus stood out better from the "background" of matching stimuli.

Zentall et al. (1980b) tested this hypothesis further, using an array of 25 stimuli, with 24 matching stimuli and one odd stimulus. Although the probability of choosing the correct location by chance was now only 4% and of choosing the correct color by chance was only 50%, the pigeons learned this task very quickly. When such a phenomenon has been reported in humans it has been referred to as *visual pop out* (Treisman, 1985). Although this phenomenon might be considered perceptual rather than cognitive, it is another example of a similarity between humans and other animals. Simple learning theory would predict that with 25 possible response locations, the oddity task would be harder than with only three locations—certainly it should not be any easier.

Timing

In our modern culture, time plays a very important role. Our ability to keep track of the passage of time, however, is not very good. To aid us, we use watches, clocks, and smart phones. When we were hunter gatherers and until quite recently, external cues like where the sun was in the sky, day/night cycles, and the phases moon, were sufficient because short time intervals were likely not very important. What about other animals? To what extent are they able to discriminate the passage of time?

One measure of short-interval animal timing is the performance of an animal on a fixed interval schedule. For example, if a pigeon receives a reinforcer for the first response after 1 min, with adequate training, one typically sees what has been called a fixed interval *scallop*. Responding does not start immediately after the last reinforcer but then increases, first slowly and then faster, as the time since the last reinforcer approaches 1 min. To get a better measure of the animal's timing ability one can start the fixed interval trial with the onset of a stimulus and turn it off with reinforcement. After some training, one can introduce *empty* trials in which reinforcement is omitted but the stimulus stays on. If one averages the pecking over a series of such empty trials, the plot of response rate as a function of time since the start of the trial has a peak very close to the time that the reinforcer would have occurred on a fixed interval trial.

Another measure of short-interval animal timing involves the use a temporal discrimination. For example, for pigeons, after having experienced a short interval sample (e.g., 2 s), choice of the red comparison would be reinforced, whereas having experienced a longer sample (e.g., 8 s) choice of the green comparison would be reinforced. After sufficient training, to get an idea of the underlying scale of time for the pigeon, one can present the pigeon with sample durations between 2 and 8 s. The psychophysical function that results when plotting the probability of a *long* response, as a function of the sample duration, provides a measure of the animal's scale of timing. In particular, the sample duration to which the animal distributes its responses equally between *short* and *long* is referred to as the *point of subjective equality*. Although one might expect that point to be the arithmetic mean of the two training durations (in this case 5 s), it is typically closer to the geometric mean (4 s), suggesting that the pigeons' judgment of the passage of time is not linear but is logarithmic. In the example given, the geometric mean is at 4 s because the ratio of 2–4 is the same as the ratio of 4–8. Similar psychophysical function have been found for humans.

Do Animals Represent Time Categorically?

When humans are given a temporal discrimination like the one described for pigeons above, they are very likely to describe the intervals relationally, as *short* and *long*, rather than in terms of their absolute duration (2 and 8 s). We were interested in whether pigeons also represent intervals relationally as *short* and *long* (Zentall et al., 2004). To answer this question, we trained pigeons on two temporal discriminations, one involving 2 and 8 s samples (with red and green comparison stimuli), and the other involving 4 and 16 s samples (with vertical and horizontal stripes). Note that the 4 s sample falls at the geometric mean of the 2–8 s discrimination, and the 8 s sample falls at the geometric mean of the 4–16 s discrimination (see **Figure 6**). On probe trials, we presented the 4 s sample with the comparisons from the 2 to 8 s discrimination and the 8 s sample with the comparisons from the 4 to 16 s discrimination.

Normally, presenting durations that correspond to the geometric mean should result equal choice of *long* and *short*. If the pigeons represented the 4 s sample as *short*, however, they might be expected to choose the colored comparison associated with the short, 2 s sample. And if the pigeons represented the 8 s sample as *long*, they might be expected to choose the line comparison associated with the long, 16 s sample. In fact, such a bias was found. Thus, similar to humans, pigeons show some evidence of representing time intervals relationally.

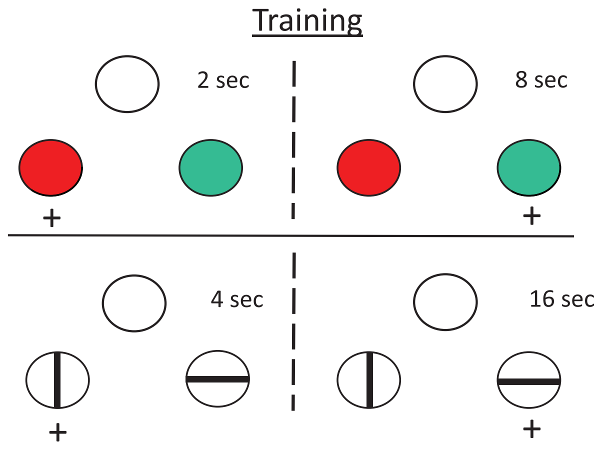


FIGURE 6 | Relative timing experiment (Zentall et al., 2004). Pigeons were trained to discriminate 2 s samples from 8 s sample (top) and 4 s samples from 16 s samples (bottom). They were then tested with 4 and 8 s samples and the comparison stimuli from the other discrimination. Although those sample durations were at the geometric mean of the other discrimination, the pigeons tended to choose "short" (red) when the sample was 4 s and choose "long" (horizontal) when the sample was 8 s.

Is Subjective Time Affected by What the Animal Is Doing?

As noted earlier, we humans are not very good at estimating the passage of time. For example. when taking an exam, students are often surprised at how much time has elapsed since the start of the exam (time flies when one is cognitively involved). On the other hand, if students are attending a boring lecture, time seems to pass very slowly. Do animals experience a similar effect? Does time pass by faster when pigeons are behaviorally involved than when they are not?

To test this possibility, Zentall and Singer (2008) trained pigeons on a temporal discrimination involving 2 and 10 s samples. When the samples were white, the pigeons were required to refrain from pecking, but when the samples were blue, the pigeons were required to peck them at least once per sec (see **Figure 7**). On test trials, white and blue samples were presented for durations between 2 and 10 s. The question of interest was the effect that sample pecking (and the absence of pecking) had on the psychophysical function (relating choice of long to sample duration), and specifically on the point of subjective equality (see **Figure 8**). Relative to a group of pigeons that were free to peck or not, they found that when the pigeons were required to peck the temporal samples, the point of subjective equality shifted to longer durations. That is, the pigeons judged that less time had elapsed. Whereas, when the pigeons were required to refrain from pecking the temporal samples, the point of subjective equality shifted to shorter durations. That is, the pigeons judged that more time had elapsed. These results indicate that animals judge the passage of time with biases similar to those of humans. The

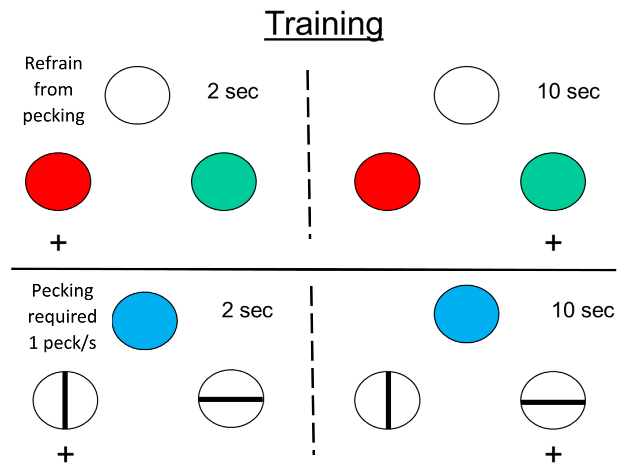


FIGURE 7 | Pigeons were trained on two temporal discriminations involving 2 and 10 s samples. When the sample was white, the pigeons were required to refrain from pecking it. When the sample was blue, the pigeons were required to peck it (once per s). On test trials, when durations between 2 and 10 s were presented. The pigeons tended to treat the white sample durations as longer than the blue sample durations (after Zentall and Singer, 2008).

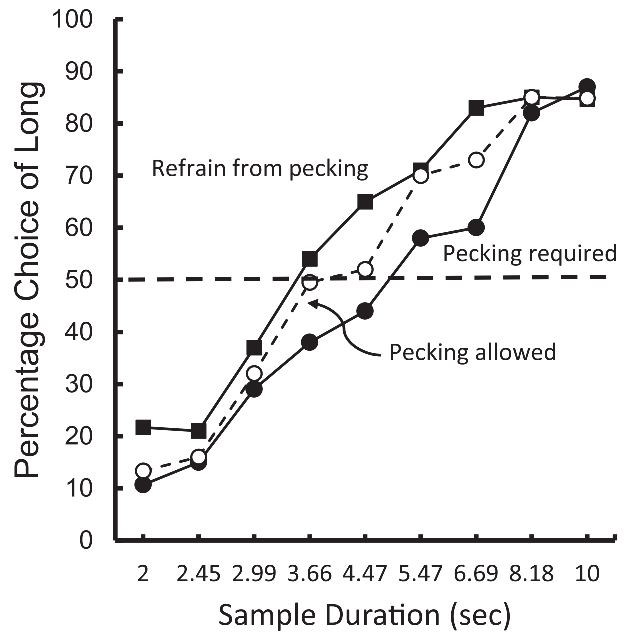


FIGURE 8 | Choice of "long" as a function of sample duration. Pigeons were trained to discriminate between samples of 2 and 10 s. When the sample was white the pigeons were required to refrain from pecking. When the sample was blue the pigeons were required to peck at a rate of 1 peck per s. On test trials, the sample duration was varied between 2 and 10 s. For comparison purposes data are also presented from pigeons for which pecking was allowed. After Zentall and Singer (2008).

implication of this research is that, much like humans, pigeons appear to judge the passage of time in terms of the rate at which relevant events occur.

Counting

The ability to count or to use the number of objects or events as a cue is a quality that adult humans perform routinely and efficiently. However, the degree to which non-verbal organisms have this ability is more controversial. Although relative numerosity judgments have been studied extensively in animals (see e.g., Meck and Church, 1983), absolute numerosity judgments are quite a bit more difficult. After extensive training, Xia et al. (2000) had some success in training pigeons to respond a fixed number of times, defined by the specific Arabic numeral displayed.

In keeping with MacPhail's suggestion that it is important to find the context appropriate to the animal, Seligman and Meyer (1970) found that after rats had been trained to press a lever for occasional delivery of food, in each session they were delivered exactly three shocks, randomly spaced throughout the session. The introduction of shocks produced suppression in responding; however, once the rats had had some experience with this procedure, they began responding at a higher rate after the third (last) shock had been administered. Thus, they understood when shocks would no longer occur. Similarly, Capaldi and Miller (1988) found that rats trained on a series of four straight-alley runs, in which a reinforcer was found on the first three runs but not on the fourth, ran slower on the fourth run. In both cases rather than asking the animal to count the number of responses they made, these studies had the animals count the number of biologically meaningful events (food). Furthermore, rather than use a discrete measure of counting, they used a continuous measure, response rate or running speed.

We took a similar approach and asked if pigeons could learn that they would be fed after each of the first three 10-peck sequences in a trial, but not after the fourth (Rayburn-Reeves et al., 2010). We used the time to complete each 10-peck sequence as a measure of their counting ability and found that the pigeons completed the 10-peck requirement relatively quickly for each of the first three sequences (about 5.5 s per sequence) but they took almost twice as long to complete the fourth sequence. When one is assessing an animal's ability to keep track of successively experienced events it is important to control for the time it takes to experience the events because the animals may be judging the passage of time instead of the number of events that it experienced. To control for time between the start of a trial and the fourth sequence (timing rather than counting) we started random trials with a non-reinforced 10-peck sequence. Thus, on those trials, the third reinforced sequence actually occurred at a time when non-reinforcement would have occurred on a standard four-sequence trial. The large increase in latency to complete fifth sequence suggested that the pigeons were counting the number of reinforcers rather than the time from the start of the trial. This line of research demonstrates that tapping into the appropriate motivational system may be critical in assessing the cognitive ability of an animal.

Imitation

It is well known that some animal are behavioral copiers. We even acknowledge copying by animals with expressions such as to "ape" someone and "monkey see monkey do." But what psychological processes are involved in the copying of behavior? Piaget (1962) suggested that true imitation involves being able to take the perspective of another. That is, something like "if I put myself in his place, what would I have to do to get the outcome that he is getting." It is difficult to imagine a young child who is imitating an adult reasoning of that kind and certainly not a non-verbal animal. But the question of whether animals are capable of imitating the behavior of a conspecific was of interest to us.

One could start by asking if an animal could learn a response after seeing another animal perform that response. But of course, one should ask, relative to what. Relative to an animal learning on its own by trial and error, perhaps. But the mere presence of another animal might facilitate learning (an effect known as social facilitation). Furthermore, if the imitation involved the manipulation of an object, the sight of that manipulation could attract the observer to that object (a phenomenon known as stimulus enhancement). Finally, facilitated acquisition could be attributed to what developmental psychologists call learned affordances, learning how the environment works, independent of the action that led to the result (e.g., learning that the up and down movement of a lever leads to the appearance of food). The question is how to test for *true imitation* while controlling for these other presumably less cognitive mechanisms.

Zentall et al. (1996) used a method they referred to as the two-action procedure to control for those non-imitative processes. They trained demonstrator pigeons to obtain food, either by pecking at a treadle (a flat metal plate located near the floor of the chanber), or by stepping on the treadle. Then they allowed observer pigeons to observe one of those behaviors (or the other). Finally, they allowed the observers to operate the treadle with either response. Zentall et al. found that the observers showed a significant tendency to operate the treadle in the same manner that they had observed it performed by the demonstrators. Using this procedure, an even stronger imitative response was found in Japanese quail, a species known to demonstrate imprinting (Akins and Zentall, 1996). The beauty of the two-action procedure is that it controls for social facilitation, stimulus enhancement, and learned affordances. That is, each group serves as a control for the other, the only difference being the manner in which the treadle was operated by the demonstrator, with its foot or with its beak.

Further research on imitation found that observers would not imitate if the demonstrator did not receive a reinforcer for their treadle response (Akins and Zentall, 1998). Nor would the observer imitate if, at the time of observation, it was not motivated by the reinforcer obtained by the demonstrator (i.e., if the observer had been prefed; Dorrance and Zentall, 2001).

Another interesting distinction related to imitation was suggested by Bandura (1969). In describing imitation by children, he distinguished between imitation and observational learning. Bandura claimed that imitation that occurred at the time of observation could be *reflexive* and was perhaps genetically predisposed (copying behavior sometimes referred to as response enhancement), whereas *observational learning* represented the internalization of the observed response, such that it could be performed at a later time.

Although in the research described above the observation and observer's performance did not occur at the same time, not much time elapsed between the two. However, as part of a larger study (Dorrance and Zentall, 2001), observers that were tested 30 min following observation showed significant copying of the stepping or pecking behavior that they had earlier observed. Thus, according to Bandura, such copying should qualify as observational learning, a more cognitive behavior than "simple" imitation.

COGNITIVE BIASES

Certain human behaviors would be described as biased or even suboptimal because they appear to be inconsistent with basic principles of associative learning. Although these behaviors do not represent an accurate assessment of the contingencies of reinforcement, they are thought to result from the cognitive *misunderstanding* of the context. We have studied four of these in pigeons: Justification of effort (a version of cognitive dissonance), base rate neglect, unskilled gambling behavior, and sunk cost (the tendency to persist in a task based on past investment, rather than the future contingencies of reinforcement).

Justification of Effort

When humans behave in ways that are inconsistent with their beliefs it is thought to create *cognitive dissonance*. This dissonance may be a social phenomenon resulting from an attempt to avoid being considered a hypocrite. Do animals have beliefs? If so, are they concerned about the consistency between their beliefs and their behavior? How would one go about evaluating their beliefs to determine whether they are consistent with their behavior? And how would one measure the presumed dissonance that might result from that inconsistency?

One version of cognitive dissonance, called *justification of effort*, may provide a tractable approach to study this behavior in animals. *Justification of effort* is the tendency to prefer reinforcers that one has worked harder to obtain. If the reinforcers are of equal value, a preference should not be found. In fact, one might expect that if one had to work hard for a reinforcer, it might reduce the value of the reinforcer and thus, it should not be preferred. If, however, there is a tendency to justify the effort put into obtaining the reinforcer, the reinforcer might be preferred. Cognitive dissonance theory would suggest that if one had to work harder to obtain the reinforcer, it must have more value. If not, the theory suggests, why did one work so hard to obtain it.

To study justification of effort in pigeons Clement et al. (2000) trained them to peck a white light. On half of the trials, the pigeon was required to peck the white light once, and then it changed to red. A single peck to the red light was reinforced. On the remaining trials, the pigeon was required to peck the white light 20 times, and then it changed to green. A single peck to the green light produced the same reinforcer. The purpose of the red and green lights was to have a way to distinguish between the two conditions of reinforcement because the reinforcers were exactly the same.

When Clement et al. (2000) tested the pigeons by giving them a choice between the red and green lights, the pigeons showed a significant preference for the green light, the color that they had to work harder to obtain (see also Friedrich and Zentall, 2004). Furthermore, they did so independently of the number of times they had to peck the white light on the test trial.

Other research indicated that other relatively less preferred events had the same effect. For example, pigeons generally prefer immediate over delayed reinforcement. However, they preferred a stimulus that followed a delay over one that did not (DiGian et al., 2004). Similarly, pigeons prefer food over the absence of food, yet they preferred a stimulus that followed the absence of food, over a stimulus that followed food (Friedrich et al., 2005).

Although this procedure fits the design of a justification of effort experiment, given that it was conducted with pigeons, one might not be inclined to interpret the results in terms of an inconsistency between the pigeon's belief (that fewer pecks are better than more pecks) and its behavior (having pecked many more times to obtain the same reinforcer). Instead, one is likely to consider the preference in terms of a more behavioral mechanism. A likely alternative mechanism is positive contrast. In the case of 1 vs. 20 pecks, it would be the contrast between the effort expended in responding to the white stimulus and the appearance of the signal for reinforcement. One can think of the contrast in terms of frustration that occurs on the high effort trials, that is relieved upon the appearance of the stimulus signaling reinforcement.

If this effect occurs in pigeons, could positive contrast also account for examples of justification of effort in experiments with human subjects? If so, one might expect humans to show a similar effect when trained on the task we used with pigeons. Alessandri et al. (2008) tested this procedure with children using computer mouse clicks as the response requirement and found similar results, a preference for the stimulus that followed greater effort. Furthermore, Klein et al. (2005) used a similar procedure with human adults and found the same result. Interestingly, when the adult subjects were asked why they had chosen the stimulus that they had worked harder to obtain, most of them said that they did not know. This finding suggests that the bias to prefer the stimulus that the subjects had worked harder for was learned implicitly (unconsciously), and it suggests that it is likely to be a mechanism similar to that used by pigeons.

Base-Rate Neglect

Base-rate neglect is a bias that fails to sufficiently include the original rate of a probable event. For example, let's say that given certain symptoms, the probability of having the flu is 0.10. Let's say, as well, that there is a test that is 80% accurate in diagnosing the flu. The test gives false positives 20% of the time and false negatives 20% of the time. Let's say a patient tests positive, what is the likelihood that they have the flu? Many people would say 80% (or maybe somewhat less). In fact, the probability of having the flu is much less (see **Figure 9**). Given the probabilities described, it is the probability of an accurate diagnosis of the flu (0.08), divided by the sum of the accurate diagnosis of the flu (0.08) plus the probability of a false positive (an inaccurate diagnosis that it is the flu, 0.18). Thus, given a positive test, the probability that the patient actually has the flu is better than without the test (0.10), but it is actually only.31!

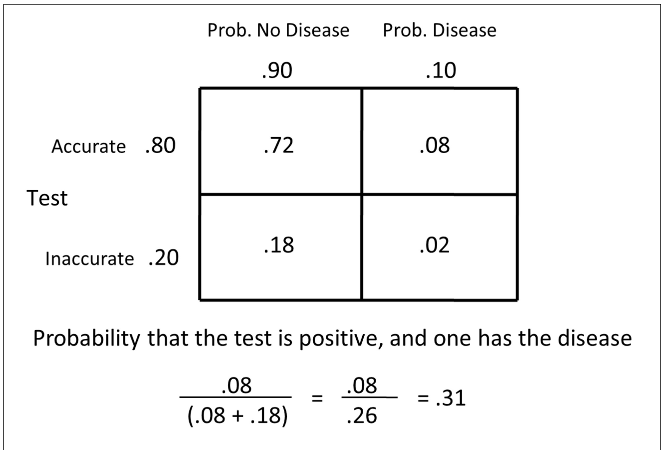


FIGURE 9 | Matrix representing the probability of having and not having the disease by the accuracy of the test. The probability of having the disease equals the probability of having the disease, given the test is accurate, divided by the sum of the probabilities of having the disease, given the test is accurate, plus the probability of not having the disease, given that the test is inaccurate.

How would one create a simulation of base-rate neglect for animals? The idea would be to create a condition in which, similar to the flu test, errors can occur. One such task might be delayed matching to sample involving, for example, red and green comparison stimuli equally associated with reinforcement (Zentall et al., 2008; see **Figure 10**). This would be considered the base rate because in the absence of the sample (the flu test) the red and green comparison stimuli should be equally associated with reinforcement. However, the samples (representing the flu test) would not be equally presented. On one third of the trials, the sample stimulus is green, and choice of the green comparison stimulus is reinforced. On another third of the trials, the sample stimulus is red, but correct choice of the red comparison stimulus is reinforced only 50% of the time. On the final third of the trials, the sample stimulus is yellow and correct choice of the red comparison stimulus is reinforced only 50% of the time. Thus on 1/3 of the trials choice of the green comparison is reinforced and on 2/3 of trials choice of the red comparison stimulus is reinforced but only 50% of the time, thus equally often as the green comparison.

Thus, when the sample stimulus (the flu test) is available, reinforcement can be obtained 67% of the time (all of the 33% of the trials with green samples and half of the 67% of the trials with red or yellow trials). But what should the pigeon do in the absence of memory for the sample? Which comparison should the pigeon choose? In the absence of a sample, because each of the comparison stimuli would be associated with 33% reinforcement, there should be no bias. This is the base rate. In training, however, the red comparison would have been chosen twice as often as the green comparison. Thus, in spite of the fact that correct choice of the red comparison was reinforced only 33% of the

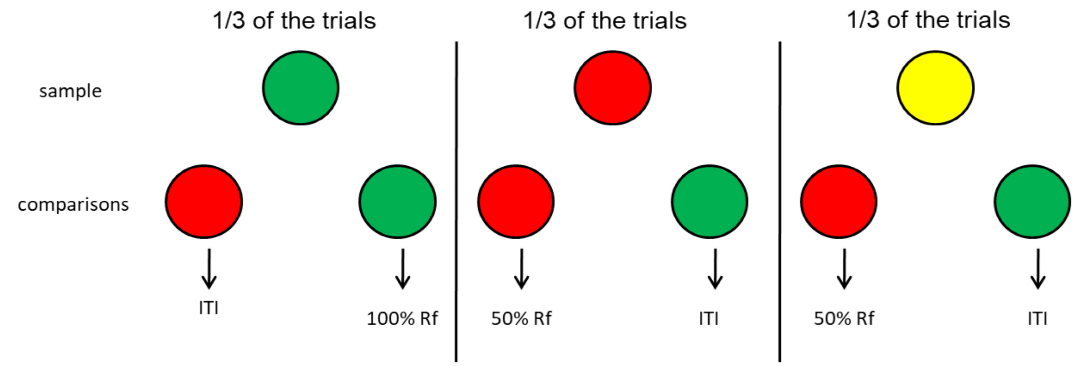


FIGURE 10 | Pigeon base rate neglect experiment (Zentall et al., 2008). Pigeons had an equal probability of getting a red, green, or yellow sample. However, the green comparison was correct when the sample was green and correct choice was reinforced 100% of the time, and the red comparison was correct when the sample was red or yellow but correct choice was reinforced only 50% of the time. Thus, in the absence of memory for the sample, the probability of reinforcement associated with each of the comparisons (the base rate) was 33% for both comparison stimuli. Yet, with increasing delay between the offset of the sample and the onset of the comparison stimuli, the pigeons showed a strong preference for the red comparison stimulus.

time, when there was a delay between the offset of the sample and onset of the comparison stimuli, the pigeons showed a strong preference for the red comparison stimulus. That is, they showed clear evidence of base-rate neglect. Similar findings were reported by Zentall and Clement (2002) and DiGian and Zentall (2007) using somewhat different designs. Thus, in these experiments, the pigeons were unduly influenced by the frequency with which they had responded to the two comparison stimuli (the equivalent of the accuracy of the flu test in the human example).

Unskilled Gambling

When humans are engaged in unskilled gambling (e.g., slot machines, lotteries, roulette) their choice is almost always suboptimal (their investment is almost always greater than the return). Those who engage in such activities claim that they do it because gambling is entertaining. Thus, generally one would not expect non-human animals to engage in such behavior. If animals are hungry and working for food, entertainment should not be a factor. Furthermore, optimal foraging theory (Stephens and Krebs, 1986) proposes that animals have evolved to forage for food in the most effective way because more efficient foragers would survive and reproduce better.

Research with pigeons, however, suggests otherwise. For example, Spetch et al. (1990) found that some pigeons preferred an alternative that provided a signal for 50% reinforcement (the gamble) over an alternative that provided a signal for 100% reinforcement (the non-gamble). In another experiment, involving manipulation of the magnitude of reinforcement, perhaps a better analog of human gambling behavior, Zentall and Stagner (2011) gave pigeons a choice between two alternatives. Choice of one alternative, 20% of the time, gave the pigeon a green light signaling that it would receive a "jackpot" of 10 pellets of food, but 80% of the time it would receive a red light signaling that it would get no food. Choice of the other alternative, 100% of the time, gave the pigeon a light signaling that it would get 3 pellets (the non-gambling option; see **Figure 11**). In this experiment, the pigeons showed a very strong preference for the gambling option that provided the pigeons with an average of 2 pellets per trial, over the non-gambling option that provided the pigeons with 3 pellets per trial (see also Stagner and Zentall, 2010).

These results suggest that the pigeons were not choosing between the values of the alternatives at the time of choice but between the value of the *signals for reinforcement* that followed that choice (Smith and Zentall, 2016). Thus, they appeared to be choosing, not between an average of 2 pellets vs. 3 pellets but between the occasional 10 pellets and 3 pellets. Interestingly, problem gamblers show a similar bias. When the value of a lottery is described in the media, the amount of a winning ticket is announced but rarely is one privy to the very low probability of such a win.

The theory that is the value of the signal for reinforcement that determines the value of the choice suggests that the signal for non-reinforcement, the one that occurred on 80% of the choices of the suboptimal alternative, has little inhibitory effect on choice. This hypothesis was confirmed by Laude et al. (2014), who found

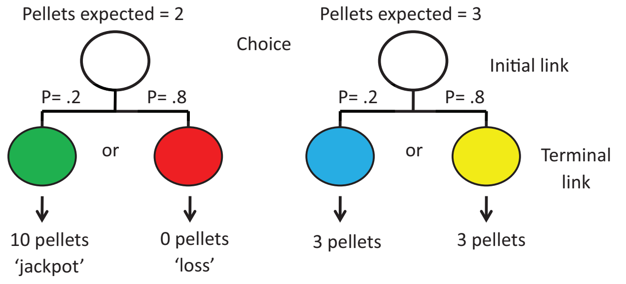


FIGURE 11 | Pigeons were given a choice between the left side and the right side. If they chose the left side there was a 20% chance that they would receive a green light followed by 10 pellets of food but an 80% chance that they would receive a red light followed by no food. If they chose the right side, whether they received a blue light or a yellow light they always received 3 pellets of food. Thus, although they received 50% more food if they chose the right side, they showed a strong preference for the left side. After Zentall and Stagner (2011).

virtually no inhibition to the stimulus associated with the absence of reinforcement.

The hypothesis that it is the value of the signal for reinforcement that determine choice, was further tested by Case and Zentall (2018). They gave pigeons a choice between 50% signaled reinforcement and 100% signaled reinforcement. The hypothesis proposed by Smith and Zentall (2016) that it is the value of the signal for reinforcement that determines choice suggests that pigeons should be indifferent between the two alternatives, because although the choice of the 50% reinforcement alternative provided only half as much reinforcement as the 100% reinforcement alternative, both of the signals for reinforcement were associated with the same 100% reinforcement. Surprisingly, after a large amount of training on this task, the pigeons actually developed a significant preference for the suboptimal, 50% reinforcement, alternative!

Thus, it appears that there is a second mechanism involved in pigeons' suboptimal choice. Case and Zentall (2018) proposed that positive contrast between the expected probability of reinforcement at the time of choice and the probability of reinforcement signaled by the conditioned stimulus (when it occurred) was responsible for the suboptimal preference.

To test the hypothesis that positive contrast is responsible for the preference for 50% reinforcement alternative over the 100% reinforcement alternative, Zentall et al. (2019) reduced the presumed amount of contrast, by *increasing* the probability of the signal for reinforcement associated with the suboptimal alternative from 50 to 75%. Thus, instead of a change in the probability of reinforcement from 50% at the time of choice to 100% upon the appearance of the conditioned stimulus (a difference of 50%) there was only a 25% difference. In spite of fact that the suboptimal alternative was now associated with *more* reinforcement (75%), there was a significant *reduction* in the preference for the suboptimal alternative. Thus, positive contrast between what is expected and what occurs appears to make an important contribution to the choice of the suboptimal alternative.

The positive contrast that pigeons show when the value of the conditioned reinforcer exceeds the expected value of choice may help to explain why humans engage in unskilled gambling. The few times that human gamblers win (or perhaps even imagine winning) may provide positive contrast (the feeling that gamblers express of being entertained) similar to that of pigeons.

The Sunk Cost Fallacy

A sunk cost is an expenditure of resources that has already occurred. The sunk cost fallacy occurs when one allows a sunk cost to determine the future investment of resources. According to economic theory, the decision to invest further in a project should depend solely on the future likelihood of its success. However, humans often continue to invest in a losing project to avoid feeling that the project was a failure, but further investment is often likely to produce additional losses. The sunk cost fallacy also may result from the cultural admonition to avoid wasting resources. But in the case of a bad investment, the resources expended are already lost. Behavioral economists often point to the sunk cost fallacy as evidence that humans do not always behave rationally (Arkes and Blumer, 2000). Non-human animals, however, should be sensitive to future reinforcement contingencies and should not be affected by cultural factors like the sunk cost fallacy.

However, several experiments have demonstrated the sunk cost fallacy in animals. For example, Navarro and Fantino (2005) examined the sunk cost effect in pigeons in which, on each trial there was a 50% chance that a small number (10) of pecks would be required for reinforcement and a decreasing probability that many more responses (40, 80, or 160) would be required. At any time, the pigeon could peck a different response key that would start a new trial, thereby potentially getting a trial with a smaller number of pecks to reinforcement. The optimal strategy would be to peck 10 times and, in the absence of reinforcement, start a new trial. Surprisingly, the pigeons in that study generally persisted and rarely choose to start a new trial (see also Magalhães and White, 2014).

This task is similar to the economic sunk cost effect with humans because in both cases there is some uncertainty about the likelihood that persistence will not pay off. For the pigeon, after 10 pecks, it would be best to start a new trial, however, it is possible that persisting will produce food after 40 more pecks, whereas by starting a new trial it could take 80 or even 160 pecks to produce food. Also, starting a new trial required the pigeon to stop pecking, move to the other response key, peck it, and then move back to the original response key.

Pattison et al. (2012) asked if pigeons would show a sunk cost effect even if there was no uncertainty about the results of persisting and no differential cost to switch to the other response key. In one experiment, they first trained pigeons to peck a green key 30 times for food on some trials and peck a red key 10 times for food on other trials. Then they trained the pigeons to peck a green light on a side key a variable number of times to turn off the green light and light a white key in the middle. A single peck to the white key relit the green key and also lit a red key on the other side of the white key. Now from the middle white key, the pigeon could choose to go back and peck the green key enough times to total 30 pecks (the initial investment plus the remaining pecks had to equal 30) to obtain a reinforcer. Or it could switch to the red key for 10 pecks to obtain a reinforcer. The question was, would the pigeon switch to the red key that required 10 pecks for reinforcement when going back to the green key meant it would have to make more than 10 pecks for reinforcement.

The results indicated that when the pigeons had invested as few as 10 pecks to the green key, they preferred to return to the green key for the remaining 20 pecks, rather than switch to the red key for 10 pecks. Under these conditions, at the time of choice, there should have been no uncertainty about the number of remaining pecks and the pigeons were equally distant from the green and red keys. Thus, they were biased to return to the green key, even though it required more pecks. Only when there were no initial pecks to the green key (no prior investment) did the pigeons prefer the red key over the green key. Thus, much like humans, the pigeons preferred to complete a task already started, rather than switch to another task.

The sunk cost fallacy may be related to a human gambling phenomenon, known as *chasing losses*. When gamblers start to lose, they often show a tendency to keep gambling, with the intent to recoup the money that they have lost. As a result, they typically get further into debt.

Uniquely Human Fallacies

There are several fallacies shown by humans that other animals do not appear to show, fallacies that appear to result from human experience or knowledge, or that may be cultural in nature. One of these is the Monte Carlo fallacy. Another is the near miss fallacy.

The Monte Carlo Fallacy

The *Monte Carlo* or *gambler's fallacy* occurs when, over the short-term, a series of events appears to show a bias for one outcome over the other. For example, if one flips a coin 4 times and each time it comes up heads, many people believe that the probability that it will come up heads again is now less than 50%. They believe that it should come up tails, to make up for the unlikely outcome of 4 heads in succession. But the coin tosses are independent—the coin has no memory of its past behavior. So, this is a fallacy.

Most animals live in an often-changing world but one in which purely random, independent events are rare. Thus, their sensitivity to short-term changes in the probability of events is likely to bias them in the direction of those changes, rather than in the reverse direction. That is, if they are exposed to random events in which a particular event occurred 4 times in succession (e.g., reinforcement occurring for a response to the left key of a response panel), they are likely to show a greater tendency to make that same response again, rather than make a response to the other alternative. Pigeons are not likely to *know* that the probability of a coin toss is equally likely to be heads as tails.

The Near Miss Fallacy

The near miss fallacy (more accurately called the near hit fallacy) can most easily be seen in the way that slot machines function. Traditional slot machines have three spinning reels. The player wins if the images on the three wheels match when the reels stop spinning. All other patterns of images on the three reels usually indicate a loss. The interesting case is when the first two images match but the third one does not. This is referred to as the near hit outcome. If the first two reels do not match, it is already a loss, but if the first two reels match, a win is still possible. For this reason, it has been found that, with equal probability of winning, people prefer to play slot machines with a higher rate of near hit outcomes (MacLin et al., 2007).

The slot-machine task easily can be modified for use with pigeons. In several experiments, pigeons were trained on such a task, with two alternatives, equated for wins and losses (Stagner et al., 2015; Fortes et al., 2017). The pigeons could choose between an alternative with near hit trials and another with a different pattern of losses (see **Figure 12**). Unlike humans, however, the pigeons tended to avoid the alternative with the near hit trials. Apparently, for pigeons the similarity of the appearance of near hit losses to win trials actually devalued the effects of a win. That is, to a pigeon it may be that red—red—green appears more similar to red—red—red than to, for example, red—green—red.

It is interesting to speculate about the mechanisms responsible for the difference between humans and pigeons with this task. One possibility is that humans have considerable experience with tasks in which losses, similar to a near hit, represent progress toward a goal.

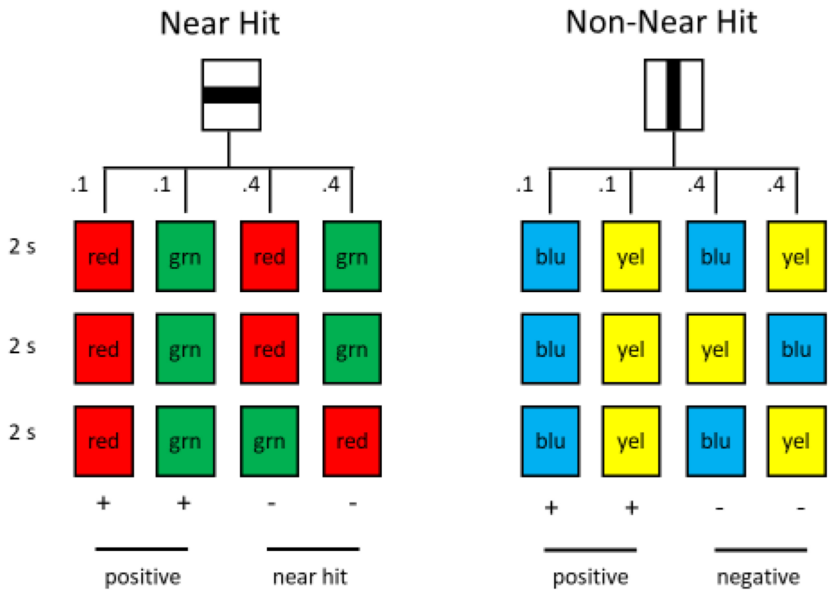


FIGURE 12 | Pigeons had a choice between the horizontal line and the vertical line. Choice of the horizontal line led to (1) the successive presentation of three red or three green stimuli, each with a probability of 0.10 or (2) the successive presentation of two red and one green or two green and one red stimulus, each with a probability of 0.40. Choice of the vertical line led to (1) the successive presentation of three blue or three yellow stimuli, each with a probability of 0.10 or (2) the successive presentation of blue, yellow, blue or yellow, blue, yellow, each with a probability of 0.40 (after Fortes et al., 2017).

Consider trying to get a basketball into a hoop. Initially, one might miss the hoop entirely. With practice one should be able to get closer to putting the ball into the hoop, but still not get it in. That improvement, still involving losses, would be evidence that one is making progress. Thus, the near hit in basketball represents an improvement in one's performance. In the slot machine task in which there is no skill, however, the near hit does not represent progress. To a gambler, however, it may feel like improvement. It would be interesting to know if pigeons that were trained on several tasks in which gradual learning of skill was needed, would also develop a preference for near hit trials when transferred to a slot-machine like task.

CONCLUSION

The research described in the present article, together with a great deal of related research on comparative cognition, suggests that Macphail's hypothesis that all vertebrates have similar cognitive capacities may not be as implausible as it may at first appear. Once one has accounted for important differences in contextual variables concerned with perception, motor skills, and motivation, many of the presumed differences may be more quantitative than qualitative. By its nature, it may not be possible to demonstrate that Macphail's hypothesis is false because one may not ever be able to ensure that the contextual variables are all appropriate for the species in question. However, whatever the outcome of the quest to test Macphail's challenge, I have found it to have resulted in a wealth of informative research on comparative cognition.

AUTHOR CONTRIBUTIONS

TZ prepared the manuscript.

REFERENCES

Akins, C. K., and Zentall, T. R. (1996). Imitative learning in male Japanese quail (*Coturnix japonica*) involving the two-action method. *J. Comp. Psychol.* 110, 316–320. doi: 10.1037/0735-7036.110.3.316

Akins, C. K., and Zentall, T. R. (1998). Imitation in Japanese quail: the role of reinforcement of demonstrator responding. *Psychon. Bull. Rev.* 5, 694–697. doi: 10.3758/bf03208847

Alessandri, J., Darcheville, J.-C., and Zentall, T. R. (2008). Cognitive dissonance in children: justification of effort or contrast? *Psychon. Bull. Rev.* 15, 673–677. doi: 10.3758/pbr.15.3.673

Arkes, H. R., and Blumer, C. (2000). "The psychology of sunk cost," in *Judgement and Decision Making: An Interdisciplinary Reader*, eds T. Connolly, H. R. Arkes, and K. R. Hammond (Cambridge: Cambridge University Press).

Bandura, A. (1969). "Social learning theory of identificatory processes," in *Handbook of Socialization Theory and Research*, ed. D. A. Goslin (Chicago, IL: Rand-McNally), 213–262.

Capaldi, E. J., and Miller, D. J. (1988). Counting in rats: its functional significance and the independent cognitive processes that constitute it. *J. Exp. Psychol. Anim. Behav. Process.* 14, 3–17. doi: 10.1037/0097-7403.14.1.3

Case, J. P., and Zentall, T. R. (2018). Suboptimal choice in pigeons: does the predictive value of the conditioned reinforcer alone determine choice? *Behav. Processes* 157, 320–326. doi: 10.1016/j.beproc.2018.07.018

Clement, T. S., Feltus, J., Kaiser, D. H., and Zentall, T. R. (2000). 'Work ethic' in pigeons: reward value is directly related to the effort or time required to obtain the reward. *Psychon. Bull. Rev.* 7, 100–106. doi: 10.3758/bf03210727

Cook, R. G., Brown, M. F., and Riley, D. A. (1985). Flexible memory processing by rats: use of prospective and retrospective information in a radial maze. *J. Exp. Psychol. Anim. Behav. Process.* 11, 453–469. doi: 10.1037/0097-7403.11.3.453

DiGian, K. A., Friedrich, A. M., and Zentall, T. R. (2004). Discriminative stimuli that follow a delay have added value for pigeons. *Psychon. Bull. Rev.* 11, 889–895. doi: 10.3758/bf03196717

DiGian, K. A., and Zentall, T. R. (2007). Matching-to-sample in pigeons: in the absence of sample memory, sample frequency is a better predictor of comparison choice than the probability of reinforcement for comparison choice. *Learn. Behav.* 35, 242–261. doi: 10.3758/bf03206430

Dorrance, B. R., and Zentall, T. R. (2001). Imitative learning in Japanese quail depends on the motivational state of the observer at the time of observation. *J. Comp. Psychol.* 115, 62–67. doi: 10.1037/0735-7036.115.1.62

Edwards, C. A., Jagielo, J. A., Zentall, T. R., and Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching-to-sample by pigeons: mediation by reinforcer-specific expectancies. *J. Exp. Psychol. Anim. Behav. Process.* 8, 244–259. doi: 10.1037/0097-7403.8.3.244

Fortes, I., Case, J. P., and Zentall, T. R. (2017). Pigeons, unlike humans, do not prefer near hits in a slot-machine-like task. *Behav. Processes* 138, 67–72. doi: 10.1016/j.beproc.2017.02.012

Friedrich, A. M., Clement, T. S., and Zentall, T. R. (2004). Functional equivalence in pigeons involving a four-member class. *Behav. Processes* 67, 395–403. doi: 10.1016/j.beproc.2004.06.008

Friedrich, A. M., Clement, T. S., and Zentall, T. R. (2005). Discriminative stimuli that follow the absence of reinforcement are preferred by pigeons over those that follow reinforcement. *Learn. Behav.* 33, 337–342. doi: 10.3758/bf03192862

Friedrich, A. M., and Zentall, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behav. Processes* 67, 405–415. doi: 10.1016/j.beproc.2004.07.001

Golding, J. M., and MacLeod, C. M. (eds) (1998). *Intentional Forgetting: Interdisciplinary Approaches*. Mahwah, NJ: Erlbaum.

Harlow, H. F. (1949). The formation of learning sets. *Psychol. Rev.* 56, 51–65. doi: 10.1037/h0062474

Kay, H., and Oldfield-Box, H. (1965). A study of learning-sets in rats with an apparatus using 3-dimensional shapes. *Anim. Behav.* 13, 19–24. doi: 10.1016/0003-3472(65)90066-7

Klein, E. D., Bhatt, R. S., and Zentall, T. R. (2005). Contrast and the justification of effort. *Psychon. Bull. Rev.* 12, 335–339. doi: 10.3758/bf03196381

Laude, J. R., Stagner, J. P., and Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *J. Exp. Psychol. Anim. Behav. Process.* 40, 12–21. doi: 10.1037/xan0000010

MacLin, O. H., Dixon, M. R., Daugherty, D., and Small, S. (2007). Using a computer simulation of three slot machines to investigate a gambler's preference among varying densities of near-miss alternatives. *Behav. Res. Methods* 39, 237–241. doi: 10.3758/bf03193153

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behav. Brain Sci.* 10, 645–695.

Magalhães, P., and White, K. G. (2014). The effect of a prior investment on choice: the sunk cost effect. *J. Exp. Psychol. Anim. Learn. Cogn.* 40, 22–37. doi: 10.1037/xan0000007

Maki, W. S. (1981). "Directed forgetting in animals," in *Information Processing in Animals: Memory Mechanisms*, eds N. E. Spear and R. R. Miller (Hillsdale, NJ: Erlbaum), 199–226.

Meck, W. H., and Church, R. M. (1983). A mode control model of counting and timing processes. *J. Exp. Psychol. Anim. Behav. Process.* 9, 320–334. doi: 10.1037/0097-7403.9.3.320

Navarro, A. D., and Fantino, E. (2005). The sunk cost effect in pigeons and humans. *J. Exp. Anal. Behav.* 83, 1–13. doi: 10.1901/jeab.2005.21-04

Pattison, K. F., Zentall, T. R., and Watanabe, S. (2012). Sunk cost: pigeons (*Columba livia*) too show bias to complete a task rather than shift to another. *J. Comp. Psychol.* 126, 1–9. doi: 10.1037/a0023826

Peterson, G. B. (1984). "How expectancies guide behavior," in *Animal Cognition*, eds H. L. Roitblat, T. G. Bever, and H. S. Terrace (Hillsdale, NJ: Erlbaum), 135–148.

Piaget, J. (1962). *Play, Dreams, and Imitation in Childhood*. New York, NY: Horton.

Rayburn-Reeves, R. M., Miller, H. C., and Zentall, T. R. (2010). "Counting" by pigeons: discrimination of the number of biologically relevant sequential events. *Learn. Behav.* 38, 169–176. doi: 10.3758/lb.38.2.169

Roper, K. L., Kaiser, D. H., and Zentall, T. R. (1995). True directed forgetting in pigeons may occur only when alternative working memory is required on forget-cue trials. *Anim. Learn. Behav.* 23, 280–285. doi: 10.3758/bf03198924

Roper, K. L., and Zentall, T. R. (1994). Directed forgetting in pigeons: the role of retention interval key pecking on delayed matching accuracy. *Learn. Motiv.* 25, 26–44. doi: 10.1006/lmot.1994.1002

Seligman, M. E. P., and Meyer, B. (1970). Chronic fear and ulceration in rats as a function of unpredictability of safety. *J. Comp. Physiol. Psychol.* 73, 202–207. doi: 10.1037/h0030219

Skinner, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* 57, 193–216. doi: 10.1037/h0054367

Slotnick, B. M., and Katz, H. M. (1974). Olfactory learning-set formation in rats. *Science* 185, 796–798. doi: 10.1126/science.185.4153.796

Smith, A. P., and Zentall, T. R. (2016). Suboptimal choice in pigeons: choice is based primarily on the value of the conditioned reinforcer rather than overall reinforcement rate. *J. Exp. Psychol. Anim. Behav. Process.* 42, 212–220. doi: 10.1037/xan0000092

Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., and Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: effects of signal condition and terminal link length. *J. Exp. Anal. Behav.* 53, 219–234. doi: 10.1901/jeab.1990.53-219

Stagner, J. P., Case, J. P., Sticklen, M. F., Duncan, A. K., and Zentall, T. R. (2015). Do pigeons prefer alternatives that include near-hit outcomes? *J. Exp. Psychol. Anim. Behav. Process.* 41, 247–254. doi: 10.1037/xan0000069

Stagner, J. P., and Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychon. Bull. Rev.* 17, 412–416. doi: 10.3758/pbr.17.3.412

Stephens, D. W., and Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.

Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learn. Motiv.* 1, 129–140. doi: 10.1016/0023-9690(70)90079-2

Treisman, A. (1985). Preattentive processing in vision. *Comput. Vis. Graph. Image Process.* 31, 156–177. doi: 10.1016/s0734-189x(85)80004-9

Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., and Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: retention, intertrial interference, and transfer. *J. Exp. Psychol. Anim. Behav. Process.* 15, 264–273. doi: 10.1037/0097-7403.15.3.264

Xia, L., Siemann, M., and Delius, J. D. (2000). Matching of numerical symbols with number of responses bt pigeons. *Anim. Cog.* 3, 35–43. doi: 10.1007/s100710050048

Zentall, T. R., Andrews, D. M., and Case, J. P. (2018). Sameness may be a natural concept that does not require learning. *Psychol. Sci.* 29, 1185–1189. doi: 10.1177/0956797618758669

Zentall, T. R., Andrews, D. M., and Case, J. P. (2019). Contrast between what is expected and what occurs increases pigeon's suboptimal choice. *Anim. Cogn.* 22, 81–87. doi: 10.1007/s10071-018-1223-x

Zentall, T. R., and Clement, T. S. (2002). Memory mechanisms in pigeons: evidence of base-rate neglect. *J. Exp. Psychol. Anim. Behav. Process.* 28, 111–115. doi: 10.1037/0097-7403.28.1.111

Zentall, T. R., Edwards, C. A., Moore, B. S., and Hogan, D. E. (1981). Identity: the basis for both matching and oddity learning in pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 7, 70–86. doi: 10.1037/0097-7403.7.1.70

Zentall, T. R., and Hogan, D. E. (1974). Abstract concept learning in the pigeon. *J. Exp. Psychol.* 102, 393–398. doi: 10.1037/h0035970

Zentall, T. R., and Hogan, D. E. (1976). Pigeons can learn identity, difference, or both. *Science* 191, 408–409. doi: 10.1126/science.191.4225.408

Zentall, T. R., Hogan, D. E., and Edwards, C. A. (1980a). Oddity learning in the pigeon: effect of negative instances, correction, and number of incorrect alternatives. *Anim. Learn. Behav.* 8, 621–629. doi: 10.3758/bf03197778

Zentall, T. R., Hogan, D. E., Edwards, C. A., and Hearst, E. (1980b). Oddity learning in the pigeon as a function of the number of incorrect alternatives. *J. Exp. Psychol. Anim. Behav. Process.* 6, 278–299. doi: 10.1037/0097-7403.6.3.278

Zentall, T. R., Hogan, D. E., and Holder, J. (1974). Comparison of two oddity tasks with pigeons. *Learn. Motiv.* 5, 106–117. doi: 10.1016/0023-9690(74)90041-1

Zentall, T. R., Sherburne, L. M., and Urcuioli, P. J. (1995). Coding of hedonic and nonhedonic samples by pigeons in many-to-one delayed matching. *Anim. Learn. Behav.* 23, 189–196. doi: 10.3758/bf03199934

Zentall, T. R., and Singer, R. A. (2008). Required pecking and refraining from pecking alters judgments of time by pigeons. *Learn. Behav.* 36, 55–61. doi: 10.3758/lb.36.1.55

Zentall, T. R., Singer, R. A., and Miller, H. C. (2008). Matching-to-sample by pigeons: the dissociation of samples from reinforcement contingencies. *Behav. Process.* 78, 185–190. doi: 10.1016/j.beproc.2008.01.015

Zentall, T. R., and Stagner, J. P. (2011). Maladaptive choice behavior by pigeons: an animal analog of gambling (sub-optimal human decision making behavior). *Proc. Biol. Sci.* 278, 1203–1208. doi: 10.1098/rspb.2010.1607

Zentall, T. R., Steirn, J. N., and Jackson-Smith, P. (1990). Memory strategies in pigeons' performance of a radial-arm-maze analog task. *J. Exp. Psychol. Anim. Behav. Process.* 16, 358–371. doi: 10.1037/0097-7403.16.4.358

Zentall, T. R., Sutton, J. E., and Sherburne, L. M. (1996). True imitative learning in pigeons. *Psychol. Sci.* 7, 343–346. doi: 10.1111/j.1467-9280.1996.tb00386.x

Zentall, T. R., Weaver, J. E., and Clement, T. S. (2004). Pigeons group time intervals according to their relative duration. *Psychon. Bull. Rev.* 11, 113–117. doi: 10.3758/bf03206469

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS

Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION

Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices

FOLLOW US

@frontiersin

IMPACT METRICS

Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK

Our network
increases your
article's readership