frontiers | Research Topics

# Animacy in cognition: effects, mechanisms, and theories

**Edited by**
Michael J. Serra, Josefa N. S. Pandeirada and Joshua VanArsdall

**Published in**
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-5696-2
DOI 10.3389/978-2-8325-5696-2

## 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

# Animacy in cognition: effects, mechanisms, and theories

**Topic editors**
Michael J. Serra — Texas Tech University Health Sciences Center, United States
Josefa N. S. Pandeirada — University of Aveiro, Portugal
Joshua VanArsdall — Elmhurst University, United States

**Citation**
Serra, M. J., Pandeirada, J. N. S., VanArsdall, J., eds. (2024). *Animacy in cognition: effects, mechanisms, and theories*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-5696-2

# Table of contents

Frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 05 November 2024
DOI 10.3389/fpsyg.2024.1508218

OPEN ACCESS

EDITED AND REVIEWED BY
Antonino Vallesi,
University of Padua, Italy

*CORRESPONDENCE
Michael J. Serra
✉ michael.serra@ttuhsc.edu

RECEIVED 09 October 2024
ACCEPTED 16 October 2024
PUBLISHED 05 November 2024

CITATION
Serra MJ, Pandeirada JNS and VanArsdall JE (2024) Editorial: Animacy in cognition: effects, mechanisms, and theories.
*Front. Psychol.* 15:1508218.
doi: 10.3389/fpsyg.2024.1508218

# Editorial: Animacy in cognition: effects, mechanisms, and theories

Michael J. Serra[1]*, Josefa N. S. Pandeirada[2] and Joshua E. VanArsdall[3]

[1]Department of Medical Education, School of Medicine, Texas Tech University Health Sciences Center, Lubbock, TX, United States, [2]William James Center for Research, Department of Education and Psychology, University of Aveiro, Aveiro, Portugal, [3]Department of Psychology, Elmhurst University, Elmhurst, IL, United States

KEYWORDS

**animacy, animacy effects, animate, inanimate, cognition, perception, memory, language**

Editorial on the Research Topic
Animacy in cognition: effects, mechanisms, and theories

## 1 Animacy in cognition

The distinction between living (animate) and non-living (inanimate) things is a crucial part of our cognition, with animate things typically receiving more attention in our thoughts and actions (Blakemore et al., 2003; Bugaiska et al., 2019; Nairne et al., 2017; Rakison and Poulin-Dubois, 2001). Beyond simply being "alive" or "not alive", animates differ from inanimates in various ways—they can think, reproduce, move purposefully, and are perceived as being similar to humans (VanArsdall and Blunt, 2022). Living things might have driven the evolution of our cognitive processes given their greater relevance to our survival and reproduction (Nairne et al., 2013, 2017).

Our Research Topic was motivated by two main goals. First, we wanted to highlight new findings on animacy's role in cognition. While cognitive scientists have long studied animacy's influence on attention, perception, language, categories, memory, and other cognitive functions, we continue to refine our understanding of the concept and its influence. Second, we aimed to bridge researchers from various fields—cognitive psychology, linguistics, computer science, human factors, robotics, and more—to deepen our understanding of animacy's effects on our thoughts and actions. Despite varying in scope and topicality, at a higher level, the articles published in this Research Topic all focused on animacy's effects on attention, perception, memory, or language.

## 2 Articles in this Research Topic

### 2.1 Animacy, attention, and perception

Animates naturally capture our attention more than inanimates, and we often perceive animacy in non-living or artificial stimuli that display animate qualities (Rakison and Poulin-Dubois, 2001). However, Loucks et al. showed that not all animate things receive equal attention—mammals, for example, might be prioritized over insects. And though we usually think that perceiving animacy draws our attention, Saito et al. found that the reverse can also happen: we may perceive greater animacy in things that receive continued attention.

Research on animacy perception often focuses on the role of motion (Blakemore et al., 2003). Parovel reviewed how we automatically perceive animacy in simple "Heider-Simmel" animations, arguing that motion helps us identify living things and infer their psychological, emotional, and social characteristics. Torabian and Grossman discussed how children learn to see such movements as goal-directed and eventually attribute them to mental states like beliefs or desires. Animacy perception also has downstream consequences, as Mayer et al. found that people perceive anthropomorphized self-driving vehicles similarly to humans, and that humanlike qualities influence social judgments like responsibility and morality.

### 2.2 Animacy and memory

People tend to remember animate concepts better than inanimate ones (Nairne et al., 2024). While this effect is well-documented in adults, Bugaiska et al. found it occurs in older children but possibly not younger ones, likely due to their still-developing episodic memory skills. Serra and DeYoung showed that the animacy advantage in free-recall exists under both computer-paced and self-paced conditions, and that while participants' beliefs about animacy do not impact the animacy effect directly (DeYoung and Serra, 2021), they can influence processing decisions (e.g., self-paced study) and the size of the effect as a result. Mah et al. replicated Popp and Serra's (2016) finding of an *inanimate* advantage in cued-recall tasks, investigating (and ruling out) semantic similarity among animates as an explanation.

### 2.3 Animacy and language

Living things tend to take precedence over non-living things in our speech and writing (Branigan et al., 2008). Czypionka et al. examined how easily people process German noun–noun pairs and found greater processing fluency when more animate words were included (e.g., "food bowl" vs. "dog food" vs. "sheep dog"). Lobben and Laeng used Construal Level Theory to explain linguistic puzzles involving prominence hierarchies (like animacy), concluding prominent concepts are less psychologically distant from the self. Sá-Leite et al. reviewed the picture-word interference paradigm, a tool for measuring retrievability, and noted that many studies have neglected animacy despite its known enhancement of cognitive and linguistic processing. Westbury explored how people decide if something is animate or not, challenging the notion that this is a simple, binary classification (see also VanArsdall and Blunt, 2022). His analyses suggest that people rely heavily on categorical family resemblance to judge animacy.

## 3 Final thoughts

Together, the articles in this Research Topic highlight key findings and new insights on animacy's role in cognition. The articles on attention and perception not only identify factors that lead to the perception of animacy, but more uniquely how animacy affects downstream judgments and decisions that we make. The memory studies identify new conditions that augment, suppress, and even moderate the animacy advantage in memory; these are important for understanding the process(es) responsible for the effects of animacy on memory. The reasons for the prominence of animacy in language, and the downstream effects of that prioritization, are explored in the articles on language. We hope that by bringing together these diverse insights, this Research Topic deepens our understanding of how animacy influences cognition and inspires further research.

## Author contributions

MS: Writing – review & editing, Writing – original draft, Validation, Project administration, Conceptualization. JP: Writing – review & editing, Validation, Funding acquisition, Conceptualization. JV: Writing – review & editing, Validation, Conceptualization.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. JP's contribution was funded with national funds from Fundação para a Ciência e Tecnologia (refs: UID/04810/2020 and CEECIND/01914/2017).

## Acknowledgments

The authors thank Frontiers and the editors of *Frontiers in Psychology: Cognition* for supporting the present Research Topic. They also thank all the researchers who submitted manuscripts and those who reviewed manuscripts for this Research Topic.

## 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.

## Publisher's note

## References

Blakemore, S. J., Boyer, P., Pachot-Clouard, M., Meltzoff, A., Segebarth, C., and Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cereb. Cortex* 13, 837–844. doi: 10.1093/cercor/13.8.837

Branigan, H. P., Pickering, M. J., and Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118, 172–189. doi: 10.1016/j.lingua.2007.02.003

Bugaiska, A., Grégoire, L., Camblats, A. M., Gelin, M., Méot, A., and Bonin, P. (2019). Animacy and attentional processes: evidence from the Stroop task. *Q. J. Exp. Psychol.* 72, 882–889. doi: 10.1177/1747021818771514

DeYoung, C. M., and Serra, M. J. (2021). Judgments of learning reflect the animacy advantage for memory, but not beliefs about the effect. *Metacogn. Learn.* 16, 711–747. doi: 10.1007/s11409-021-09264-w

Nairne, J. S., Pandeirada, J. N. S., Fernandes, N. L., and Félix, S. B. (2024). "Adaptive memory," in *Reference Module in Neuroscience and Biobehavioral Psychology*, ed. J. Stein (Amsterdam: Elsevier). doi: 10.1016/B978-0-443-15754-7.00016-X

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017). Remembering the living: episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory: the mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

Popp, E. Y., and Serra, M. J. (2016). Adaptive memory: animacy enhances free-recall but impairs cued-recall. *J. Exp. Psychol.* 42, 186–201. doi: 10.1037/xlm0000174

Rakison, D. H., and Poulin-Dubois, D. (2001). Developmental origin of the animate–inanimate distinction. *Psychol. Bull.* 127, 209–228. doi: 10.1037/0033-2909.127.2.209

VanArsdall, J. E., and Blunt, J. R. (2022). Analyzing the structure of animacy: exploring relationships among six new animacy and 15 existing normative dimensions for 1,200 concrete nouns. *Mem. Cognit.* 50, 997–1012. doi: 10.3758/s13421-021-01266-y

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fpsyg.2022.1052729

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University,
United States

REVIEWED BY
Sandra Grinschgl,
University of Graz,
Austria
Robert Hugh Thomson,
United States Military Academy,
United States

*CORRESPONDENCE
Maike M. Mayer
maike.mayer@hhu.de

SPECIALTY SECTION
This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

RECEIVED 24 September 2022
ACCEPTED 24 November 2022
PUBLISHED 04 January 2023

CITATION
Mayer MM, Buchner A and Bell R (2023)
Humans, machines, and double standards?
The moral evaluation of the actions of
autonomous vehicles, anthropomorphized
autonomous vehicles, and human drivers in
road-accident dilemmas.
*Front. Psychol.* 13:1052729.
doi: 10.3389/fpsyg.2022.1052729

# Humans, machines, and double standards? The moral evaluation of the actions of autonomous vehicles, anthropomorphized autonomous vehicles, and human drivers in road-accident dilemmas

Maike M. Mayer*, Axel Buchner and Raoul Bell

Department of Experimental Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

A more critical evaluation of the actions of autonomous vehicles in comparison to those of human drivers in accident scenarios may complicate the introduction of autonomous vehicles into daily traffic. In two experiments, we tested whether the evaluation of actions in road-accident scenarios differs as a function of whether the actions were performed by human drivers or autonomous vehicles. Participants judged how morally adequate they found the actions of a non-anthropomorphized autonomous vehicle (Experiments 1 and 2), an anthropomorphized autonomous vehicle (Experiment 2), and a human driver (Experiments 1 and 2) in otherwise identical road-accident scenarios. The more lives were spared, the better the action was evaluated irrespective of the agent. However, regardless of the specific action that was chosen, the actions of the human driver were always considered more morally justifiable than the corresponding actions of the autonomous vehicle. The differences in the moral evaluations between the human driver and the autonomous vehicle were reduced, albeit not completely eliminated, when the autonomous vehicle was anthropomorphized (Experiment 2). Anthropomorphizing autonomous vehicles may thus influence the processes underlying moral judgments about the actions of autonomous vehicles such that the actions of anthropomorphized autonomous vehicles appear closer in moral justifiability to the actions of humans. The observed differences in the moral evaluation of the actions of human drivers and autonomous vehicles could cause a more critical public response to accidents involving autonomous vehicles compared to those involving human drivers which might be reduced by anthropomorphizing the autonomous vehicles.

KEYWORDS

**autonomous agents, autonomous vehicle, human driver, anthropomorphism, moral evaluation**

## Introduction

In recent years differences in the cognitive processing of information about humans and animals in comparison to inanimate objects has gained increasing attention (Nairne et al., 2017). Whereas these differences had long been ignored in cognitive research, there has been a surge of interest in the prioritization of humans and animals over inanimate objects in memory and attention in recent years (New et al., 2007; Nairne et al., 2013; Altman et al., 2016; Popp and Serra, 2016; Komar et al., 2022). The cognitive mechanisms underlying these differences are hotly debated and remain yet to be identified (VanArsdall et al., 2017; Meinhardt et al., 2018; Popp and Serra, 2018; Bonin et al., 2022). Moral judgement is a domain in which it seems quite obvious to distinguish between humans and inanimate agents such as machines, drones, or artificial intelligence algorithms. At first glance, intuition may suggest that humans are held to a higher moral standard than machines which implies that the actions of humans should be judged more harshly than those of machines (*cf.* Li et al., 2016; Gill, 2020). However, the scientific literature on this issue is mixed. When differences were found, the actions of humans were often judged more leniently than those of machines (e.g., Young and Monroe, 2019). However, it has also been observed that the moral judgement of humans and machines depends on the type of decision that is made. For example, a plausible possibility is that it is specifically the self-sacrifice of a human that may be evaluated more favorably than that of a machine but not the sacrifice of others (*cf.* Sachdeva et al., 2015). Furthermore, it has been proposed that the actions of machines are more likely to be judged according to utilitarian standards than the actions of humans (Malle et al., 2015, 2016). The aim of the present study is to evaluate how robust the differences in the moral evaluation of humans and machines are by testing whether there are reliable differences in the evaluations of the actions of human drivers and autonomous vehicles in road-accident scenarios across conditions that differ in the degree to which they involve utilitarian action and self-sacrifice.

The question of how people judge the actions of autonomous vehicles in comparison to those of humans is of high applied relevance as well. Even though it may yet take several years of development until fully autonomous driving will have reached an acceptable level of safety and reliability (for analyses of accident reports with autonomous vehicles see, e.g., Favarò et al., 2017; Wang et al., 2020), autonomous-driving technology promises to bring many benefits eventually, such as less traffic congestion, potentially resulting in less pollution and reduced energy consumption (Bagloee et al., 2016). Autonomous vehicles may also open a new chapter in mobility-on-demand and car-sharing services that might reduce the individual and societal costs of mobility (Spieser et al., 2014). Once driving technologies will have reached an automation level that does not require humans to intervene, these technologies could increase the comfort of daily driving: Being freed of the driving task, passengers of autonomous vehicles could use the driving time for other activities (Anderson et al., 2016; Bagloee et al., 2016). Given that human error is a major cause of road accidents (National Highway Traffic Safety Administration, 2008), autonomous vehicles are also expected to increase traffic safety in the future (e.g., Anderson et al., 2016). However, accidents cannot be completely avoided. Apart from the fact that no technology will ever function without flaws (e.g., Lin, 2016; Gogoll and Müller, 2017), there is another reason why autonomous vehicles cannot avoid all accidents regardless of their driving performance: they share the roads with human road users whose behaviors are hard to predict (e.g., Lin, 2016; Koopman and Wagner, 2017; Nyholm, 2018).

Fatal accidents with autonomous vehicles can be expected to attract strong media attention during the first years of introducing automated driving technologies into daily traffic (e.g., Shariff et al., 2017; Jelinski et al., 2021). Two of the best-known examples of accidents involving vehicles with automated driving technologies are the 2016 Tesla accident and the 2018 Uber accident. In 2016, a Tesla Model S collided with a semitrailer, resulting in the death of the Tesla's driver (National Transportation Safety Board, 2017). The Tesla accident likely represents the first fatal crash involving a vehicle with automated driving technologies (Yadron and Tynan, 2016). The 2018 Uber accident—in which an Uber vehicle struck and killed a pedestrian (National Transportation Safety Board, 2019)—might be the first fatal crash of a vehicle with automated driving technologies involving a non-motorized road user (Levin and Wong, 2018; Wakabayashi, 2018). The critical coverage of accidents in the media can negatively affect the public perception and acceptance of autonomous vehicles (Shariff et al., 2017; Anania et al., 2018). Currently, the public's opinion on automated vehicles is mixed (Becker and Axhausen, 2017). Some studies indicate prevailing positive anticipation (Winkler et al., 2019) but others show more negative than positive emotions (Hassol et al., 2019; Tennant et al., 2019). People who are skeptical about using automated driving technologies often cite an unwillingness to yield control over the driving task to the autonomous vehicle as a reason for their skeptical attitude (Smith and Anderson, 2017; Winkler et al., 2019). The prospect of machines making decisions that might harm or kill humans might contribute to the discomfort of handing over the control of driving to autonomous vehicles (Li et al., 2016; Malle et al., 2016; Bigman and Gray, 2018). This widespread discomfort with autonomous vehicles making life-and-death decisions may—together with other unsolved problems such as legal issues—delay the adoption of automated driving technologies (e.g., Li et al., 2016).

Therefore, it is of interest to understand how people morally evaluate, in fatal accident scenarios, the actions of autonomous vehicles in comparison to those of human drivers. A large body of the literature is focused on the action that people think autonomous vehicles and humans ought to choose. What is considered the moral choice does not need to be identical for humans and autonomous vehicles. However, the results of several studies suggest that people want humans and machines to make similar choices in road-accident dilemmas (e.g., Bonnefon et al., 2016; Li et al., 2016; Kallioinen et al., 2019; Young and Monroe, 2019). Most of these studies are modeled after the Trolley Dilemma (Foot, 1967; Thomson, 1976, 1985) which can be used to assess moral preferences. In the Trolley Dilemma, a trolley is racing towards five people on the tracks. It is possible to divert the trolley to a sidetrack which will, however, result in the death of an unsuspecting track worker. Is it morally permissible to sacrifice one person to save five? Or should the trolley continue on its path and kill five people? According to utilitarianism, sacrificing one life to save many is morally justifiable based on the principle that decisions should minimize harm and death (e.g., Bentham, 1789; Mill, 2010) while deontology, which focuses on moral rights and duties (e.g., Kant, 2011), may declare the same action as impermissible as it violates the duty not to kill otherwise uninvolved people as a means to an end. A road-accident scenario with an autonomous vehicle fashioned after the trolley dilemma could be the following: An autonomous vehicle is about to crash into one or more pedestrian/s on the road. The only other option being left is to crash the vehicle into a road block which results in the death of the passenger of the autonomous vehicle. Even though there is some degree of variability in people's preferences for the action of the autonomous vehicle in such a moral dilemma (e.g., Awad et al., 2018), one of the most pervasive preferences that have been identified is the utilitarian preference to minimize the number of deaths that result from the accident with the autonomous vehicle (e.g., Li et al., 2016; Awad et al., 2018; Kallioinen et al., 2019; Mayer et al., 2021).

In the present study we focus on the *moral evaluation* of the actions of autonomous vehicles and human drivers in accidents that have already occurred. Whether people evaluate the actions of autonomous vehicles and human drivers differently is a two-part question: First, is there a general cognitive tendency towards evaluating the actions of human drivers less critically than those of autonomous vehicles? Second, are the same moral principles applied to human drivers and autonomous vehicles to morally evaluate their actions? In several studies, Malle and colleagues have used different versions of trolley-type moral dilemmas to test whether people evaluate the actions of robots, artificial intelligence agents, and drones differently than those of humans (Malle et al., 2015, 2016, 2019). Interestingly, the results suggest that the moral evaluation of actions may differ between humans and machines. Specifically, the results of Malle et al. (2015) suggest that "robots are expected—and possibly obligated—to make utilitarian choices" (p. 122) and thus people "regarded the act of sacrificing one person in order to save four (a 'utilitarian' choice) as more permissible for a robot than for a human" (p. 122). There is also evidence indicating that people have a general tendency for blaming autonomous vehicles more harshly than human drivers for their actions in road-accident scenarios (Young and Monroe, 2019). If the latter result turns out to be a robust finding and people are more critical of the actions of autonomous vehicles than of the actions of human drivers, then the question arises as to whether anthropomorphizing autonomous vehicles (that is, assigning humanlike characteristics or properties to them; Epley et al., 2007; Bartneck et al., 2009) will help to shift the moral evaluation of autonomous vehicles closer to the moral evaluation of human drivers. Young and Monroe (2019) found evidence suggesting that describing the decision-making process of the autonomous vehicle in mentalistic terms (i.e., ascribing thoughts and feelings to the autonomous vehicle) may reduce the differences in blame between the autonomous vehicle and a human driver, and may make people's responses to the autonomous vehicle's decisions less negative. In a similar way, Malle et al. (2016) found that presenting a robot with a more human-like appearance reduced the differences in blame for the decisions of robots and humans in comparison to presenting a robot with a mechanical appearance.

The empirical evidence is thus as yet mixed. No overall difference in the evaluation of humans and machines has been found in some studies, but the evaluation may differ depending on whether the action conforms to utilitarian standards or not (Malle et al., 2015, 2016). In another study results have been found that are more in line with a general bias for judging human agents more favorably than machines (Young and Monroe, 2019). It also seems conceivable, if not plausible, that self-sacrifice may play a special role in the moral evaluation of humans and machines (Sachdeva et al., 2015). Specifically, a human driver who sacrifices their own life to spare the lives of others may be more morally praiseworthy than an autonomous vehicle that sacrifices the passenger whom it was designed to protect. These hypotheses were put to an empirical test in the present two experiments.

The primary aim was to test whether there are differences in the moral evaluation of actions of autonomous vehicles and human drivers in road-accident scenarios that differed in whether or not a self-sacrifice of the human driver was involved and in the degree to which utilitarian principles favored this option. Participants morally evaluated the actions of either a human driver or the actions of an autonomous vehicle. They were presented with road-accident scenarios in which the life of the person inside the vehicle was weighted against the lives of one, two, or five pedestrians. If there is a general *bias* toward evaluating humans more favorably than machines (Young and Monroe, 2019), then the actions of autonomous vehicles should be evaluated as more reprehensible and less morally justifiable than those of human drivers, irrespective of whether the action involves a self-sacrifice of the human driver and irrespective of the number of pedestrians on the road. However, a different hypothesis can be derived from the position that people are more likely to judge the actions of machines according to utilitarian standards in comparison to those of humans (Malle et al., 2015). If utilitarian actions are rated more favorably if the agent is a machine rather than a human, then the differences in the moral evaluation between human drivers and autonomous vehicles should crucially depend on the number of lives that can be saved by the action. Specifically, the moral evaluation should be biased in favor of the autonomous vehicle the more lives are spared and biased in favor of the human driver the more deaths are caused by the decision. Finally, based on the hypothesis that there is a special role of self-sacrifice in moral evaluations (Sachdeva et al., 2015), it can be hypothesized that the self-sacrifice of a human driver should be rated as more morally praiseworthy than the sacrifice of the person inside the vehicle by the autonomous vehicle. To anticipate, the results lend clear support to the hypothesis that the actions of the autonomous vehicle are evaluated as less morally justifiable and more reprehensible than those of the human driver. In Experiment 2 we tested whether this negative evaluation tendency can be reduced by anthropomorphizing the autonomous vehicle by assigning a first name ("Alina") to it and by describing it in mentalizing terms (compare Waytz et al., 2014; Hong et al., 2020). The actions of the anthropomorphized autonomous vehicle were indeed evaluated more positively than the actions of the non-anthropomorphized autonomous vehicle which provides further support of the hypothesis that the difference in the moral evaluation of the actions of human drivers and autonomous vehicles can be reduced by assigning humanlike characteristics or properties to the autonomous vehicle (Young and Monroe, 2019). The hypotheses that the moral evaluation of the actions of human drivers versus those of autonomous vehicles may depend on the involvement of utilitarian standards and self-sacrifice received some initial support in Experiment 1 but the associated sample effect sizes were only small and Experiment 2 showed that these interactions were not reliable. We thus conclude that the dominant pattern is that of a general bias towards judging the actions of human drivers more favorably than those of autonomous vehicles.

# Experiment 1

## Methods

The experiment was conducted online using *SoSci Survey* (Leiner, 2019). In total, participation took about 10 min. Both experiments reported here were conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from each participant before the experiment.

### Participants

The sample was recruited via online advertisements. Undergraduate Psychology students received course credit for participating; other participants could enter a lottery to win a € 20 voucher for an online store. To be able to detect even small differences between the judgments of the actions of the human driver and the autonomous vehicle, valid data from 350 participants were necessary (see the next paragraph). Considering the typical data loss in online studies we continued data collection until the end of the week at which data from about 20 percent more than that figure were collected. Of the 444 participants who had started the study, 79 did not complete the experiment. In addition, five participants did not meet the a-priori defined inclusion criteria (being of legal age, having sufficient German language skills, and being able to read the text on screen according to self-reports). Valid data sets of 360 participants (266 women, 94 men), aged between 18 and 80 years ($M = 27$, $SD = 11$) were included in the analyses. Participants were randomly assigned to the human-driver condition ($n = 187$) or the autonomous-vehicle condition ($n = 173$).

We conducted a sensitivity power analysis with *G*Power* (Faul et al., 2007) in which we focused on the agent variable (human driver, autonomous vehicle; between-subjects) and on the action variable (sacrifice the pedestrian/s, sacrifice the person inside the vehicle; within-subject). Given a total sample size of $N = 360$, $\alpha = \beta = 0.05$, and assuming a correlation of $\rho = 0.20$ between the levels of the action variable (estimated based on related results), small effects of about $f = 0.15$ (Cohen, 1988) could be detected for the agent variable. Note that due to the exclusion of the within-subjects number-of-pedestrians variable that was technically necessary to perform the analysis, the power analysis can only give an approximate indication of the sensitivity underlying this study.

### Materials and procedure

First, participants read an introductory text. Depending on the assigned condition, the text stated that human drivers or autonomous vehicles have to handle different traffic situations, including inevitable accidents. The instructions were identical in both conditions, with the only exception that the instructions in the autonomous-vehicle condition included the definition of autonomous vehicles as fully self-driving cars capable of participating in traffic without the need of human intervention (see definition of level 5 driving automation, SAE International, 2021). Participants were then provided with an exemplary description of the accident scenarios they were asked to evaluate later in the experiment. The instructions in the human-driver condition read:

> You will now see various traffic situations in which an accident with a vehicle is unavoidable.
>
> In these situations, a person is driving along a road. Suddenly an obstacle and one or more people appear on the road. Neither timely braking nor evasive action is possible. This means that the driving person only has two options for action:
>
> 1. He/she drives into the obstacle. The person inside the vehicle dies.
>
> 2. He/she drives into the person or persons on the road who dies or die in the process, respectively.
>
> Your task is to evaluate the action of the driving person in the presented traffic situations.

In the autonomous-vehicle condition, the instructions were identical, but "person" was replaced by "autonomous vehicle".

In each of the scenarios (see Figure 1 for an example), the agent (either a human driver or an autonomous vehicle) drove on a single-lane road and was suddenly confronted with an obstacle and at least one pedestrian on the road. As the agent could neither brake nor swerve, only two actions remained: The agent could either sacrifice the person inside the vehicle to save the pedestrian/s by crashing into the obstacle or sacrifice the pedestrian/s to save the person inside the vehicle. The scenarios were depicted as abstract sketches from a bird's eye view. There were either one pedestrian, two pedestrians, or five pedestrians on the road. The agent had already taken one of the two available actions, represented by a yellow arrow. In each scenario, the agent either sacrificed the person inside the vehicle (Figure 1A) or the pedestrian/s (Figure 1B) who died because of the accident. The fatal consequence of the decision was illustrated by a red skull that was presented next to the person inside the vehicle or the pedestrian/s, depending on who was sacrificed. The visual depiction of the scenario was accompanied by a text vignette describing the situation, the action taken, and the action's consequences. For example, if the autonomous vehicle sacrificed five pedestrians to save the person inside the vehicle, the text stated: "The autonomous vehicle drives into the persons on the street. The person inside the vehicle remains unharmed. The five persons on the street are killed." Six different scenarios were obtained by combining two actions and three different numbers of pedestrians. The positions of the obstacle and the pedestrian/s (left or right side of the road) were counterbalanced. Altogether, four presentations of each of the six scenarios were presented, yielding 24 evaluations in total. The scenarios were presented in random order.

Below each image and the corresponding text vignette, participants were asked to evaluate the action (sacrifice the person inside the vehicle vs. sacrifice the pedestrian/s) of the agent (human driver vs. autonomous vehicle) from a moral perspective. The question repeated the agent, the action, and the action's consequences for the two involved parties. For example, if the autonomous vehicle decided to sacrifice five pedestrians to save the person inside the vehicle, the question was: "How do you evaluate, from a moral point of view, the action of the autonomous vehicle to save the person inside the vehicle and to sacrifice the five persons on the street?". Participants were asked to complete the sentence "From a moral point of view, I perceive the action as …" by choosing a category on a scale ranging from "very reprehensible" (1) to "very justifiable" (6). These labels were chosen based on a pilot study ($N = 16$) in which participants were asked to choose from six pairs of negative and positive labels the pair that best captured their moral evaluation of actions in road-accident dilemmas.

As an attention check at the end of the study, the participants were asked to indicate the type of agent that had been involved in the presented scenarios ("A human driver," "An autonomous vehicle," "I do not know"). As the statistical conclusions did not change in both experiments if participants who failed the attention check were included in the statistical analysis, we decided against the exclusion of data, following a recommendation of Elliott et al. (2022).

## Results

In our analyses, we used the multivariate approach to repeated-measures analyses described, for instance, in a primer by

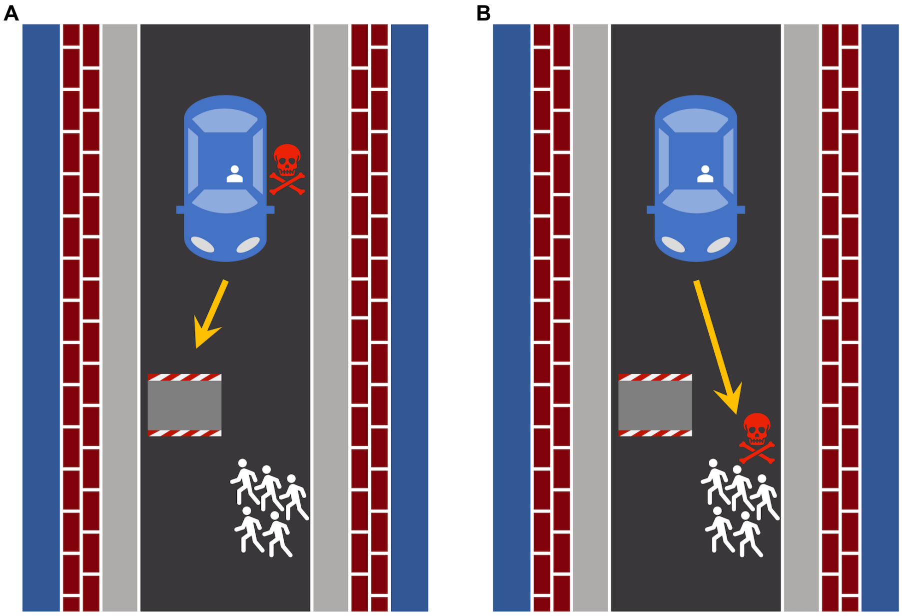


FIGURE 1
Two examples of the illustrations of the road-accident scenarios employed in the experiment. The images depict the two available actions for a road-accident scenario with five pedestrians on the road. **(A)** The person inside the vehicle is sacrificed to save the five pedestrians. **(B)** The five pedestrians are sacrificed to save the person inside the vehicle. The scenarios were created using Microsoft PowerPoint® and Apple Keynote®.

O'Brien and Kaiser (1985). In contrast to the so-called univariate approach to repeated-measures analyses, the multivariate approach does not require the sphericity assumption to be met. This is a major advantage given that the sphericity assumption is violated in almost all repeated measures designs. Exact *F* statistics are reported. The $\alpha$ level was set to 0.05 and all post-hoc comparisons were Bonferroni-Holm adjusted (Holm, 1979). The partial eta squared is used as a sample effect size measure. The mean moral evaluation of the actions as well as the standard errors of the means are depicted in Figure 2.

A 2 (agent: human driver, autonomous vehicle; between-subjects) × 2 (action: sacrifice the pedestrian/s, sacrifice the person inside the vehicle; within-subject) × 3 (number of pedestrians: one pedestrian, two pedestrians, five pedestrians; within-subject) analysis showed that the actions of the human driver ($M = 3.64$, $SE = 0.05$) were evaluated as more morally justifiable than the actions of the autonomous vehicle ($M = 3.18$, $SE = 0.05$), $F(1,358) = 40.51$, $p < 0.001$, $\eta_p^2 = 0.10$. Sacrificing the person inside the vehicle was evaluated more favorably ($M = 4.23$, $SE = 0.06$) than sacrificing the pedestrian/s ($M = 2.61$, $SE = 0.06$), $F(1,358) = 340.82$, $p < 0.001$, $\eta_p^2 = 0.49$. The interaction between agent and action was statistically significant as well, $F(1,358) = 4.72$, $p = 0.030$, $\eta_p^2 = 0.01$. Simple main effect analyses revealed that the human driver's actions were always evaluated more favorably ($M = 4.54$, $SE = 0.08$, for sacrificing the person inside the vehicle; $M = 2.74$, $SE = 0.08$, for sacrificing the pedestrian/s) than those of the autonomous vehicle ($M = 3.89$, $SE = 0.09$, for sacrificing the person inside the vehicle; $M = 2.48$, $SE = 0.08$, for sacrificing the pedestrian/s), but the difference between agents was more pronounced for the decision to sacrifice the person inside the vehicle ($\eta_p^2 = 0.08$) than for the decision to sacrifice the pedestrian/s ($\eta_p^2 = 0.02$).

In addition, there was a significant main effect of the number of pedestrians on the road, $F(2,357) = 23.94$, $p < 0.001$, $\eta_p^2 = 0.12$ ($M = 3.51$, $SE = 0.04$, for one pedestrian; $M = 3.40$, $SE = 0.04$, for two pedestrians; $M = 3.35$, $SE = 0.04$, for five pedestrians). The direction of this effect, however, depended on the action that was taken, $F(2,357) = 187.20$, $p < 0.001$, $\eta_p^2 = 0.51$. An increase in the number of pedestrians led to a significant increase in the moral evaluation of sacrificing the person inside the vehicle ($M = 3.94$, $SE = 0.06$, for one pedestrian; $M = 4.19$, $SE = 0.06$, for two pedestrians; $M = 4.55$, $SE = 0.06$, for five pedestrians; all comparisons $p < 0.001$) while it led to a significant decrease in the moral evaluation of sacrificing the pedestrian/s ($M = 3.08$, $SE = 0.06$, for one pedestrian; $M = 2.61$, $SE = 0.06$, for two pedestrians; $M = 2.15$, $SE = 0.06$, for five pedestrians; all comparisons $p < 0.001$). The effect of the number of pedestrians did not differ between agents, $F(2,357) = 2.73$, $p = 0.067$, $\eta_p^2 = 0.02$.

Finally, there was a significant three-way interaction, $F(2,357) = 4.54$, $p = 0.011$, $\eta_p^2 = 0.02$. We conducted a 2 (action: sacrifice the pedestrian/s, sacrifice the person inside the vehicle; within-subject) × 3 (number of pedestrians: one pedestrian, two pedestrians, five pedestrians; within-subject) repeated-measures analysis for each of the two agents separately. The action of sacrificing the person inside the vehicle was evaluated as significantly more morally justifiable than the action of sacrificing the pedestrian/s for both the human driver, $F(1,186) = 241.01$, $p < 0.001$, $\eta_p^2 = 0.56$, and the autonomous vehicle, $F(1,172) = 117.40$, $p < 0.001$, $\eta_p^2 = 0.41$. There was a significant main effect of the number of pedestrians on the road for both the human driver,

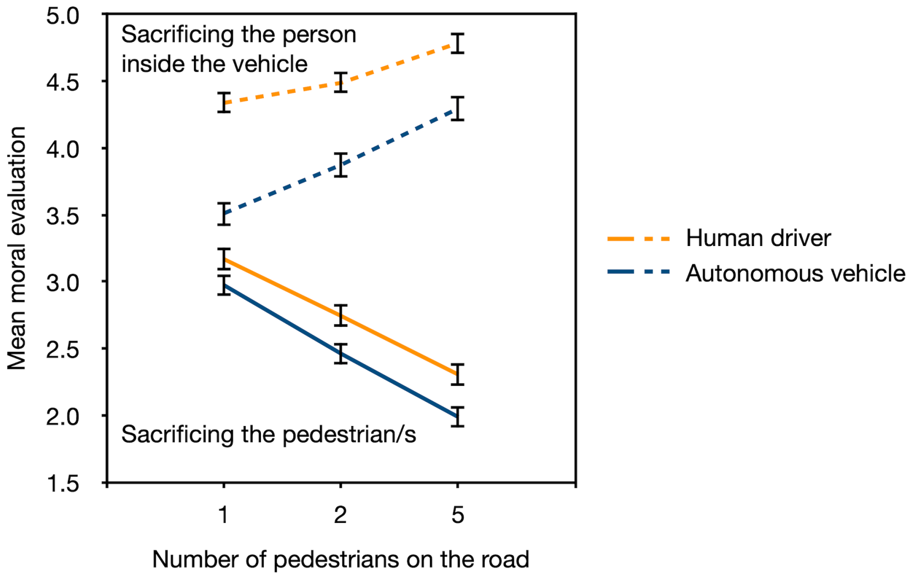


FIGURE 2
The mean moral evaluation of the actions (sacrificing the person inside the vehicle [dashed lines], sacrificing the pedestrian/s [solid lines]) as a function of the number of pedestrians on the road (1, 2, and 5) and the agent (human driver, autonomous vehicle). The moral-evaluation scale ranged from "very reprehensible" (1) to "very justifiable" (6). The error bars represent standard errors of the means.

$F(2,185) = 18.70$, $p < 0.001$, $\eta_p^2 = 0.17$, and the autonomous vehicle, $F(2,171) = 6.79$, $p = 0.001$, $\eta_p^2 = 0.07$. Finally, the interaction between action and the number of pedestrians on the road was significant for both the human driver, $F(2,185) = 90.23$, $p < 0.001$, $\eta_p^2 = 0.49$, and the autonomous vehicle, $F(2,171) = 96.98$, $p < 0.001$, $\eta_p^2 = 0.53$. When the person inside the vehicle was sacrificed, the action was evaluated as significantly more morally justifiable with an increasing number of pedestrians on the road for both the human driver ($M = 4.34$, $SE = 0.08$, for one pedestrian; $M = 4.49$, $SE = 0.08$, for two pedestrians; $M = 4.79$, $SE = 0.08$, for five pedestrians; all comparisons $p < 0.001$) and the autonomous vehicle ($M = 3.51$, $SE = 0.09$, for one pedestrian; $M = 3.87$, $SE = 0.09$, for two pedestrians; $M = 4.30$, $SE = 0.10$, for five pedestrians; all comparisons $p < 0.001$) while the reverse pattern emerged when the decision was to sacrifice the pedestrian/s for both the human driver ($M = 3.17$, $SE = 0.09$, for one pedestrian; $M = 2.74$, $SE = 0.08$, for two pedestrians; $M = 2.31$, $SE = 0.08$, for five pedestrians; all comparisons $p < 0.001$) and the autonomous vehicle ($M = 2.98$, $SE = 0.09$, for one pedestrian; $M = 2.46$, $SE = 0.08$, for two pedestrians; $M = 1.99$, $SE = 0.08$, for five pedestrians; all comparisons $p < 0.001$). The three-way interaction thus does not indicate that fundamentally different moral principles were applied to the evaluation of the actions of the human driver and to the evaluation of the actions of the autonomous vehicle, but the effect of the number of pedestrians on the moral evaluation of the action of sacrificing the person inside the vehicle was somewhat less pronounced for the human driver than for the autonomous vehicle.

## Discussion

The present study served to test whether there are differences in the moral evaluation of the actions of human drivers and autonomous vehicles. The most important finding is that the actions of the human driver were always evaluated as more morally justified than the actions of the autonomous vehicle, which suggests that there is a moral-evaluation bias in favor of the human driver.

Another aim of the present study was to evaluate whether the actions of human drivers are evaluated according to different moral principles than those of autonomous vehicles. Before addressing the qualitative differences in the moral evaluation of the human driver and the autonomous vehicle, we want to draw attention to the fact that there are striking similarities. Overall, the moral evaluations of the actions of the human driver and the autonomous vehicle depended on both the type of action that was evaluated (sacrificing the person inside the vehicle or the pedestrian/s) and the number of pedestrians on the road. Regardless of whether the actions of the human driver or the autonomous vehicle were evaluated, participants regarded actions that spared the maximum number of lives as more morally justifiable than other actions. The favorable evaluations of utilitarian actions are in line with demonstrations of overall preferences for utilitarian actions of human and machine agents in other studies (e.g., Li et al., 2016; Kallioinen et al., 2019). There was an interaction between agent and action, indicating that the decision of the human driver to self-sacrifice was evaluated more favorably than the action of the autonomous vehicle to sacrifice the person inside the vehicle, in line with a special role of self-sacrifice in moral judgement (Sachdeva et al., 2015). Furthermore, there was a three-way interaction between agent, action, and number of pedestrians, suggesting that the evaluation of the human driver's decisions to sacrifice themselves was less dependent on the number of pedestrians on the road than the evaluation of the autonomous vehicle's decisions to sacrifice the person inside the vehicle. At first glance, this finding is in line with the assumption that the moral evaluation of the actions of the autonomous vehicle depends more on utilitarian standards than the moral evaluation of the actions of the human driver. However, this finding can easily be explained by the fact that the decision of the human driver to self-sacrifice received favorable moral evaluations already when this meant sparing the life of only one pedestrian, and this favorable evaluation was hard to boost when more pedestrians were saved at the expense of the driver. It is also worth pointing out that the sample effect sizes of these interactions are quite small (the sample effect sizes of the two-way interaction between agent and action and the three-way interaction between agent, action, and number of pedestrians were $\eta_p^2 = 0.01$ and $\eta_p^2 = 0.02$, respectively). Therefore, it seems questionable whether interactions of such small magnitude can be robustly replicated in future experiments (see the Discussion of Experiment 2). Furthermore, the evaluations of the human driver's actions were always more favorable than those of the autonomous vehicle irrespective of whether self-sacrifice or utilitarian actions were involved or not (*cf.* Figure 1). The dominant finding is thus that there is an overall bias towards a more favorable evaluations of the actions of the human driver over those of the autonomous vehicle.

Experiment 2 had two main aims. The first aim was to test whether the differences in the moral evaluations of the actions of the human driver and the autonomous vehicle found in Experiment 1 could be replicated. Due to the small sample effect sizes of the critical two-way and three-way interactions observed in Experiment 1, we thought it important to perform a high-powered replication before drawing any firm conclusions. Based on the sample effect sizes observed in Experiment 1, we expected that the main effect of agent—reflecting a more critical evaluation of the actions of the autonomous vehicle in comparison to those of the human driver—should also be obtained in Experiment 2 whereas it was questionable whether the two-way interaction between agent and action and the three-way interaction between agent, action, and number of pedestrians—that were both associated with small sample effect sizes—could be replicated. The second aim of Experiment 2 was to test whether anthropomorphizing the autonomous vehicle may help to narrow the gap between the moral evaluation of actions taken by an autonomous vehicle and a human driver in inevitable accidents with human fatalities. If the difference in the moral evaluation in the actions of the human driver and the autonomous vehicle is caused by some fundamental difference in the moral evaluation of humans and machines, anthropomorphizing the autonomous vehicle (that is, making it more similar to human agents) should reduce the differences in the moral evaluations.

# Experiment 2

## Methods

### Participants

Participants were recruited from the online research panels of GapFish GmbH (Berlin, Germany). Of the 892 participants who started the study, 80 did not complete the experiment, 10 did not meet the *a-priori* defined inclusion criteria (being of legal age, having sufficient German language skills, and being able to read the text on screen according to self-reports), and 37 either withdrew their consent to the processing of their data or reported that not all pictures had been displayed during the study. Additionally, 10 participants were excluded due to double participation. The final sample consisted of 755 participants (317 women, 437 men, and 1 diverse), aged between 18 and 87 years ($M = 46$, $SD = 15$). Participants were randomly assigned to the human-driver condition ($n = 248$), the anthropomorphized-autonomous-vehicle condition ($n = 250$), or the autonomous-vehicle condition ($n = 257$).

Given the goal to test whether anthropomorphizing the autonomous vehicle would cause the moral evaluations of the autonomous vehicle to shift towards the more favorable evaluation of the human driver, it seemed important to increase the sensitivity of the statistical tests in Experiment 2. We decided to collect data from at least twice as many participants as in Experiment 1 and stopped data collection at the end of the week this criterion was surpassed. A sensitivity power analysis parallel to that conducted for Experiment 1 showed that, given a total sample size of $N = 755$ and otherwise identical assumptions, small effects of about $f = 0.10$ (Cohen, 1988) could be detected for comparisons involving two levels of the agent variable (e.g., anthropomorphized autonomous vehicle vs. autonomous vehicle) on the moral evaluations.

### Materials and procedure

Materials and procedure—including the descriptions of the autonomous vehicle and the human driver—were identical to those of Experiment 1 with one exception. In addition to the two experimental conditions used in the first experiment (human driver and autonomous vehicle), we included a third condition with an anthropomorphized autonomous vehicle. This vehicle was introduced as a self-driving vehicle controlled by an intelligent driving system called "Alina." Subsequently, the vehicle was only referred to by its name.

## Results

The data were analyzed in the same way as in Experiment 1. The mean moral evaluation of the actions as well as the standard errors of the means are depicted in Figure 3.

As in Experiment 1, there was a significant main effect of the agent, $F(2,752) = 24.72$, $p < 0.001$, $\eta_p^2 = 0.06$. Orthogonal Helmert contrasts showed that the actions of the human driver ($M = 3.39$, $SE = 0.05$) were evaluated more favorably from a moral perspective than the actions of both vehicle types together, $F(1,752) = 37.76$,

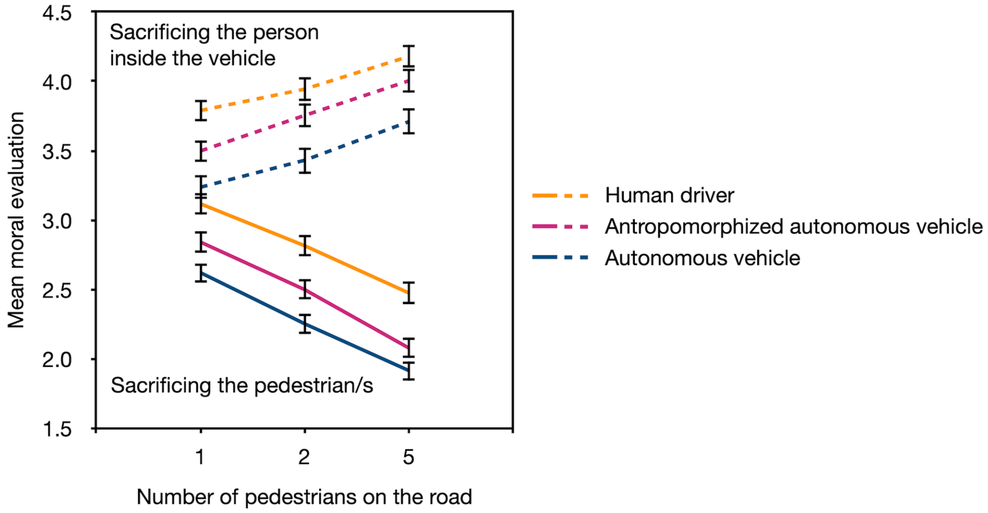


FIGURE 3
The mean moral evaluation of the actions (sacrificing the person inside the vehicle [dashed lines], sacrificing the pedestrian/s [solid lines]) as a function of the number of pedestrians on the road (1, 2, and 5) and the agent (human driver, anthropomorphized autonomous vehicle, autonomous vehicle). The moral-evaluation scale ranged from "very reprehensible" (1) to "very justifiable" (6). The error bars represent standard errors of the mean.

$p < 0.001$, $\eta_p^2 = 0.05$, and that the actions of the anthropomorphized autonomous vehicle ($M = 3.11$, $SE = 0.05$) were evaluated more favorably than the actions of the autonomous vehicle ($M = 2.86$, $SE = 0.06$), $F(1,752) = 11.35$, $p = 0.001$, $\eta_p^2 = 0.01$. Sacrificing the person inside the vehicle ($M = 3.72$, $SE = 0.05$) was evaluated as more morally justifiable than sacrificing the pedestrian/s ($M = 2.51$, $SE = 0.04$), $F(1,752) = 399.57$, $p < 0.001$, $\eta_p^2 = 0.35$. The interaction between these two variables was not significant, $F(2,752) = 0.30$, $p = 0.742$, $\eta_p^2 < 0.01$.

The main effect of the number of pedestrians on the road was significant, $F(2,751) = 28.18$, $p < 0.001$, $\eta_p^2 = 0.07$ ($M = 3.18$, $SE = 0.03$, for one pedestrian; $M = 3.11$, $SE = 0.03$, for two pedestrians; $M = 3.06$, $SE = 0.03$, for five pedestrians). As in Experiment 1, the direction of this effect depended on the action that was taken, $F(2,751) = 219.12$, $p < 0.001$, $\eta_p^2 = 0.37$. An increase in the number of pedestrians led to a significant increase in the moral evaluation of the act of sacrificing the person inside the vehicle ($M = 3.50$, $SE = 0.05$, for one pedestrian; $M = 3.71$, $SE = 0.05$, for two pedestrians; $M = 3.96$, $SE = 0.05$, for five pedestrians; all comparisons $p < 0.001$) while it led to a significant decrease in the moral evaluation of sacrificing the pedestrian/s ($M = 2.86$, $SE = 0.04$, for one pedestrian; $M = 2.52$, $SE = 0.04$, for two pedestrians; $M = 2.15$, $SE = 0.04$, for five pedestrians; all comparisons $p < 0.001$). The effect of the number of pedestrians did not differ among the agents, $F(4, 1502) = 1.10$, $p = 0.353$, $\eta_p^2 < 0.01$.

The three-way interaction was also not significant, $F(4, 1502) = 0.86$, $p = 0.485$, $\eta_p^2 < 0.01$. When the person inside the vehicle was sacrificed, the action was evaluated as significantly more morally justifiable with an increasing number of pedestrians on the road for the human driver ($M = 3.79$, $SE = 0.08$, for one pedestrian; $M = 3.94$, $SE = 0.08$, for two pedestrians; $M = 4.18$, $SE = 0.08$, for five pedestrians; all comparisons $p < 0.001$), the anthropomorphized autonomous vehicle ($M = 3.50$, $SE = 0.08$, for one pedestrian; $M = 3.75$, $SE = 0.08$, for two pedestrians; $M = 4.00$, $SE = 0.09$, for five pedestrians; all comparisons $p < 0.001$), and the autonomous vehicle ($M = 3.24$, $SE = 0.09$, for one pedestrian; $M = 3.43$, $SE = 0.09$, for two pedestrians; $M = 3.71$, $SE = 0.09$, for five pedestrians; all comparisons $p < 0.001$). When the decision was to sacrifice the pedestrian/s, the opposite pattern was found for the human driver ($M = 3.12$, $SE = 0.08$, for one pedestrian; $M = 2.81$, $SE = 0.07$, for two pedestrians; $M = 2.47$, $SE = 0.08$, for five pedestrians; all comparisons $p < 0.001$), the anthropomorphized autonomous vehicle ($M = 2.84$, $SE = 0.08$, for one pedestrian; $M = 2.50$, $SE = 0.07$, for two pedestrians; $M = 2.08$, $SE = 0.07$, for five pedestrians, respectively; all comparisons $p < 0.001$), and the autonomous vehicle ($M = 2.62$, $SE = 0.08$, for one pedestrian; $M = 2.25$, $SE = 0.07$, for two pedestrians; $M = 1.91$, $SE = 0.07$, for five pedestrians; all comparisons $p < 0.001$).

### Discussion

Despite differences in the distributions of age and gender between the samples of Experiments 1 and 2, the overall pattern of results is very consistent. The global difference in the moral evaluation of the actions of the human driver and the autonomous vehicle observed in Experiment 1 was replicated in Experiment 2. Regardless of the type of action or the number of pedestrians on the road, the actions of the human driver were evaluated most favorably while the actions of the non-anthropomorphized autonomous vehicle were evaluated least favorably. This finding suggests that the actions of autonomous vehicles are more likely to be evaluated as morally reprehensible than the actions of human drivers. We were interested in whether it would be possible to narrow this evaluation gap by anthropomorphizing the autonomous vehicle. The anthropomorphization significantly reduced the evaluation gap between the human driver and the autonomous vehicle. The difference in the moral evaluation of the human driver and the autonomous vehicle was not completely eliminated but note that stronger manipulations (e. g., adding a human voice or other characteristics) may have stronger effects.

In addition, the results of Experiment 2 add to the evidence suggesting that utilitarian considerations are involved in the moral evaluations of both the human driver and the autonomous vehicles. Specifically, participants evaluated the actions of both the human driver and the autonomous vehicles more favorably if they were compatible with the utilitarian principle of saving more lives. Furthermore, if the life of one person inside the vehicle had to be weighed against the life of one pedestrian, participants evaluated the action that spared the life of the pedestrian more favorably than the action that spared the life of the person inside the vehicle. Despite the large sample size, we found no evidence that either of these effects differed as a function of whether the agent was a human driver or an autonomous vehicle. In terms of statistical tests, a statistically significant two-way interaction between agent and action and a statistically significant three-way interaction between agent, action, and number of pedestrians could have been interpreted as evidence of qualitative differences in the moral evaluation of the actions between human drivers and autonomous vehicles. These interactions were statistically significant but numerically small in Experiment 1 and clearly failed to replicate in Experiment 2 despite an increase in sample size which, in turn, resulted in an increased sensitivity to detect such effects. Overall, the evidence suggests that these interactions are negligeable.

## General discussion

The aim of the present study was to test whether differences in the moral evaluation of the actions of humans and machines can robustly be found in road-accident scenarios. The results provide support of the hypothesis that the actions of human drivers are judged more leniently than those of autonomous vehicles. Two main results support this hypothesis: (1) The results of both experiments consistently show that the actions of human drivers are judged as morally superior to those of autonomous vehicles. (2) Experiment 2 shows that anthropomorphizing the autonomous vehicle is effective in reducing the difference in the moral evaluations between the human driver and the autonomous vehicle. The results of

Experiment 1 suggested that the more favorable moral evaluation of human drivers in comparison to autonomous vehicles may be modulated by utilitarian standards and self-sacrifice as suggested by previous theorizing (Malle et al., 2015; Sachdeva et al., 2015). However, both the two-way interaction between agent (human driver, autonomous vehicle) and action (sacrificing the person inside the vehicle, sacrificing the pedestrian/s) and the three-way interaction between agent, action, and number of pedestrians on the road (one pedestrian, two pedestrians, or five pedestrians) were only small in terms of sample effect sizes, and both interactions were not replicated in Experiment 2 despite the large sample size which provided favorable conditions for replicating the interactions if they were robust. The dominant pattern in both experiments was thus that the actions of human drivers were evaluated more favorably than those of autonomous vehicles. Descriptively, this pattern was present irrespective of whether the person inside the vehicle or the pedestrian/s on the road were sacrificed and irrespective of whether the action was in line with utilitarian standards or not. The more favorable evaluation of the actions of human drivers over actions of autonomous vehicles that was observed in the present experiments is thus primarily due to a general bias rather than to a differential reliance on utilitarian principles or the specific moral admiration of the self-sacrifice of the human driver.

The present study thus helps to determine the nature of the differences in the moral evaluations of human drivers versus autonomous vehicles. The dominant pattern is that of a general bias toward judging the actions of human drivers as more morally permissible than those of autonomous vehicles. More research is necessary to understand the exact processes that underlie the bias toward the more favorable evaluation of the actions of human drivers in comparison to those of autonomous vehicles. One possibility is that of a moral-evaluation bias, that is, a general aversion against machines making life-and-death decisions (Bigman and Gray, 2018). This interpretation is in line with the finding of Young and Monroe (2019) that people blame autonomous vehicles more harshly than human drivers for their decisions in accident scenarios and with the finding of Gogoll and Uhl (2018) that people are reluctant to delegate a moral task to a machine. Possibly, this critical view of machines making moral decisions may stem from the fact that it is relatively easy for participants to put themselves in the human driver's shoes and to imagine having experienced a conflict when making the decision which seems to be harder in case of a machine agent (Scheutz and Malle, 2021). Consequently, people might more easily justify (and potentially condone) the actions of a human agent compared to the actions of an autonomous vehicle. For example, participants might take into consideration that human drivers have to make spontaneous split-second decisions in critical traffic situations while autonomous vehicles are programmed in advance, the latter of which might make it easier to suspect bad intent. Further, the prospect of increased traffic safety—which is often linked to the introduction of autonomous vehicles—might also hint at higher expectations regarding the driving performance of autonomous vehicles compared to the performance of human drivers. There is, for example, evidence to suggest that the risks associated with autonomous vehicles are tolerated less than the risks associated with human drivers (Liu et al., 2020). A fatal accident might therefore represent an expectation violation in the case of an autonomous vehicle, which might contribute to the more negative moral evaluation of the actions of an autonomous vehicle compared to the actions of a human driver. Here it seems relevant that manipulations that make machines more human-like, for example by ascribing mental properties such as thoughts and feelings to them (Young and Monroe, 2019), reduce the evaluation gap between human drivers and autonomous vehicles. In line with this interpretation, Experiment 2 showed that anthropomorphizing the autonomous vehicle shifted the critical moral evaluation of the actions of the vehicle in the road-accident scenarios towards the more positive moral evaluations of the same actions performed by a human driver. This finding is in line with the observation that anthropomorphizing can positively affect the perception of a machine agent (e.g., Gong, 2008; Lee et al., 2015; Niu et al., 2018). A limitation of the present study is that the specific attributes that are responsible for the more human-like judgement of the anthropomorphized vehicle are yet to be determined. The autonomous vehicle was assigned a name and described as an intelligent driving system. Humans and machines differ in a number of characteristics such as their perceived agency and the mental capacities (perception, emotion, learning, and thinking) which people might attribute to them. These attributions may change as the technology and people's experience with autonomous driving systems and computer algorithms evolves. We do not know how exactly the attribution of humanlike properties to the autonomous vehicles have influenced participants' assumptions about these characteristics and, consequently, the exact cognitions that underlie the observed moral-evaluation bias. Understanding these underlying processes is an important goal for future studies.

When interpreting the present results, it should be considered that participants were asked to evaluate abstract road-accident scenarios fashioned after moral dilemmas which is a common research paradigm to examine moral evaluations and decisions (for other studies employing abstract scenarios see, e.g., Bonnefon et al., 2016; Awad et al., 2018; Frank et al., 2019). Moral dilemmas are useful to identify factors of a scenario that are relevant for its evaluation (e.g., Hauser et al., 2007; Keeling, 2020), to probe different ethical principles or theories and to investigate moral intuitions and moral decision making (e.g., Hauser et al., 2007; Cushman and Greene, 2012; Goodall, 2016; Wolkenstein, 2018). Abstract scenarios obviously fall short of real-life accidents experienced first-hand but they bear resemblance to newspaper reports on accidents. Newspaper reports probably are associated with low levels of immersion as they primarily describe the accident itself and perhaps the accident's causes and consequences. In that sense abstract scenarios seem suitable for investigating how the public will react to accidents with autonomous vehicles they read about in the newspaper. This seems quite relevant given that it is more likely for the majority of people to learn about accidents from newspaper reports than by witnessing, or being directly involved in, an accident. Nevertheless, it has to be counted among the limitations of the present study that we cannot draw conclusions about situations in which there is a more direct involvement in the accidents. Furthermore, the conclusions of the present study are necessarily limited by the specific conditions that were included in the present experiments. While the involvement of utilitarian standards and the presence or absence of self-sacrifice were varied, moral evaluations may depend on many other factors such as the violation of rules and obligations or social prejudices and biases (e.g., Awad et al., 2018; Frank et al., 2019). The present study thus cannot shed light on the degree to which the moral evaluations of humans and machines are influenced by these factors.

The presents study's aim was neither to develop guidelines for programming autonomous vehicles (e.g., Wolkenstein, 2018) nor to determine whether autonomous vehicles or other machines can be regarded as moral agents (for some points of view see, e.g., Li et al., 2016; Gogoll and Müller, 2017; Bonnefon et al., 2019; Scheutz and Malle, 2021) and in how far concepts such as responsibility, liability, or blame can or should be assigned to machines. We focused on the moral evaluation of actions in critical traffic situations as this might represent a first step in understanding the public's reaction to accidents with autonomous vehicles. The evaluation of the agent itself or questions of blame and responsibility are separate issues. Investigating how actions of different agents are perceived in critical traffic situations is important in order to anticipate potential problems regarding the acceptance of autonomous vehicles. In this respect, the perception and opinion of ordinary people is especially relevant as they have to accept the technology (Malle et al., 2019). Gogoll and Uhl (2018) have argued that a disliking of autonomous vehicles making moral decisions has the potential to slow down automation in driving. Considering and openly addressing differences in the moral evaluation of human drivers and autonomous vehicles could thus be beneficial for the introduction and the success of autonomous driving technologies.

In conclusion, the present study contributes to our understanding of how moral norms are applied to machine agents. People have a bias toward judging actions of human drivers as morally superior to identical actions (with identical consequences) of autonomous vehicles. Accidents resembling moral dilemmas might be rare but they are emotionally salient (Bonnefon et al., 2016) and there is evidence to suggest that moral dilemmas are regarded as an important challenge for autonomous vehicles (Gill, 2021). Moral decisions—which include decisions about how to distribute harm in accident situations—have the potential to affect the perception of autonomous vehicles via media coverage of accidents (e.g., Anania et al., 2018). At least during the early introduction phases, a strong media attention to accidents involving autonomous vehicles seems likely (Shariff et al., 2017). A more negative moral evaluation of the actions of autonomous vehicles in comparison to those of human drivers may have negative effects on the acceptance of autonomous driving technologies (see also Gogoll and Uhl, 2018). Therefore, it seems relevant to search for interventions that may decrease the differential moral evaluations of human drivers and autonomous vehicles. The results of Experiment 2 suggest that anthropomorphizing autonomous vehicles can reduce the action evaluation gap between autonomous vehicles and human drivers. Thus, assigning human characteristics to autonomous vehicles might represent a promising intervention for transferring some of the leniency people display towards human drivers to autonomous vehicles.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/3xamb/.

## 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

MM, AB, and RB participated in planning the research, designing the experiments, analyzing the collected data, and interpreting the results. MM created the materials, collected the data, and wrote the initial manuscript. AB and RB provided feedback and contributed to the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

The publication of this article is supported by the open access fund of Heinrich Heine University Düsseldorf.

## Acknowledgments

The presented experiments are part of the dissertation of MM (Mayer, 2022).

## 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.

## Publisher's note

## References

Altman, M. N., Khislavsky, A. L., Coverdale, M. E., and Gilger, J. W. (2016). Adaptive attention: how preference for animacy impacts change detection. *Evol. Hum. Behav.* 37, 303–314. doi: 10.1016/j.evolhumbehav.2016.01.006

Anania, E. C., Rice, S., Walters, N. W., Pierce, M., Winter, S. R., and Milner, M. N. (2018). The effects of positive and negative information on consumers' willingness to ride in a driverless vehicle. *Transp. Policy* 72, 218–224. doi: 10.1016/j.tranpol.2018.04.002

Anderson, J. M., Nidhi, K., Stanley, K. D., Sorensen, P., Samaras, C., and Oluwatola, O. A. (2016). *Autonomous Vehicle Technology: A Guide for Policymakers*. Santa Monica, CA: RAND Corporation.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., et al. (2018). The moral machine experiment. *Nature* 563, 59–64. doi: 10.1038/s41586-018-0637-6

Bagloee, S. A., Tavana, M., Asadi, M., and Oliver, T. (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *J. Mod. Transport.* 24, 284–303. doi: 10.1007/s40534-016-0117-3

Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* 1, 71–81. doi: 10.1007/s12369-008-0001-3

Becker, F., and Axhausen, K. W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. *Transportation* 44, 1293–1306. doi: 10.1007/s11116-017-9808-9

Bentham, J. (1789). *An Introduction to the Principles Of Morals and Legislation*. London: T. Payne and sons.

Bigman, Y. E., and Gray, K. (2018). People are averse to machines making moral decisions. *Cognition* 181, 21–34. doi: 10.1016/j.cognition.2018.08.003

Bonin, P., Thiebaut, G., Bugaiska, A., and Méot, A. (2022). Mixed evidence for a richness-of-encoding account of animacy effects in memory from the generation-of-ideas paradigm. *Curr. Psychol.* 41, 1653–1662. doi: 10.1007/s12144-021-02666-8

Bonnefon, J.-F., Shariff, A., and Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science* 352, 1573–1576. doi: 10.1126/science.aaf2654

Bonnefon, J.-F., Shariff, A., and Rahwan, I. (2019). The trolley, the bull bar, and why engineers should care about the ethics of autonomous cars [point of view]. *Proc. IEEE* 107, 502–504. doi: 10.1109/JPROC.2019.2897447

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hilsdale, NJ: Lawrence Earlbaum Associates.

Cushman, F., and Greene, J. D. (2012). Finding faults: how moral dilemmas illuminate cognitive structure. *Soc. Neurosci.* 7, 269–279. doi: 10.1080/17470919.2011.614000

Elliott, E. M., Bell, R., Gorin, S., Robinson, N., and Marsh, J. E. (2022). Auditory distraction can be studied online! A direct comparison between in-person and online experimentation. *J. Cogn. Psychol.* 34, 307–324. doi: 10.1080/20445911.2021.2021924

Epley, N., Waytz, A., and Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychol. Rev.* 114, 864–886. doi: 10.1037/0033-295X.114.4.864

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Favarò, F. M., Nader, N., Eurich, S. O., Tripp, M., and Varadaraju, N. (2017). Examining accident reports involving autonomous vehicles in California. *PLoS One* 12:e0184952. doi: 10.1371/journal.pone.0184952

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Rev.* 5, 5–15.

Frank, D.-A., Chrysochou, P., Mitkidis, P., and Ariely, D. (2019). Human decision-making biases in the moral dilemmas of autonomous vehicles. *Sci. Rep.* 9:13080. doi: 10.1038/s41598-019-49411-7

Gill, T. (2020). Blame it on the self-driving car: how autonomous vehicles can alter consumer morality. *J. Consum. Res.* 47, 272–291. doi: 10.1093/jcr/ucaa018

Gill, T. (2021). Ethical dilemmas are really important to potential adopters of autonomous vehicles. *Ethics Inf. Technol.* 23, 657–673. doi: 10.1007/s10676-021-09605-y

Gogoll, J., and Müller, J. F. (2017). Autonomous cars: in favor of a mandatory ethics setting. *Sci. Eng. Ethics* 23, 681–700. doi: 10.1007/s11948-016-9806-x

Gogoll, J., and Uhl, M. (2018). Rage against the machine: automation in the moral domain. *J. Behav. Exp. Econ.* 74, 97–103. doi: 10.1016/j.socec.2018.04.003

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Comput. Human Behav.* 24, 1494–1509. doi: 10.1016/j.chb.2007.05.007

Goodall, N. J. (2016). Away from trolley problems and toward risk management. *Appl. Artif. Intell.* 30, 810–821. doi: 10.1080/08839514.2016.1229922

Hassol, J., Perlman, D., Chajka-Cadin, L., and Shaw, J. (2019). Understanding surveys of public sentiment regarding automated vehicles: Summary of results to date and implications of past research on the dynamics of consumer adoption (DOT-VNTSC-FHWA-20-03/FHWA-JPO-19-764). Cambridge (MA): U.S. Department of Transportation & John A. Volpe National Transportation Systems Center. Available at: https://rosap.ntl.bts.gov/view/dot/43628 (Accessed November 28, 2022).

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind Lang.* 22, 1–21. doi: 10.1111/j.1468-0017.2006.00297.x

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hong, J.-W., Wang, Y., and Lanz, P. (2020). Why is artificial intelligence blamed more? Analysis of faulting artificial intelligence for self-driving car accidents in experimental settings. *Int. J. Hum. Comput. Interact.* 36, 1768–1774. doi: 10.1080/10447318.2020.1785693

Jelinski, L., Etzrodt, K., and Engesser, S. (2021). Undifferentiated optimism and scandalized accidents: the media coverage of autonomous driving in Germany. *J. Sci. Commun.* 20, 1–25. doi: 10.22323/2.20040202

Kallioinen, N., Pershina, M., Zeiser, J., Nosrat Nezami, F., Pipa, G., Stephan, A., et al. (2019). Moral judgements on the actions of self-driving cars and human drivers in dilemma situations from different perspectives. *Front. Psychol.* 10, 1–15. doi: 10.3389/fpsyg.2019.02415

Kant, I. (2011). *Groundwork of the Metaphysics of Morals. A German-English Edition*. New York, NY: Cambridge University Press.

Keeling, G. (2020). Why trolley problems matter for the ethics of automated vehicles. *Sci. Eng. Ethics* 26, 293–307. doi: 10.1007/s11948-019-00096-1

Komar, G. F., Mieth, L., Buchner, A., and Bell, R. (2022). Animacy enhances recollection but not familiarity: convergent evidence from the remember-know-guess paradigm and the process-dissociation procedure. *Mem. Cognit.*, 1–17. doi: 10.3758/s13421-022-01339-6

Koopman, P., and Wagner, M. (2017). Autonomous vehicle safety: an interdisciplinary challenge. *IEEE Intell. Transp. Syst. Mag.* 9, 90–96. doi: 10.1109/MITS.2016.2583491

Lee, J.-G., Kim, K. J., Lee, S., and Shin, D.-H. (2015). Can autonomous vehicles be safe and trustworthy? Effects of appearance and autonomy of unmanned driving systems. *Int. J. Hum. Comput. Interact.* 31, 682–691. doi: 10.1080/10447318.2015.1070547

Leiner, D. J. (2019). *SoSci Survey (Computer Software)*. Munich: SoSci Survey Gmb H.

Levin, S., and Wong, J. C. (2018). Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian. *The Guardian*. Available at: https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe. (Accessed October 15, 2021).

Li, J., Zhao, X., Cho, M.-J., Ju, W., and Malle, B. F. (2016). From trolley to autonomous vehicle: perceptions of responsibility and moral norms in traffic accidents with self-driving cars (SAE technical paper 2016-01-0164). SAE International, 1–8.

Lin, P. (2016). "Why ethics matters for autonomous cars" in *Autonomous Driving*. eds. M. Maurer, J. C. Gerdes, B. Lenz and H. Winner (Berlin Heidelberg: Springer), 69–85.

Liu, P., Wang, L., and Vincent, C. (2020). Self-driving vehicles against human drivers: equal safety is far from enough. *J. Exp. Psychol. Appl.* 26, 692–704. doi: 10.1037/xap0000267

Malle, B. F., Magar, S. T., and Scheutz, M. (2019). "AI in the sky: how people morally evaluate human and machine decisions in a lethal strike dilemma" in *Robotics and Well-being*. eds. M. I. Aldinhas Ferreira, J. Silva Sequeira, G. Singh Virk, M. O. Tokhi and E. E. Kader (Cham: Springer), 111–133.

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., and Cusimano, C. (2015). "Sacrifice one for the good of many? People apply different moral norms to human and robot agents". In *HRI'15 Proceedings of the 2015 ACM/IEEE International Conference on Human-Robot Interaction*. Association for Computing Machinery. pp. 117–124.

Malle, B. F., Scheutz, M., Forlizzi, J., and Voiklis, J. (2016). "Which robot am I thinking about? The impact of action and appearance on people's evaluations of a moral robot". In *HRI'16 the Eleventh ACM/IEEE International Conference on Human Robot Interaction*. Institute of Electrical and Electronics Engineers, Inc. p. 125–132.

Mayer, M. M. (2022). Autonome Fahrzeuge und moralische Dilemmas: Einflüsse der Perspektive, sozialer Erwünschtheit und der Handelnden. [dissertation]. [Düsseldorf]: Heinrich Heine University Düsseldorf.

Mayer, M. M., Bell, R., and Buchner, A. (2021). Self-protective and self-sacrificing preferences of pedestrians and passengers in moral dilemmas involving autonomous vehicles. *PLoS One* 16:e0261673. doi: 10.1371/journal.pone.0261673

Meinhardt, M. J., Bell, R., Buchner, A., and Röer, J. P. (2018). Adaptive memory: is the animacy effect on memory due to emotional arousal? *Psychon. Bull. Rev.* 25, 1399–1404. doi: 10.3758/s13423-018-1485-y

Mill, J. S. (2010). *Utilitarianism/Der Utilitarismus*. Stuttgart: Reclam.

Nairne, J. S., Van Arsdall, J. E., and Cogdill, M. (2017). Remembering the living: episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Nairne, J. S., Van Arsdall, J. E., Pandeirada, J. N. S., Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory: the mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

National Highway Traffic Safety Administration (2008). *National Motor Vehicle Crash Causation Survey—Report to Congress (DOT HS 811 059)*. Washington, DC: National Highway Traffic Safety Administration.

National Transportation Safety Board (2017). *Collision Between a Car Operating With Automated Vehicle Control Systems and a Tractor-semitrailer Truck Near Williston*. Washington, DC: National Transportation Safety Board.

National Transportation Safety Board (2019). *Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian, Tempe, Arizona, march 18, 2018 (NTSB/HAR-19/03, product no. PB2019-101402)*. Washington, DC: National Transportation Safety Board.

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

Niu, D., Terken, J., and Eggen, B. (2018). Anthropomorphizing information to enhance trust in autonomous vehicles. *Hum. Factors Ergon. Manuf. Service Industries* 28, 352–359. doi: 10.1002/hfm.20745

Nyholm, S. (2018). The ethics of crashes with self-driving cars: a roadmap, I. *Philos. Compass* 13, 1–10. doi: 10.1111/phc3.12507

O'Brien, R. G., and Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychol. Bull.* 97, 316–333. doi: 10.1037/0033-2909.97.2.316

Popp, E. Y., and Serra, M. J. (2016). Adaptive memory: Animacy enhances free recall but impairs cued recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 186–201. doi: 10.1037/xlm0000174

Popp, E. Y., and Serra, M. J. (2018). The animacy advantage for free-recall performance is not attributable to greater mental arousal. *Memory* 26, 89–95. doi: 10.1080/09658211.2017.1326507

Sachdeva, S., Iliev, R., Ekhtiari, H., and Dehghani, M. (2015). The role of self-sacrifice in moral dilemmas. *PLoS One* 10:e0127409. doi: 10.1371/journal.pone.0127409

SAE International (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (SAE ground vehicle standard J3016_202104). Available at: https://www.sae.org/standards/content/j3016_202104/. (Accessed March 11, 2022).

Scheutz, M., and Malle, B. F. (2021). "May machines take lives to save lives? Human perceptions of autonomous robots (with the capacity to kill)" in *Lethal Autonomous Weapons: Re-examining the Law and Ethics of Robotic Warfare*. eds. J. Galliott, D. Mac Intosh and J. D. Ohlin (New York, NY: Oxford University Press), 89–101.

Shariff, A., Bonnefon, J.-F., and Rahwan, I. (2017). Psychological roadblocks to the adoption of self-driving vehicles. *Nat. Hum. Behav.* 1, 694–696. doi: 10.1038/s41562-017-0202-6

Smith, A., and Anderson, M. (2017). *Automation in Everyday Life*. Washington, DC: Pew Research Centers.

Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., and Pavone, M. (2014). "Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: a case study in Singapore" in *Road Vehicle Automation*. eds. G. Meyer and S. Beiker (Cham: Springer), 229–245.

Tennant, C., Stares, S., and Howard, S. (2019). Public discomfort at the prospect of autonomous vehicles: building on previous surveys to measure attitudes in 11 countries. *Transp. Res. Part F Traffic Psychol. Behav.* 64, 98–118. doi: 10.1016/j.trf.2019.04.017

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *Monist* 59, 204–217. doi: 10.5840/monist197659224

Thomson, J. J. (1985). The trolley problem. *Yale Law J.* 94, 1395–1415. doi: 10.2307/796133

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Cogdill, M. (2017). A categorical recall strategy does not explain animacy effects in episodic memory. *Q. J. Exp. Psychol.* 70, 761–771. doi: 10.1080/17470218.2016.1159707

Wakabayashi, D. (2018). Self-driving Uber car kills pedestrian in Arizona, where robots roam. *The New York Times* 19:2018. Available at: https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html (Accessed October 15, 2021).

Wang, J., Zhang, L., Huang, Y., and Zhao, J. (2020). Safety of autonomous vehicles. *J. Adv. Transp.* 2020, 1–13. doi: 10.1155/2020/8867757

Waytz, A., Heafner, J., and Epley, N. (2014). The mind in the machine: anthropomorphism increases trust in an autonomous vehicle. *J. Exp. Soc. Psychol.* 52, 113–117. doi: 10.1016/j.jesp.2014.01.005

Winkler, M., Mehl, R., Erander, H., Sule, S., Buvat, J., KVJ, S., et al. (2019). The autonomous car: A consumer perspective: Capgemini Research Institute. Available at: https://www.capgemini.com/wp-content/uploads/2019/05/30min-–-Report.pdf (Accessed October 19, 2021).

Wolkenstein, A. (2018). What has the trolley dilemma ever done for us (and what will it do in the future)? On some recent debates about the ethics of self-driving cars. *Ethics Inf. Technol.* 20, 163–173. doi: 10.1007/s10676-018-9456-6

Yadron, D., and Tynan, D. (2016). Tesla driver dies in first fatal crash while using autopilot mode. *The Guardian*. Available at: https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk (Accessed October 15, 2021).

Young, A. D., and Monroe, A. E. (2019). Autonomous morals: inferences of mind predict acceptance of AI behavior in sacrificial moral dilemmas. *J. Exp. Soc. Psychol.* 85:103870. doi: 10.1016/j.jesp.2019.103870

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 January 2023
DOI 10.3389/fpsyg.2022.1017685

# Facilitating animacy perception by manipulating stimuli exposure time

Toshiki Saito[1,2]*, Kosuke Motoki[3], Rui Nouchi[4] and Motoaki Sugiura[4]

[1]School of Fundamental Science and Engineering, Waseda University, Tokyo, Japan, [2]Japan Society for the Promotion of Science, Tokyo, Japan, [3]Department of Management, The University of Tokyo, Tokyo, Japan, [4]Institute for Development and Ageing and Cancer, Tohoku University, Sendai, Japan

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University,
United States

REVIEWED BY
Jianping Huang,
Soochow University,
China
Silvia Guerra,
University of Padua,
Italy

*CORRESPONDENCE
Toshiki Saito
✉ t.saito18@kurenai.waseda.jp

SPECIALTY SECTION
This article was submitted to Cognition,
a section of the journal
Frontiers in Psychology

RECEIVED 12 August 2022
ACCEPTED 15 December 2022
PUBLISHED 12 January 2023

CITATION
Saito T, Motoki K, Nouchi R and Sugiura M (2023) Facilitating animacy perception by manipulating stimuli exposure time.
*Front. Psychol.* 13:1017685.
doi: 10.3389/fpsyg.2022.1017685

Animacy perception—discriminating between animate and inanimate visual stimuli—is the basis for engaging in social cognition and for our survival (e.g., avoiding potential danger). Previous studies indicate that factors in a target, such as the features or motion of a target, enhance animacy perception. However, factors in a perceiver, such as the visual attention of a perceiver to a target, have received little attention from researchers. Research on judgment, decision-making, and neuroeconomics indicates the active role of visual attention in constructing decisions. This study examined the role of visual attention in the perception of animacy by manipulating the exposure time of targets. Among Studies 1a to 1c conducted in this study, participants saw two face illustrations alternately; one of the faces was shown to be longer than the other. The participants chose the face that they considered more animated and rounder. Consequently, longer exposure time toward targets facilitated animacy perception and preference rather than the perception of roundness. Furthermore, preregistered Study 2 examined the underlying mechanisms. The results suggest that mere exposure, rather than orienting behavior, might play a vital role in the perception of animacy. Thus, in the reverse relationship between visual attention and animacy perception, animate objects capture attention—attention results in the perception of animacy.

KEYWORDS
**animacy perception, visual attention, gaze manipulation, mere exposure, orienting behavior**

## 1. Introduction

Animacy perception, which distinguishes animate from inanimate visual stimuli (Rutherford and Kuhlmeier, 2013), is a necessary component of social interaction. Evidence shows that such perceptions emerge even in infancy (Leslie, 1982; Gergely et al., 1995; Rochat et al., 1997) but are disrupted by developmental disorders (Rutherford et al., 2006) and amygdala damage (Heberlein and Adolphs, 2004).

Previous research on factors driving the perception of animacy mainly focused on the properties of target stimuli, such as human-like appearances (e.g., a face) and motion (e.g., interactive motion between two geometric shapes). For instance, people consider an object comparatively more animate when the object has unique human features such as eyes and mouth (Looser and Wheatley, 2010), intelligence (Bartneck et al., 2009), and facial expressions of happiness (Bowling and Banissy, 2017; Krumhuber et al., 2019; Saito et al., 2022). Furthermore, individuals consider moving objects more animate when the motion seems to have specific goals, such as chasing and helping (Heider and Simmel, 1944; Rochat et al., 1997; Castelli et al., 2000; Scholl and Tremoulet, 2000; Tremoulet and Feldman, 2000; Kuhlmeier et al., 2003).

Although most previous research has shown that factors in a target (e.g., human-like features of targets) play crucial roles in animacy perception, these factors do not necessarily facilitate it. According to the uncanny valley theory (Mori et al., 2012), inanimate objects (e.g., robots) resemblance to humans increases the perception of animacy. However, when the resemblance reaches a certain point, it provokes uncanny or strange feelings and hinders the perception. The uncanny valley theory suggests the importance of focusing not only on factors in a target but also on factors in a perceiver. However, minimal extant literature focuses on the factors (e.g., knowledge and mental state of participants) in animacy perception. For example, beliefs about the origin of moving objects (i.e., humans or robots; Cross et al., 2016) and the state of participants (i.e., loneliness) affect animacy perception (Powers et al., 2014). Thus, it is also essential to focus on factors affecting animacy perception in perceivers.

Notably, attention, which can be counted as a factor in perceivers, has a critical relationship with animacy perception and might be a causal effect in animacy perception. Previous studies show that animate objects capture attention (Pratt et al., 2010; Yang et al., 2012; Jackson and Calvillo, 2013; Calvillo and Jackson, 2014). For example, when individuals are tasked with finding a category exemplar and are unexpectedly exposed to either an animate or inanimate object, they are more likely to notice the animate object (Calvillo and Jackson, 2014). Thus, the authors concluded that these findings reflect that detecting animate objects is vital in ancestral hunter-gatherer environments and is consistent with the animate-monitoring hypothesis (New et al., 2007). As mentioned above, animate objects attract attention. However, is there also a reversal relationship between them? Specifically, does attracting attention lead to animacy perception? This prediction could be the case considering recent research on judgment and decision-making.

The growing body of research on judgment, decision-making, and neuroeconomics highlights the crucial role of visual attention in decision-making (Armel et al., 2008; Krajbich et al., 2010; Glöckner and Herbold, 2011; Krajbich and Rangel, 2011; Krajbich et al., 2012; Orquin and Loose, 2013; Cavanagh et al., 2014; Thomas et al., 2019). In particular, the attentional drift-diffusion model (aDDM), proposed by Krajbich and his peers, incorporates the role of visual attention in traditional decision-making models (i.e., the drift-diffusion model; Krajbich et al., 2010; Krajbich and Rangel, 2011; Krajbich et al., 2012; Krajbich, 2019). The aDDM is a decision-making model assuming that the evidence of an item for reaching a decision is amplified when the item receives more attention. Notably, assuming that visual attention modulates the accumulation of evidence to reach a threshold to decide, decision times and choices can accurately be predicted (Krajbich et al., 2010). Previous neural studies present supportive evidence that there was neural activity related to fixation-dependent value coding but did not examine the validity of aDDM (Lim et al., 2011; McGinty et al., 2016). Furthermore, numerous behavioral studies have shown that a longer gaze duration toward one option results in a higher choice probability for that option (Shimojo et al., 2003; Armel et al., 2008; Krajbich et al., 2010, 2012; Saito et al., 2017, 2020; Thomas et al., 2019; Motoki et al., 2021). Moreover, behavioral studies indicate the causal role of attention in decision-making by manipulating the gaze toward options and that the probabilities of choices have changed (Shimojo et al., 2003; Armel et al., 2008; Pärnamets et al., 2015). As mentioned above, visual attention plays a crucial role in decision-making. By employing this perspective, we sought to specify factors in a perceiver affecting animacy perception in the current study.

Gaze manipulation does not always bias decision-making (Shimojo et al., 2003). Though, it is reported that there is a consistent effect on simple perceptual choice (Tavares et al., 2017). According to Shimojo et al. (2003), gaze manipulation can influence subjective (e.g., preference) rather than objective judgments. In other words, gaze manipulation is likely to influence higher-level cognition (e.g., preference) rather than low-level perception (e.g., morphological perception). Given that preference for targets contributes to an uncanny valley feeling (Wang and Rochat, 2017), animacy perception may be influenced by gaze manipulation through a preference for targets. This study directly tests this hypothesis where gaze manipulation influences the animacy perception.

Exposure duration (i.e., mere exposure effect; Zajonc, 1968) and gaze shifting (i.e., gaze orienting) may be potential mechanisms influencing the role of visual attention in animacy perception. The account of exposure duration is based on the mere exposure effect (Zajonc, 1968). Specifically, the more people look at a stimulus, the more they like it. It has also been assumed that gaze orienting is a precursor to higher-level cognition (e.g., preferences; Shimojo et al., 2003; Simion and Shimojo, 2006; Simion and Shimojo, 2007). More prolonged exposure durations with orientation (i.e., gaze shifting) can induce a preference shift. In contrast, longer exposure durations without orientation do not result in a preference shift (Shimojo et al., 2003). We further elucidated the potential underlying mechanisms by manipulating exposure duration and gaze orientation.

This study examines whether stimuli exposure time influences the perception of animacy. In particular, we investigated whether the manipulation influenced high-level perceptions, animacy perception (Study 1a), and preference (Study 1b). Additionally, we examined the effect on roundness judgment (Study 1c), which we considered low-level perception. We expected high-level cognition (i.e., animacy and preference), rather than low-level perception, would be biased by gaze manipulation. Furthermore, Study 2 examined the underlying mechanisms by separating the factors of gaze manipulation into exposure duration and arbitrary eye movements. As we mentioned above, it is reported that arbitrary eye movement is necessary for biasing high-level cognition (Shimojo et al., 2003). However, contradicting findings reported that extended exposure duration, regardless of gaze orientation, biased decision-making (e.g., Nittono and Wada, 2009; Bird et al., 2012). Thus, we examined which factor of gaze manipulation, exposure duration, or arbitrary eye movements influence animacy perception in Study 2.

## 2. Study 1a to 1c

Study 1a examined the effects of gaze manipulation on animacy perception. The participants viewed two facial images with artificial features and then chose the image perceived as more animate. While viewing the images, participants' eye movements were manipulated using the paradigm of a previous study (Shimojo et al., 2003). Study 1b was designed to replicate the effect of gaze manipulation on preference judgment (Shimojo et al., 2003) in the current experimental procedure. The procedure was similar, except that it made participants choose their preferred facial images. Study 1c was designed to confirm the specificity of gaze manipulation for both preference and animacy perception. The procedure was almost the same, except that it required participants to choose a rounder facial image.

### 2.1. Methods

### 2.1.1. Participants

To the best of our knowledge, because no prior study has examined the effect of exposure time on animacy perception, we did not formally calculate the sample size for Studies 1a to 1c. We recruited university students who participated in each study during the 1st wave of the recruitment period. Finally, 43 participants for Study 1a (11 women, 32 men; mean age, 20.78; SD of age, 1.38), 61 participants for Study 1b (20 women, 41 men; mean age, 21.13; SD of age, 2.75), and 29 participants for Study 1c (12 women, 17 men; mean age, 21.41; SD of age, 1.37) were selected. We considered those sample sizes (i.e., 29–61) almost sufficient to detect the effect given previous studies' sample sizes ranged from 10 to 100 (Shimojo et al., 2003; Armel et al., 2008). The participants were all university students recruited *via* a university bulletin board and mailing list. After completing the study, participants received a small monetary compensation for their participation. This study was approved by the ethics committees of Tohoku University (Number: UMIN000025712) and Waseda University [Number 2019-357(1)] and conducted per the Declaration of Helsinki. For each study, the participants gave their free and informed consent.

### 2.1.2. Stimuli

In this study, we used 40 pairs of facial images (20 female and 20 male face pairs). To create these images, we selected 45 male and 45 female faces from the Chicago Face Database (Ma et al., 2015). All facial images displayed no emotional expression (i.e., neutral expression). The images depicted real human faces and might cause a ceiling effect on animacy perception that prevented the effects of gaze manipulation. Therefore, we modified these images to add artificial features using non-photorealistic rendering methods (Rosin and Lai, 2015). This method produces realistic cartoons from real images of the same identity (Figure 1). The images were resized to a uniform width of 600 pixels and height of 450 pixels.

We further conducted an online pre-experiment to manipulate the attractiveness of the images using Qualtrics. We recruited 40 participants *via* Lancers[1] and asked them to rate the attractiveness

FIGURE 1
Examples of modified face images used in the studies. Adapted with permission from Chicago Face Database, available at https://www.chicagofaces.org/ (Ma et al., 2015).

---

[1] https://www.lancers.jp

of the images on a 7-point Likert scale (1 = very unattractive to 7 = very attractive). Based on these ratings, we created 40 pairs of facial images. Stimulus codes for the exact stimuli employed are available in the online supplemental material.[2] The average ratings of the faces in a pair were matched such that the difference in the average rating in each pair was <0.10 points. The average rating for all faces was 3.12 (SD = 0.49). The faces in a pair were also matched in terms of sex. There was an equal number of face pairs in each sex (20 male and 20 female face pairs).

### 2.1.3. Procedure

We used similar experimental procedures and conducted the experiments almost concurrently. Based on a previous study (Shimojo et al., 2003), we manipulated stimuli exposure time to participants while perceiving a pair of faces (Figure 2A).

Participants completed the experiment individually on a computer (display resolution 14-inch, 1920 × 1080). The distance between the participant's eyes and the display was approximately 60 cm. After showing the fixation cross for 500 ms, we presented each face six times to the participants. Faces alternated between the left and right halves of the screen. Therefore, participants had to shift their gaze toward the visible face on the screen. The presentation duration for each face in a pair was different, 900 ms for one face and 300 ms for another face. At one trial, one face was shown for 5400 ms (900 ms × 6 times) and another for 1800 ms (300 ms × 6 times). These durations were identical to those of the previous study (Shimojo et al., 2003). Faces that were shown longer than other faces were counterbalanced across the participants. After viewing a pair of faces, participants chose a face in which they perceived animacy more (Study 1a), preferred more (Study 1b), or perceived rounder (Study 1c) by pressing the corresponding keys. For instance, the "f" key for the left-sided face and the "j" key for the right-sided face. The reaction time was not constrained, and the order of face pair presentation was randomized across trials. The total number of trials was set to 40. Before the experiment, we explained the procedure to participants and confirmed their understanding of the instruction by asking them.

### 2.1.4. Statistical analysis

Through Studies 1a to 1c, we used mixed logistic models to predict the choice of the target presented on the left side (1: left-sided target, 0: right-sided target), with the left-sided target shown for a long or short duration (1: shown longer, 0: shown shorter) as a fixed effect, and participants and pairs of stimuli included as a

---

[2] https://osf.io/cr4yx/?view_only=633225a44c9f455993688e2c96ea382c

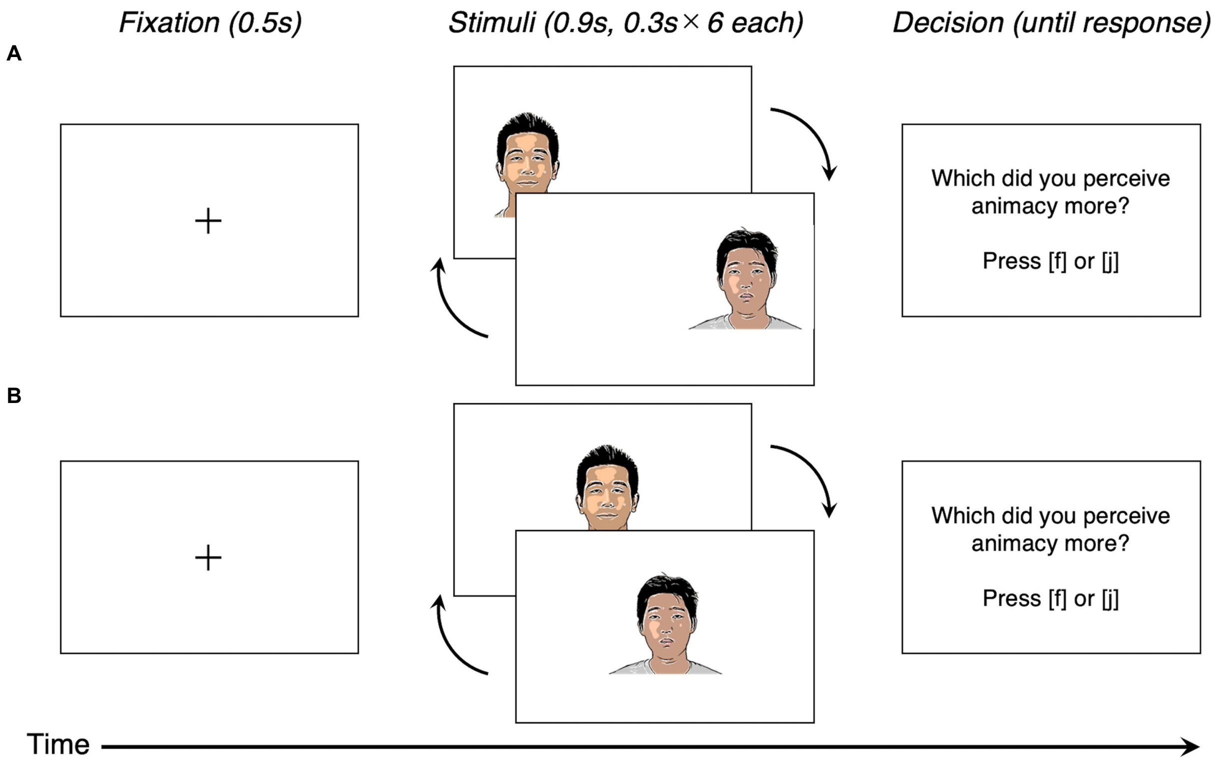


FIGURE 2
Example of task flow in Studies 1a to 1c and Study 2. Note. In Studies 1b, 1c, and 2, the instruction for decision-making was changed depending on the conditions [which did you prefer?] for Study 1b; and which did you perceive as rounder?; for Studies 1c and 2. **(A)** the arbitrary eye movement condition in Studies 1a-1c, 2 and **(B)** the fixed eye movement condition in Study 2. Adapted with permission from Chicago Face Database, available at https://www.chicagofaces.org/ (Ma et al., 2015).

random slope and a random intercept. All analyses were conducted using the lme4 package (Bates et al., 2014) in the R software (R Core Team, 2021). Regarding the analysis in Study 1c, we excluded 12 trials in which the stimuli were not presented for the intended duration owing to technical issues. In conclusion, we analyzed the data of 28 trials from each participant in Study 1c. The data analyzed in this study were made available at the Open Science Framework.[3]

## 2.2. Results

### 2.2.1. Study 1a (Animacy judgment)

The results of the analysis showed that participants tended to choose longer-shown faces as more animated faces in Study 1a (53.89, 95% CI [51.89–55.89]; $b = 0.34$, $z = 3.19$, $p < 0.001$). This result suggests that gaze bias influenced perceptions of animacy. The likelihoods of longer-shown stimuli chosen through Studies 1a to 1c are visualized in Figure 3. Table 1 further shows the details of the results.

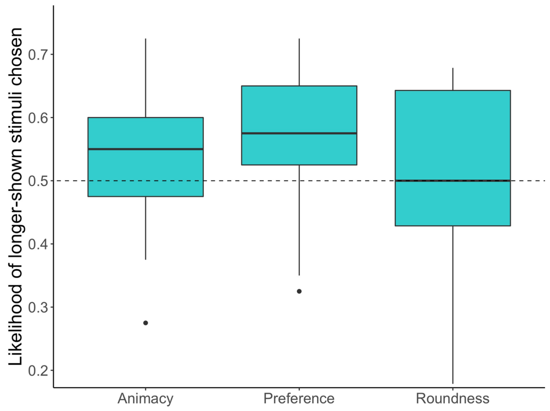


FIGURE 3
The box plots of the likelihood of the longer-shown stimuli chosen by participants across Studies 1a to 1c. The dashed line represents the chance level (50%).

### 2.2.2. Study 1b (preference judgment)

The results of the analysis revealed that participants preferred the faces that were shown longer (57.03, 95% CI [54.65–59.39]; $b = 0.53$, $z = 4.21$, $p < 0.001$). This indicates that we successfully replicated the effect of eye movement on preference judgment (Shimojo et al., 2003; Krajbich et al., 2010; Saito et al., 2017).

### 2.2.3. Study 1c (roundness judgment)

The results of the analysis indicated that participants did not tend to choose longer-shown faces as rounder faces in Study 1c (49.88, 95% CI [46.38–53.37]; $b = -0.02$, $z = -0.08$, $p = 0.98$). This result suggests that gaze bias specifically influenced both preferences and animacy perception rather than morphological perception (i.e., roundness judgment).

## 2.3. Discussion

Through Studies 1a to 1c, we observed that gaze manipulation influences animacy and preference judgments, not roundness judgments. These findings suggest the specificity of the effect of gaze manipulation on animacy and preference perceptions and that these perceptions might be affected by gaze manipulation through the exact mechanism. However, regarding the mechanism, it is unclear what aspect of gaze manipulation we used affected the perceptions because we manipulated the presentation duration (i.e., mere exposure) and arbitrary eye movements (i.e., orienting behavior). Furthermore, we did not directly compare the effects of gaze manipulation on animacy and roundness judgment. Therefore, Study 2 was designed to address these questions.

# 3. Study 2

In Study 2, we sought to solve the issues mentioned above by directly comparing the effects of (1) present duration (i.e., mere exposure) and arbitrary eye movements (i.e., orienting behavior) and (2) animacy and roundness judgment. To specify the factors of gaze manipulation on animacy perception, we used a paradigm in which participants' eye movements were fixed (Shimojo et al., 2003).

**TABLE 1** Fixed effects from the GLMM analyses through Studies 1a to 1c.

| Study | Predictor | $\beta$ | SE | $z$-value | Value of $p$ | OR | 95% (OR) |
|---|---|---|---|---|---|---|---|
| Study 1a (Animacy) | Intercept | −0.28 | 0.11 | −2.57 | 0.010 | 0.76 | [0.61, 0.94] |
| | Presentation Duration | 0.34 | 0.11 | 3.19 | 0.001 | 1.40 | [1.14, 1.72] |
| Study 1b (Preference) | Intercept | −0.33 | 0.14 | −2.41 | 0.016 | 0.72 | [0.55, 0.94] |
| | Presentation Duration | 0.53 | 0.13 | 4.21 | 0.001 | 1.71 | [1.33, 2.19] |
| Study 1c (Roundness) | Intercept | 0.19 | 0.15 | 1.28 | 0.202 | 1.20 | [0.91, 1.60] |
| | Presentation Duration | −0.02 | 0.24 | −0.08 | 0.936 | 0.98 | [0.61, 1.58] |

OR, odds ratio.

---

[3] https://osf.io/cr4yx/

## 3.1. Method

### 3.1.1. Experimental design

This study included two independent variables: the type of judgment (two levels: animacy and roundness) and gaze manipulation (two levels: arbitrary eye movement and fixed eye movement). These variables were between-participant factors. The dependent variable was the choice of stimulus.

### 3.1.2. Participants and stimuli

We conducted a simulation-based power analysis using the SIMR package (Green and MacLeod, 2016) in R and the data from Study 1a to estimate the ideal sample size. This analysis determined the expected power to secure the fixed effect of gaze manipulation for various sample sizes. The results indicated the need for a sample size of 169 to achieve over 80% at an alpha level of 0.05. Considering that the average dropout rate of a typical web experiment is approximately 30% (Musch and Reips, 2000; Zhou and Fishbach, 2016), we determined the sample size to be 220 participants (55 participants for each condition). We recruited participants using Lancers. A total of 221 participants were recruited for the study. After excluding participants who failed the attention check, data from 205 participants (63 women, 136 men, 6 preferred not to disclose; mean age, 41.98; SD of age, 8.66) were analyzed. The participants received a small monetary compensation for their participation. This study was approved by AsPredicted.org.[4] Further, we used the same stimuli as in Studies 1a to 1c.

### 3.1.3. Procedure

The procedure was almost identical to those of the previous studies, except as noted in the following text. We conducted the study online through Qualtrics[5] because it was challenging to experiment in person due to the coronavirus disease 2019 pandemic at that time (October 18–25, 2021). At the beginning of each study, participants answered a question designed to check whether they read instructions as an attention check (Oppenheimer et al., 2009). In particular, participants had to ignore the standard response format and instead provide a confirmation that they had read the instruction in the question. Then, participants were randomly assigned to one of four conditions (two types of judgment: animacy, roundness × 2, gaze manipulation: arbitrary eye movement, and fixed eye movement). In Study 2, we established a fixed eye movement condition, where the stimuli face alternated at the center of the screen (Figure 2B). Thus, the participants did not have to shift their gaze toward the visible face on the screen. After viewing each pair of faces, participants chose the face they perceived as having more animacy (animacy condition) or rounder (roundness condition) by pressing the corresponding keys. For instance, the "f" key for the face presented at last and the "j" key for the face presented before the last. In conclusion, 104 participants were assigned to the animacy judgment condition (53 in arbitrary eye movement, 51 in fixed eye movement), and 101 participants were further assigned to the roundness judgment condition (52 participants in arbitrary eye movement, 49 participants in fixed eye movement).

## 3.2. Statistical analysis

Study 2 used the preregistered analysis, which was a linear mixed model predicting the choice of one target (arbitrary eye movement condition: 1 = left-sided target, 0 = right-sided target; fixed condition: 1 = the last-presented target, 0 = before the last-presented target), with the target was shown for a long or short duration (presentation duration: 1 = shown longer, 0 = shown shorter), gaze manipulation (1 = arbitrary eye movement, 0 = fixed eye movement), types of judgment (1 = animacy, 0 = morphological perception). Further, the interactions were included as fixed effects, and participants and pairs of stimuli were included as random effects. We used the performance package (Lüdecke et al., 2021) in the R software to investigate the variance inflation factors (VIFs) and overdispersion.

In addition to the preregistered analysis, we conducted similar analyses to the previous three studies as exploratory analyses for analytical consistency across studies. For each condition, we conducted an analysis that was a mixed logistic model predicting the choice of one target (arbitrary eye movement condition: 1 = left-sided target, 0 = right-sided target; fixed condition: 1 = the last-presented target, 0 = before the last presented target), with the target shown for a long or short duration (1 = shown longer, 0 = shown shorter), as a fixed effect. Further, participants and pairs of stimuli were included as a random slope and random intercept, respectively.

## 3.3. Results

Figure 4 shows the likelihood of choosing longer-shown faces. Regarding the preregistered analysis, we confirmed that multicollinearity was not a problem by inspecting the VIFs (VIFs < 3.91). Further, overdispersion was not a problem in the overdispersion test ($\chi^2 = 7157.48$, $p = 1.00$). The result from the preregistered analysis showed neither significant effects of presentation duration, gaze manipulation, and types of judgment nor those interactions (Table 2).

Although we did not observe any significant results from the registered analysis, we conducted a mixed logistic model to predict the choice of one target for each condition, as in previous studies (Table 3).

In the arbitrary eye movement condition, we observed that participants tended to choose longer-shown faces when choosing more animate faces (54.86, 95% CI [52.71–56.99]; $b = 0.42$, $z = 2.77$,

---

[4] https://aspredicted.org/492_915

[5] https://www.qualtrics.com/jp/

$p = 0.01$) than when choosing rounder faces (50.48, 95% CI [48.31–52.65]; $b = 0.05$, $z = 1.14$, $p = 0.26$). In the fixed eye movement condition, we also observed that participants tended to choose longer-shown faces when choosing more animate faces (53.43, 95% CI [51.24–55.61]; $b = 0.29$, $z = 2.28$, $p = 0.02$) but not when choosing rounder faces (50.71, 95% CI [48.48–52.95]; $b = 0.23$, $z = 0.61$, $p = 0.54$).

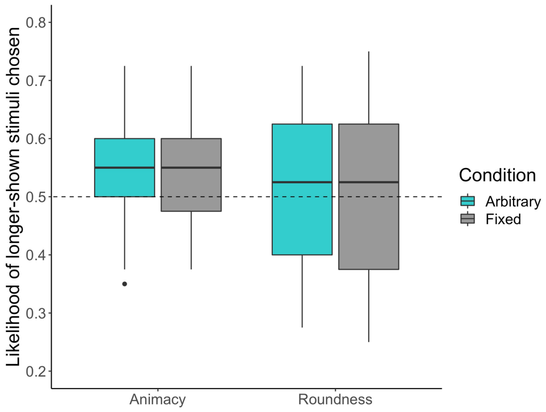


FIGURE 4
The box plots of the likelihood of the longer-shown stimuli chosen by participants in Study 2. The dashed line represents the chance level (50%).

## 3.4. Discussion

In Study 2, the preregistered analysis showed neither significant effects nor interactions between the experimental conditions. Therefore, we failed to elucidate the factors of gaze manipulation (i.e., mere exposure and orienting behavior) that influence animacy perception. However, subsequent exploratory analyses were consistent with Studies 1a to 1c, showing that gaze manipulation in arbitrary and fixed eye movement conditions influenced only animacy perception rather than the perception of roundness. These results suggest that mere exposure may be critical in facilitating animacy perception.

# 4. General discussions

Factors in a perceiver have not received sufficient attention regarding the factors that drive animacy perception. We tested whether one of the primary factors, visual attention toward stimuli, affects animacy perception. Across Studies one and two, the participants felt that cartoon faces were more animated when

TABLE 2 Fixed effects from the registered analysis predicting the choice in Study 2.

| Predictor | $\beta$ | SE | $z$-value | Value of $p$ | OR | 95% (OR) |
|---|---|---|---|---|---|---|
| Intercept | −0.13 | 0.19 | −0.72 | 0.473 | 0.87 | [0.61, 1.26] |
| Presentation duration | 0.30 | 0.28 | 1.08 | 0.282 | 1.35 | [0.78, 2.35] |
| Gaze manipulation | −0.12 | 0.26 | −0.46 | 0.643 | 0.89 | [0.53, 1.49] |
| Type of choice | 0.39 | 0.27 | 1.45 | 0.148 | 1.47 | [0.87, 2.48] |
| Presentation duration × Gaze manipulation | 0.14 | 0.40 | 0.34 | 0.732 | 1.15 | [0.53, 2.50] |
| Presentation duration × Type of choice | −0.12 | 0.40 | −0.30 | 0.762 | 0.89 | [0.40, 1.95] |
| Gaze manipulation × Type of choice | −0.35 | 0.38 | −0.93 | 0.353 | 0.70 | [0.34, 1.48] |
| Presentation duration × Gaze manipulation × Type of choice | −0.27 | 0.57 | −0.47 | 0.641 | 0.77 | [0.25, 2.34] |

OR, odds ratio.

TABLE 3 Fixed effects from the additional GLMM analyses of each condition in Study 2.

| Type | Condition | Predictor | $\beta$ | SE | $z$-value | Value of $p$ | OR | 95% (OR) |
|---|---|---|---|---|---|---|---|---|
| Animacy | Arbitrary | Intercept | −0.25 | 0.11 | −2.35 | 0.019 | 0.78 | [0.63, 0.96] |
| | | Presentation duration | 0.42 | 0.15 | 2.82 | 0.005 | 1.52 | [1.14, 2.03] |
| | Fixed | Intercept | −0.12 | 0.08 | −1.50 | 0.133 | 0.88 | [0.75, 1.04] |
| | | Presentation duration | 0.28 | 0.13 | 2.23 | 0.026 | 1.33 | [1.04, 1.70] |
| Roundness | Arbitrary | Intercept | −0.29 | 0.33 | −0.88 | 0.379 | 0.75 | [0.40, 1.42] |
| | | Presentation duration | 0.03 | 0.52 | 0.06 | 0.950 | 1.03 | [0.38, 2.84] |
| | Fixed | Intercept | 0.25 | 0.23 | 1.12 | 0.264 | 1.29 | [0.83, 2.01] |
| | | Presentation duration | 0.21 | 0.35 | 0.59 | 0.558 | 1.23 | [0.62, 2.46] |

OR, odds ratio.

manipulating their gaze to look at the faces longer. This effect of biased exposure duration was also observed in preference judgments (Study 1b) rather than in lower-level perception (i.e., roundness judgments, Studies 1c and 2). Furthermore, in the preregistered online study (Study 2), it was found that arbitrary eye movements were not necessarily needed to increase animacy perception. However, exposure duration played a crucial role in influencing it.

Our results provide evidence that gazing behavior influences the perception of animacy. In this study, manipulating the exposure duration in arbitrary and fixed eye movement conditions facilitated animacy perception. This finding is inconsistent with the claim that gaze orienting is necessary to bias higher-level cognition, such as preference judgment (Shimojo et al., 2003). Instead, this finding is consistent with studies that show that gaze orienting is not a necessary condition for forming higher-level cognition but instead demonstrates that a mere exposure effect underlies biased higher-level cognition by gaze manipulation (Glaholt and Reingold, 2009; Nittono and Wada, 2009; Glaholt and Reingold, 2011; Bird et al., 2012).

There are several potential explanations for why the mere exposure effect derives animacy perceptions. First, along with aDDM (Krajbich et al., 2010), attention would have facilitated evidence of animacy and preference. Second, mere exposure may have changed several psychological constructs, as mere exposure increases familiarity and saliency (Montoya et al., 2017). Familiarity seemed to be a crucial construct in this study, given the uncanny valley theory, where unfamiliarity or strangeness hinders the perception of animacy (Mori et al., 2012). Moreover, an empirical study indicated that people attribute fundamental capacities of the mind, which is a concept strongly related to animacy, to preferred targets (Kozak et al., 2006). Examining the relationship between the mere exposure effect and animacy will likely be a pivotal issue for future work.

We have observed that the choice probabilities of longer-shown stimuli in the animacy condition were greater than chance. Nonetheless, it is important to note that in Study 2's preregistered analysis, we could not identify the gaze manipulation components that affected animacy perception. Given that the preregistered analysis did not reveal significant effects of judgment types and gaze manipulations on the choice probabilities, we cannot conclude that gaze manipulation uniquely affected animacy perception. Instead, we need to stress that the effect of gaze manipulation on animacy perception might be limited or relatively small. The degree and the uniqueness of the gaze manipulation effect on animacy perception should be further examined in future studies.

Future work would be needed to specify the relationship between visual attention and animacy perception in detail. Firstly, it is necessary to test whether the effect of gaze manipulation occurs for completely inanimate objects (e.g., simple geometrics). The facial stimuli in the present study seemed relatively animate; therefore, it is unclear whether the gaze manipulation effect can trigger animacy perception. Thus, testing whether exposure duration facilitates the perception of animacy, even when the targets are entirely inanimate, would be an interesting direction. It is crucial to inspect the underlying mechanisms biased by gaze manipulation directly influencing animacy perception. Notably, attentional bias results in changes in the target's characteristics, such as saliency, liking (Mrkva and Van Boven, 2020), and familiarity (Montoya et al., 2017). Therefore, future studies need to examine the psychological mechanisms that mediate the relationship between gaze manipulation and animacy perception.

This study tested whether exposure time plays a role in the perception of animacy. We found evidence that biased exposure time of targets facilitated both animacy perception and preference toward targets rather than lower-level perception (i.e., morphological judgment). The underlying mechanisms biased by gaze manipulation directly influencing animacy perception are not clear. However, our findings suggest that biasing visual attention toward targets facilitates animacy perception, possibly because mere exposure increases familiarity or preference.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/cr4yx/.

## Ethics statement

The studies involving human participants were reviewed and approved by the ethics committees of Tohoku University. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

TS: conceptualization, methodology, investigation, software, formal analysis, writing—original draft preparation. KM: conceptualization, methodology, investigation, writing—reviewing and editing. RN: methodology, writing—reviewing and editing. MS: supervision, writing—review and editing. All authors contributed to the article and approved the submitted version.

## Funding

This research was supported by a Grant-in-Aid for JSPS Fellows (21J01224; TS) from the Japan Society for the Promotion of Science and JSPS KAKENHI 19H01760 [Grant-in-Aid for Scientific Research (B)] and JSPS KAKENHI 19H05003 [Grant-in-Aid for Scientific Research on Innovative Areas (Research in a Proposed Research Area)] (RN).

## 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.

## Publisher's note

## References

Armel, K. C., Beaumel, A., and Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgm. Decis. Mak.* 3, 396–403.

Bartneck, C., Kanda, T., Mubin, O., and Al Mahmud, A. (2009). Does the design of a robot influence its animacy and perceived intelligence? *Adv. Robot.* 1, 195–204. doi: 10.1007/s12369-009-0013-7

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). lme4: linear mixed-effects models using Eigen and S4. *R Package Version*, 1, 1–23.

Bird, G. D., Lauwereyns, J., and Crawford, M. T. (2012). The role of eye movements in decision making and the prospect of exposure effects. *Vis. Res.* 60, 16–21. doi: 10.1016/j.visres.2012.02.014

Bowling, N. C., and Banissy, M. J. (2017). Emotion expression modulates perception of animacy from faces. *J. Exp. Soc. Psychol.* 71, 83–95. doi: 10.1016/j.jesp.2017.02.004

Calvillo, D. P., and Jackson, R. E. (2014). Animacy, perceptual load, and inattentional blindness. *Psychon. Bull. Rev.* 21, 670–675. doi: 10.3758/s13423-013-0543-8

Castelli, F., Happé, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–325. doi: 10.1006/nimg.2000.0612

Cavanagh, J. F., Wiecki, T. V., Kochar, A., and Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *J. Exp. Psychol. Gen.* 143, 1476–1488. doi: 10.1037/a0035813

Cross, E. S., Ramsey, R., Liepelt, R., Prinz, W., and Hamilton, A. F. D. C. (2016). The shaping of social perception by stimulus and knowledge cues to human animacy. *Philos. Trans. R. Soc. B Biol. Sci* 371:20150075. doi: 10.1098/rstb.2015.0075

Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193. doi: 10.1016/0010-0277(95)00661-H

Glaholt, M. G., and Reingold, E. M. (2009). Stimulus exposure and gaze bias: a further test of the gaze cascade model. *Atten. Percept. Psychophys.* 71, 445–450. doi: 10.3758/APP.71.3.445

Glaholt, M. G., and Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *J. Neurosci. Psychol. Econ.* 4, 125–146. doi: 10.1037/a0020692

Glöckner, A., and Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. *J. Behav. Decis. Mak.* 24, 71–98. doi: 10.1002/bdm.684

Green, P., and MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* 7, 493–498. doi: 10.1111/2041-210X.12504

Heberlein, A. S., and Adolphs, R. (2004). Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7487–7491. doi: 10.1073/pnas.0308220101

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259. doi: 10.2307/1416950

Jackson, R. E., and Calvillo, D. P. (2013). Evolutionary relevance facilitates visual information processing. *Evol. Psychol.* 11, 1011–1026. doi: 10.1177/147470491301100506

Kozak, M. N., Marsh, A. A., and Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *J. Pers. Soc. Psychol.* 90, 543–555. doi: 10.1037/0022-3514.90.4.543

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Curr. Opin. Psychol.* 29, 6–11. doi: 10.1016/j.copsyc.2018.10.008

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298. doi: 10.1038/nn.2635

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* 3:193. doi: 10.3389/fpsyg.2012.00193

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13852–13857. doi: 10.1073/pnas.1101328108

Krumhuber, E. G., Lai, Y. K., Rosin, P. L., and Hugenberg, K. (2019). When facial expressions do and do not signal minds: the role of face inversion, expression dynamism, and emotion type. *Emotion* 19, 746–750. doi: 10.1037/emo0000475

Kuhlmeier, V., Wynn, K., and Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychol. Sci.* 14, 402–408. doi: 10.1111/1467-9280.01454

Leslie, A. M. (1982). The perception of causality in infants. *Perception* 11, 173–186. doi: 10.1068/p110173

Lim, S. L., O'Doherty, J. P., and Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* 31, 13214–13223. doi: 10.1523/JNEUROSCI.1246-11.2011

Looser, C. E., and Wheatley, T. (2010). The tipping point of animacy: how, when, and where we perceive life in a face. *Psychol. Sci.* 21, 1854–1862. doi: 10.1177/0956797610388044

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., and Makowski, D. (2021). Performance: an R package for assessment, comparison and testing of statistical models. *J. Open Source Softw.* 6:3139. doi: 10.21105/joss.03139

Ma, D. S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: a free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135. doi: 10.3758/s13428-014-0532-5

McGinty, V. B., Rangel, A., and Newsome, W. T. (2016). Orbitofrontal cortex value signals depend on fixation location during free viewing. *Neuron* 90, 1299–1311. doi: 10.1016/j.neuron.2016.04.045

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., and Lauber, E. A. (2017). A re-examination of the mere exposure effect: the influence of repeated exposure on recognition, familiarity, and liking. *Psychol. Bull.* 143, 459–498. doi: 10.1037/bul0000085

Mori, M., MacDorman, K., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robot. Autom. Mag.* 19, 98–100. doi: 10.1109/mra.2012.2192811

Motoki, K., Saito, T., and Onuma, T. (2021). Eye-tracking research on sensory and consumer science: a review, pitfalls, and future directions. *Food Res. Int.* 145:110389. doi: 10.1016/j.foodres.2021.110389

Mrkva, K., and Van Boven, L. (2020). Salience theory of mere exposure: relative exposure increases liking, extremity, and emotional intensity. *J. Pers. Soc. Psychol.* 118, 1118–1145. doi: 10.1037/pspa0000184

Musch, J., and Reips, U. D. (2000). "A brief history of web experimenting" in *Psychological Experiments on the Internet.* ed. M. H. Birnbaum (Cambridge, MA: Academic Press)

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

Nittono, H., and Wada, Y. (2009). Gaze shifts do not affect preference judgments of graphic patterns. *Percept. Mot. Skills* 109, 79–94. doi: 10.2466/pms.109.1.79-94

Oppenheimer, D. M., Meyvis, T., and Davidenko, N. (2009). Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* 45, 867–872.

Orquin, J. L., and Loose, S. M. (2013). Attention and choice: a review on eye movements in decision making. *Acta Psychol.* 144, 190–206. doi: 10.1016/j.actpsy.2013.06.003

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., and Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4170–4175. doi: 10.1073/pnas.1415250112

Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T., and Heatherton, T. F. (2014). Social connection modulates perceptions of animacy. *Psychol. Sci.* 25, 1943–1948. doi: 10.1177/0956797614547706

Pratt, J., Radulescu, P. V., Guo, R. M., and Abrams, R. A. (2010). It's alive! Animate motion captures visual attention. *Psychol. Sci.* 21, 1724–1730. doi: 10.1177/0956797610387440

R Core Team. (2021). R: A language and environment for statistical computing. Available at: https://www.R-project.org/ (Accessed March 1, 2021).

Rochat, P., Morgan, R., and Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cogn. Dev.* 12, 537–561. doi: 10.1016/S0885-2014(97)90022-8

Rosin, P. L., and Lai, Y. K. (2015). "Non-photorealistic rendering of portraits" in *Proceedings of the Workshop on Computational Aesthetics* (Goslar: Eurographics Association), 159–170.

Rutherford, M. D., and Kuhlmeier, V. A. (2013). *Social perception: Detection and interpretation of Animacy, agency, and intention* MIT Press.

Rutherford, M. D., Pennington, B. F., and Rogers, S. J. (2006). The perception of animacy in young children with autism. *J. Autism Dev. Disord.* 36, 983–992. doi: 10.1007/s10803-006-0136-8

Saito, T., Almaraz, S., and Hugenberg, K. (2022). Happy = human: A feeling of belonging modulates the "expression-to-mind" effect. *Soc. Cogn.* 40, 213–227. doi: 10.1521/soco.2022.40.3.213

Saito, T., Nouchi, R., Kinjo, H., and Kawashima, R. (2017). Gaze bias in preference judgments by younger and older adults. *Front. Aging Neurosci.* 9:285. doi: 10.3389/fnagi.2017.00285

Saito, T., Sudo, R., and Takano, Y. (2020). The gaze bias effect in toddlers: preliminary evidence for the developmental study of visual decision-making. *Dev. Sci.* 23:e12969. doi: 10.1111/desc.12969

Scholl, B. J., and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309. doi: 10.1016/S1364-6613(00)01506-0

Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322. doi: 10.1038/nn1150

Simion, C., and Shimojo, S. (2006). Early interactions between orienting, visual sampling and decision making in facial preference. *Vis. Res.* 46, 3331–3335. doi: 10.1016/j.visres.2006.04.019

Simion, C., and Shimojo, S. (2007). Interrupting the cascade: orienting contributes to decision making even in the absence of visual stimulation. *Percept. Psychophys.* 69, 591–595. doi: 10.3758/BF03193916

Tavares, G., Perona, P., and Rangel, A. (2017). The attentional drift diffusion model of simple perceptual decision-making. *Front. Neurosci.* 11:468.

Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R., and Mohr, P. N. C. (2019). Gaze bias differences capture individual choice behaviour. *Nat. Hum. Behav.* 3, 625–635. doi: 10.1038/s41562-019-0584-8

Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–951. doi: 10.1068/p3101

Wang, S., and Rochat, P. (2017). Human perception of animacy in light of the uncanny valley phenomenon. *Perception* 46, 1386–1411. doi: 10.1177/0301006617722742

Yang, J., Wang, A., Yan, M., Zhu, Z., Chen, C., and Wang, Y. (2012). Distinct processing for pictures of animals and objects: evidence from eye movements. *Emotion* 12, 540–551. doi: 10.1037/a0026848

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9, 1–27. doi: 10.1037/h0025848

Zhou, H., and Fishbach, A. (2016). The pitfall of experimenting on the web: how unattended selective attrition leads to surprising (yet false) research conclusions. *J. Pers. Soc. Psychol.* 111, 493–504. doi: 10.1037/pspa0000056

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 27 April 2023
DOI 10.3389/fpsyg.2023.1146248

OPEN ACCESS

EDITED BY
Josefa N. S. Pandeirada,
University of Aveiro, Portugal

REVIEWED BY
Thomas Hagen,
University of Oslo, Norway
Ben Van Buren,
The New School, United States

*CORRESPONDENCE
Jeff Loucks
✉ jeff.loucks@uregina.ca

RECEIVED 17 January 2023
ACCEPTED 05 April 2023
PUBLISHED 27 April 2023

CITATION
Loucks J, Reise B, Gahite R and Fleming S (2023) Animate monitoring is not uniform: implications for the animate monitoring hypothesis.
*Front. Psychol.* 14:1146248.
doi: 10.3389/fpsyg.2023.1146248

# Animate monitoring is not uniform: implications for the animate monitoring hypothesis

Jeff Loucks[1]*, Berit Reise[2], Rosselle Gahite[1] and Shaun Fleming[1]

[1]University of Regina, Regina, SK, Canada, [2]Universität Osnabrück, Osnabrück, Germany

The animate monitoring hypothesis (AMH) purports that humans evolved specialized mechanisms that prioritize attention to animates over inanimates. Importantly, the hypothesis emphasizes that any animate—an entity that can move on its own—should take priority in attention. While many experiments have found general support for this hypothesis, there have yet been no systematic investigations into whether the type of animate matters for animate monitoring. In the present research we addressed this issue across three experiments. In Experiment 1, participants (*N*=53) searched for an animate or inanimate entity in a search task, and the animate was either a mammal or a non-mammal (e.g., bird, reptile, insect). Mammals were found significantly faster than inanimates, replicating the basic AMH finding. However, they were also found significantly faster than non-mammals, who were not found faster than inanimates. Two additional experiments were conducted to probe for differences among types of non-mammals using an inattentional blindness task. Experiment 2 (*N*=171) compared detection of mammals, insects, and inanimates, and Experiment 3 (*N*=174) compared birds and herpetofauna (reptiles and amphibians). In Experiment 2, mammals were spontaneously detected at significantly higher rates than insects, who were detected at only slightly higher rates than the inanimates. Furthermore, when participants did not consciously identify the target, they nonetheless could correctly guess the higher level category of the target (living vs. nonliving thing) for the mammals and the inanimates, but could not do so for the insects. We also found in Experiment 3 that reptiles and birds were spontaneously detected at rates similar to the mammals, but like insects they were not identified as living things at rates greater than chance when they were not consciously detected. These results do not support a strong claim that all animates are prioritized in attention, but they do call for a more nuanced view. As such, they open a new window into the nature of animate monitoring, which have implications for theories of its origin.

KEYWORDS

**animate monitoring, attention, animacy, perception, visual search, inattentional blindness**

## 1. Introduction

For a great majority of animals across a wide variety of ecosystems, those that pay attention to other animals in their immediate environment would seemingly survive longer. New et al. (2007) proposed the animate monitoring hypothesis (AMH), which states that humans (and potentially other animals) are biologically predisposed to pay greater attention to animates over inanimates, as a result of our evolution (and which is also likely shared with nearby evolutionary cousins). Using a change detection paradigm, they found that people detected changes to animate entities in scenes more quickly and more frequently than changes to inanimate entities.

This hypothesis has garnered significant interest, and numerous additional studies have provided support for the AMH. Animates are also detected more quickly in visual search tasks (Lipp et al., 2004; Jackson and Calvillo, 2013), more frequently reported in attentional blink and inattentional blindness tasks (Calvillo and Jackson, 2014; Guerrero and Calvillo, 2016), and receive longer fixations (Yang et al., 2012). Not every investigation has found support, however. Notably, Hagen and Laeng (2016) found no animate advantage in change blindness once the visual context of the scene had been accounted for. Hagen and Laeng (2017) also found that animates do not induce or reduce attentional blinks, though they are more accurately reported in such tasks (see also Hagen et al., 2018). Avoiding such visual confounds, Loucks et al. (2020) demonstrated that 4-year-old children remember a novel, arbitrary sequence of actions better if the sequence contains an animate rather than inanimate entity, despite identical appearances. Taken together with other complementary findings (Pratt et al., 2010; Nairne et al., 2017; van Buren and Scholl, 2017; Nguyen and van Buren, in press), these results suggest that once an entity has been ascribed an animate status, heightened attention and cognition follows.

While some of the above research indicates that animates stimuli are in general prioritized in cognition, little attention has been paid to the specific animate used. Importantly, the AMH would not suggest that there should be any difference according to the type of animate; an evolved system such as this should prioritize any animate in the observer's immediate vicinity. However, in another experiment of Loucks et al. (2020), they found that using a toy dog improved children's memory better than a toy beetle. It is possible that insects may not be considered the same kind of animate as mammals. For instance, most people consider mammals to be more similar to humans than insects (Eddy et al., 1993), and more worthy of moral consideration (Kellert, 1993; Tisdell et al., 2006). The amount of visual experience observers have for each type of animate also likely differs, in addition to the particular quality of that experience (e.g., Knight, 2008). Possidónio et al. (2019) also found that various types of animals differ from one another to observers on the basis of dimensions such as valence, arousal, and dangerousness, which may also affect their cognitive processing more generally.

The difference that Loucks et al. (2020) observed between the dog and the beetle was relatively weak, in a statistical sense, and it has only been observed with children, not adults. But, in considering the possibility that humans might attend to different types of animates differently, it is striking that research on animate monitoring tends to use mammals much more often than other types of animals. Most importantly, no research to date has ever systematically compared different types of animates in terms of their capacity to capture attention (with the exception of snakes: LoBue and DeLoache, 2008; and spiders: New and German, 2015).

Research on the neural correlates of animate processing in adults supports the possibility that the type of animate may matter for animate monitoring. For example, animals appear to be processed in a graded fashion according to perceived animacy (Connolly et al., 2012), and that those judged as being more animate (e.g., humans, chimpanzees, cats) activate distinct regions of lateral occipital cortex (LOC) relative to tools, but that those judged as being less animate (e.g., fish, insects) activate overlapping regions of LOC relative to tools (Sha et al., 2015). Importantly, this concept of perceived animacy goes beyond a binary definition of animacy—the latter would refer only to whether an entity has the capacity for self-initiated movement, while the former involves a graded concept of agency in relation to humans.

However, two other bodies of work suggest that the type of animate should not matter. One is that of Thorpe and colleagues on rapid visual categorization (e.g., Macé et al., 2009; Crouzet et al., 2012; Wu et al., 2015), which has shown that adults can identify animals within approximately 120 ms. Importantly, this rapid detection occurs at the superordinate level of "animal," and not at the basic level of "dog" or "bird" (Wu et al., 2015). Thus, at early levels of awareness, adults know that there is some kind of animal present, without knowing exactly what animal it is. But as with the literature on the AMH, while a fairly wide range of animal types are used in this line of research, a large proportion are mammals, and no specific comparisons have been made between different types.

Another is the literature on visual features diagnostic of animates vs. inanimates, which are thought to be distinguishable on the basis of mid-level features such as the degree of curvilinearity (Wichmann et al., 2010). Long et al. (2017) created synthetic images of animals and objects that preserved certain texture and form information but removed basic-level diagnostic information: "texforms." They found that observers could find texforms faster when they were embedded among texforms of a different higher level category (e.g., finding an animal among objects) than when they were from the same category, and that the degree to which a texform displayed curvilinearity was predictive of whether it was classified by observers as an animal (see also Zachariou et al., 2018). However, recently He and Cheung (2019) equated animals and tools in gist statistics—by using elongated and round types of both—and found that observers were still faster at detecting animals. Taken together with some of the findings of Long et al. (2017), it appears that visual features cannot entirely account for the animate advantage. In any case, no systematic comparison between animal types has been made in this literature.

Thus, in the current research, we aimed to compare the attentional capture of (non-human) mammals against a variety of non-mammals. In contrast with the AMH, we hypothesized that mammals would hold a higher status in attention over non-mammals, and would thus be detected more easily/rapidly in a variety of tasks. We believe they hold this elevated status given their higher similarity to humans, either in form or in perceived animacy, and/or the different experiences humans have with them. In terms of the general animate advantage, we hypothesized that mammals would be detected more easily/rapidly than inanimates, but were unsure about non-mammals in this respect. In an initial experiment, we first assessed whether mammals were generally advantaged in detection over a diverse group of non-mammals in visual search. Two additional experiments were then conducted to compare detection of mammals against specific types of non-mammals using inattentional blindness, to extend the results of the first experiment by way of a different methodology.

## 2. Experiment 1

In Experiment 1 we compared search times for mammals against a heterogeneous group of non-mammals in a visual search task. Our primary interest with this first experiment was to determine whether there would be any global advantage for mammals above other animals, rather than comparing them to specific classes at the same level (e.g., reptiles, birds) in an exhaustive sense. We also did not control for all visual features in our entities, such as curvilinearity or gist, as we wanted to have entities be easily recognizable and in their typical posture. However, we did ensure that targets were equated on certain visual features, such as luminance and contrast (e.g., the SHINE toolbox, Willenbockel et al., 2010). We also specifically selected mammals that our (primarily White Canadian) participants had little experience with, due to the fact that there are likely pre-existing differences in exposure to mammals and non-mammals. We predicted that mammals would be detected faster than non-mammals and inanimates, but that non-mammals would not be detected faster than inanimates.

## 2.1. Method

### 2.1.1. Participants

Participants were 53 University of Regina undergraduate students (8 male), who earned partial course credit for their participation. We aimed for a sample size of 34, as this would allow us to detect a medium effect size ($d = 0.50$; assuming power = 0.80, and $\alpha = 0.05$), but ended up sampling more participants to satisfy student demand in our department. An additional 4 individuals participated (1 male) whose data was dropped because they provided valid reaction time data for less than 50% of trials. Self-reported race of our sample was: White ($n = 31$), South Asian ($n = 10$), Black ($n = 4$), mixed ($n = 3$), East Asian ($n = 2$), Middle Eastern ($n = 1$), and undisclosed ($n = 1$).

### 2.1.2. Stimuli

Stimuli consisted of greyscale images of 48 animates and 48 inanimates. For both groups of stimuli, 16 were target images, 16 were associated presentation images, and 16 were distractor images. Presentation images were secondary images of the targets, so that only the category of the target was cued on the presentation screen, and not the exact target image to be found on the search screen. For the animates there was a further subdivision of 8 mammals and 8 non-mammals among the targets and presentation images. We specifically wanted to avoid selecting animals that are encountered frequently by Western adults (e.g., dogs, pigs). The 8 mammal targets were an armadillo, a camel, a chinchilla, a ferret, a lemur, an okapi, a rhino, and a saiga. The 8 non-mammal targets were a centipede, a crab, a gecko, a mantis, a puffin, a squid, and a turtle. Our selection of these particular animals was based on a desire to have variety in overall appearance across mammals and non-mammals. The 16 inanimate targets were a baseball glove, a belt, a bottle opener, a watering can, a cassette tape, a drill, an egg slicer, a picture frame, a hole puncher, a blender, a lawnmower, a sled, a speaker, a staple remover, a tea infuser, and a tricycle. Distractor stimuli were other animates and inanimates that never served as targets (and half of the animate distractors were mammals, half non-mammals). We also analyzed the average luminosity and contrast (the standard deviation of the luminance distribution) of all target stimuli using the SHINE toolbox (Willenbockel et al., 2010). As a group, none of the three entity types were different from each other on either of these variables, nor were the animates as a group different from the inanimates, all $t$'s > 1.6, $p$'s > 0.11. The entire set of stimuli can be viewed at: https://osf.io/b2zpf/?view_only=4129537fbef14d018b52e90e6a4c55c1.

### 2.1.3. Design and procedure

Figure 1 displays the trial structure. Each trial consisted of a presentation screen (3 s) and a search screen (until response). The ITI was fixed at 1 s. Presentation screens displayed one animate and one inanimate image as possible targets to look for on the upcoming search screen; the specific images on the presentation screen were different than the target images (see stimuli above). Stimuli were positioned along the center of the y-axis of the screen and at 25 and 75% of the x-axis of the screen. The target on the presentation screen was randomly selected (e.g., squid), and then paired with a random target from the opposing category (e.g., tricycle). The search screen displayed the selected target and three additional distractors, arranged in a 2 × 2 grid. The stimuli were centered within each of the four quadrants defined by bisecting the screen along the x-and y-axes. Regardless of the target category, each search screen displayed two animate and inanimate entities, and the specific positioning of each stimulus within the quadrants was random. Participants were required to indicate which position on the grid contained the target entity, using the T, Y, G, and H keys on their keyboard, as quickly and accurately as they could. The right/left position of each target on the presentation screen was randomly selected. A first block of trials displayed targets in one right/left position on the presentation screen, and then a second block displayed them in the opposite position. Trial order within a block was random, and there was no break between blocks (to the participant it was all one block). Thus each target was searched for two times total, across a total of 64 trials. A video of sample trials can be found at: https://osf.io/b2zpf/?view_only=4129537fbef14d018b52e90e6a4c55c1.

The experiment was implemented in PsychoPy and was hosted on Pavlovia.org. After providing consent, participants were asked to report on their gender and race/ethnicity, and then provided an instruction screen. Participants were told that on each trial, they would be shown an animal and an object, and would have to find either the animal or the object on a subsequent screen, as quickly and accurately as they could.

## 2.2. Results

Accuracy was high: 97.2% correct in the mammal condition, 95.9% correct in the non-mammal condition, and 94.8% in the inanimate condition. Since the task was straightforward for participants in terms of accuracy, this supports the idea that reaction time would be a more sensitive measure of the search efficiency for this task.

Only reaction times for accurate identifications that occurred within 2,000 ms of the search screen onset were included in the analysis (14.1% of trials excluded). Figure 2 displays these mean reaction times for each type of entity. A repeated measures ANOVA on entity type was significant, $F(2, 104) = 8.82$, $p < 0.001$, partial $\eta^2 = 0.15$. Paired samples $t$-tests revealed that search times for mammals were significantly faster than those for non-mammals, $t(52) = 2.88$, $p = 0.006$, Cohen's $d = 0.40$, and inanimates, $t(52) = 4.14$, $p < 0.001$, $d = 0.57$, and that search times for non-mammals were not different than those for inanimates, $t(52) = 1.26$, $p = 0.21$. A Bayesian

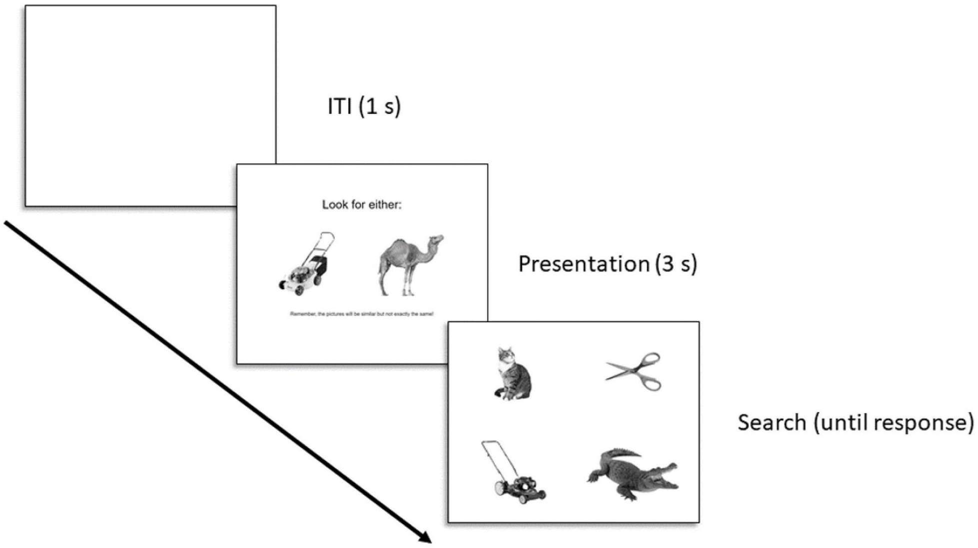


FIGURE 1
Trial structure for Experiment 1. The text below the images on the presentation display read: "Remember, the pictures will be similar but not exactly the same!".

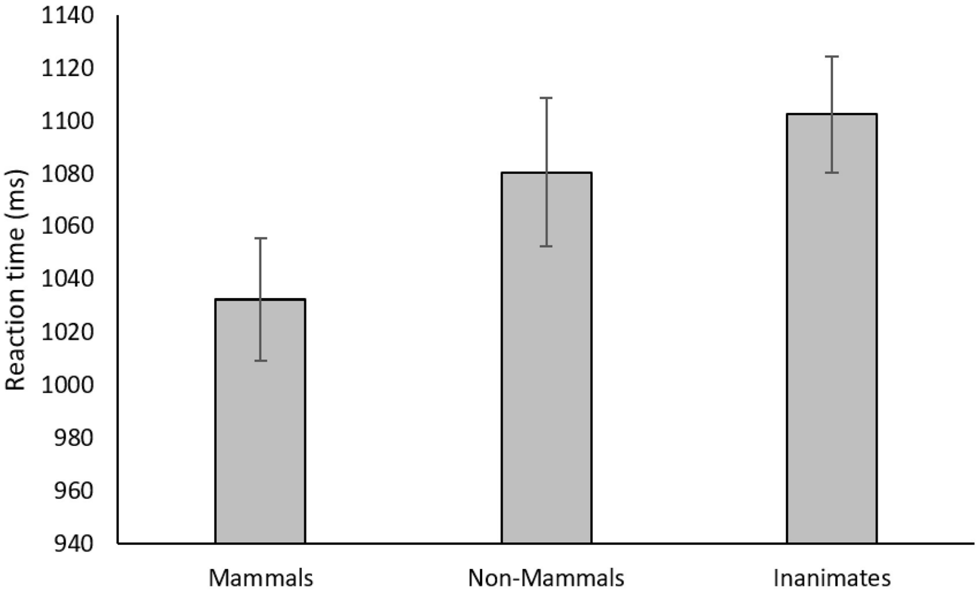


FIGURE 2
Mean reaction times as a function of entity type in Experiment 1. Error bars represent standard error.

analysis of the difference between the non-mammals and the inanimates provided moderate evidence for the null hypothesis, $BF_{01} = 3.18$.

Figure 3 displays the mean reaction times for each animal in rank order. Although there is a fairly clear separation of the mammal searches being generally faster than the non-mammal searches, it is not perfectly clean: two of the non-mammals are among the faster ranks, and two of the mammals are among the slower ranks. Although matched on luminosity and contrast, we cannot control for all visual differences between our stimuli, and thus we cannot fully determine why an individual animal may have been easier to find than another. However, our main prediction was that, even amidst the noise of different body shapes, faces, patterns, and textures, there would nonetheless be a signal that stands out as the mammal/non-mammal distinction in search times.

## 2.3. Discussion

The results of Experiment 1 indicate that not all animals appear to be equal in their ability to capture observers' attention: mammals are detected more rapidly than non-mammals in visual search. It is

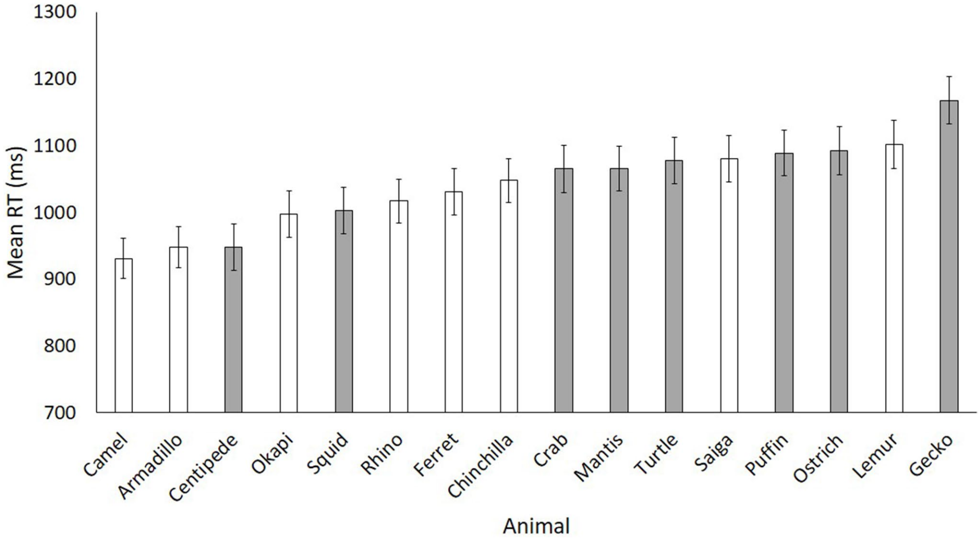


FIGURE 3
Mean reaction times for each animal in Experiment 1. Mammals are white, non-mammals are grey. Error bars represent standard error.

important to note that we selected mammals which are not common in Westerner's daily experience—in an attempt to put them on more equal footing with the non-mammals—but which were likely easily binned by our participants into the relevant categories (e.g., mammal, bird, insect). Thus, it is not likely a difference in experience with specific targets which drives the effect. Instead, it is likely the detection of perceptual features that are diagnostic of the particular category (e.g., four limbs, furry)—or the combination of more than one feature—which drives the increased fidelity of observer's attention. Non-mammals were not simply detected more slowly than mammals—importantly, they were also not detected significantly faster than inanimate objects. This result does not support the idea that animate monitoring is based on "pure" animacy—the capacity for internally generated motion. All of our non-mammals have this property, and none of our inanimates do. Although this negative result with non-mammals seems to stand in contrast to a relatively large literature supporting the AMH, recall that most studies of the AMH use primarily mammals as stimuli.

However, strong conclusions should not be drawn from a single experiment. If mammals indeed hold special status over other animals in this respect, then this advantage should also be observed using different methods. Thus, one goal of Experiments 2 and 3 was to replicate this basic advantage using an inattentional blindness task. A secondary goal of these additional experiments was to get more fine-grained data on the relative advantage of mammals over specific classes of animals—specifically, insects, birds, and herpetofauna (reptiles and amphibians).

## 3. Experiment 2

Experiments 2 and 3 are highly similar experiments, but we report them individually because they were collected at different times with slightly different samples. Both utilized an inattentional blindness paradigm to explore how well various animate and inanimate entities capture observers' attention when they appear unexpectedly amidst another task. In developing this experiment we were inspired by the work Calvillo and Jackson (2014), and utilized a task and stimuli akin to theirs. The stimuli we used were thus normed line drawings from Snodgrass and Vanderwort (1980). These stimuli contrast with those from Experiment 1, where the goal was to select animates that our participants would have less experience with; In this case the stimuli are more familiar and recognizable animals, which is important for conscious verbal reporting of the stimuli when they appear unexpectedly.

In Experiment 2, we first explored whether mammals would be detected more robustly than insects and inanimates. This was achieved by inserting an unexpected image of an entity amidst a secondary, sham task (finding a color word). In addition to the standard data on each participants' ability to spontaneously detect the unexpected image, we also asked participants to guess if the image that they saw was of a living or non-living thing, regardless of whether they spontaneously noticed the image or not. This measure assessed the ability of participants who failed to consciously recognize the image to nonetheless correctly identify a critical aspect of its identity through unconscious/partially conscious recognition. We predicted that mammals would be detected more readily than insects and inanimates, and that insects would not be detected more readily than inanimates.

### 3.1. Method

### 3.1.1. Participants

Participants for Experiment 2 were 171 University of Regina undergraduate students (42 male, 2 non-binary), who earned partial course credit for their participation. This sample size was chosen based on a power analysis assuming the same percentage difference as obtained in Calvillo and Jackson (2014) for animates vs. inanimates in their low load condition (~30%). Self-reported race of the sample was: White (*n* = 95), South Asian (*n* = 21), Black (*n* = 11), South-East Asian (*n* = 11), East Asian (*n* = 9), Middle Eastern (*n* = 7), Indigenous (*n* = 7), Métis (*n* = 6), mixed (*n* = 3), and Afro Carribean (*n* = 1). An additional four participants were sampled but dropped, as they did not notice the color word on two or more real trials, with one of these being the critical trial.

Participants were randomly assigned to one of the three target conditions. Approximately equal numbers of participants were assigned to the mammal ($n = 57$), insect ($n = 54$), and inanimate ($n = 61$) conditions.

### 3.1.2. Stimuli

The target stimuli for the inattentional blindness task were black and white line drawings of five mammals (camel, deer, fox, lion, raccoon), five insects (ant, bee, beetle, butterfly, grasshopper), and five inanimate objects (flag, helicopter, kite, sled, whistle), taken from Snodgrass and Vanderwort (1980). We selected fewer animals than in experiment 1 (from 8 to 5) as we were constrained by the available choices of this particular stimulus set. Our selection of the specific 5 for each category was initially based on our intuitions about how readily participants would be able to name the entities if they did happen to notice them. Following this initial selection, entities were chosen for all categories that had roughly similar scores for familiarity and naming agreement, in the middle range for both measures (Snodgrass and Vanderwort, 1980). Each image measured 300 × 300 pixels. They can be viewed at https://osf.io/b2zpf/?view_only=4129537fbef14d018b52e90e6a4c55c1.

The target stimuli for the sham word finding task were color words. The targets in the practice trials were green, blue, and red, and in the real trials orange, purple, and yellow. This target order was fixed, and thus yellow was the target on the critical (inattentional blindness) trial for all participants. Non-target words were various 3–6 letter words that were non-color and non-animal words. All words were presented in capital letters in black.

### 3.1.3. Design and procedure

Five of the six trials of the sham task consisted of a blank screen (1 s), a central fixation cross (1 s), a word grid surrounding the fixation cross (1 s), a perceptual mask (1 s), and an answer screen (until response). A schematic of each trial and the critical trial can be found in Figure 4. On the grid screen the words were presented in the northwest, northeast, southwest, and southeast corners of the screen, surrounding the fixation. The specific visual angle of the stimuli was dependent on the participants screen size. The sixth and final trial was the critical trial, which was identical to the first five trials except that the fixation cross on the word grid screen was replaced with one of the target images. Each participant only saw one target on this one critical trial.

The experiment was implemented and hosted using the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020). After providing consent and demographic information, participants were instructed that in each trial, four words would be quickly flashed on the screen, and their job was to find the one color word amidst this set. Participants completed three practice trials first, during which they received feedback about their performance. All participants were correct on at least 2/3 practice trials. Following practice, participants engaged in the three real trials, during which feedback was not given. All participants were correct on the critical trial. Immediately following the critical trial, participants were asked two questions across two separate screens. The first was "Did you notice anything odd on the screen which was not there on the previous trials? If so, please tell me what you saw, with as much detail as you can. If you did not notice anything, that's OK—please just type 'no.'" Participants inputted their response to this question via a text box. After this, they were told on a new screen "On this last trial an image was presented in the middle of the screen, between the words. Even if you cannot recall very much about what you saw, can you guess whether it was an image of a living thing or a non-living thing?" Participants chose between two response buttons labeled "Living" and "Non-Living." The entire session took approximately 6 min.

## 3.2. Results

All spontaneous responses were evaluated by the first and third author. Spontaneous identifications were counted if the participant correctly named the target (e.g., "camel"), or if a highly similar entity was named (e.g., "a goat" for the deer, "skateboard" for the sled), or a higher level name for the entity (most commonly this was "bug" for the insects, and "animal" for the raccoon). One participant said "bird" for the beetle, and we elected not to count this.

Overall, mammals were noticed a little over half of the time, at 58%, which was a significantly higher rate than that of the insects at 35%, test of two proportions $z = 2.40$, $p = 0.016$, Cohen's $h = 0.46$, and the inanimates at 21%, $z = 4.07$, $p < 0.001$, Cohen's $h = 0.77$. The identification rate for the insects was only marginally higher than that of the inanimates, $z = 1.66$, $p = 0.097$, Cohen's $h = 0.31$.

If a participant did not notice anything spontaneously, or noticed something but could not correctly identify it (e.g., "I saw a picture but I do not know what"), their guessing data—the second question—were eligible for analysis. This yielded sample sizes of $n = 24$ in the mammal condition, $n = 35$ in the insect condition, and $n = 48$ in the inanimate condition. For these participants, those who were shown a mammal guessed correctly that it was a living thing 79% of the time, which was significantly greater than chance, goodness of fit $\chi^2(1) = 8.17$, $p = 0.004$, and those who were shown an inanimate guessed correctly that it was a non-living thing 69% of the time, which was also significantly greater than chance, $\chi^2(1) = 6.75$, $p = 0.009$. However, those who were shown an insect guessed correctly only 63% of the time, which was not significantly greater than chance, $\chi^2(1) = 2.31$, $p > 0.12$.

## 3.3. Replication

Although mammals were clearly noticed at higher rates than inanimates, insects were only noticed at slightly higher rates, and the difference was not statistically significant. These data supported our predictions. However, as the difference with insects was marginally significant, we elected to conduct a replication of Experiment 2 with a different sample, in order to see if the same results would hold.

This sample was collected approximately 4 months after the original sample, and was an international sample of $N = 179$ from Prolific (www.prolific.co; 82 male). Self-reported race of this sample was: White ($n = 122$), Black ($n = 25$), Hispanic ($n = 9$), East Asian ($n = 9$), South Asian ($n = 5$), mixed ($n = 4$), unreported ($n = 3$), South-East Asian ($n = 1$), and Middle Eastern ($n = 1$). The country of residence of these participants was: United Kingdom ($n = 81$), elsewhere in Europe (n = 49), United States ($n = 29$), Canada ($n = 6$),

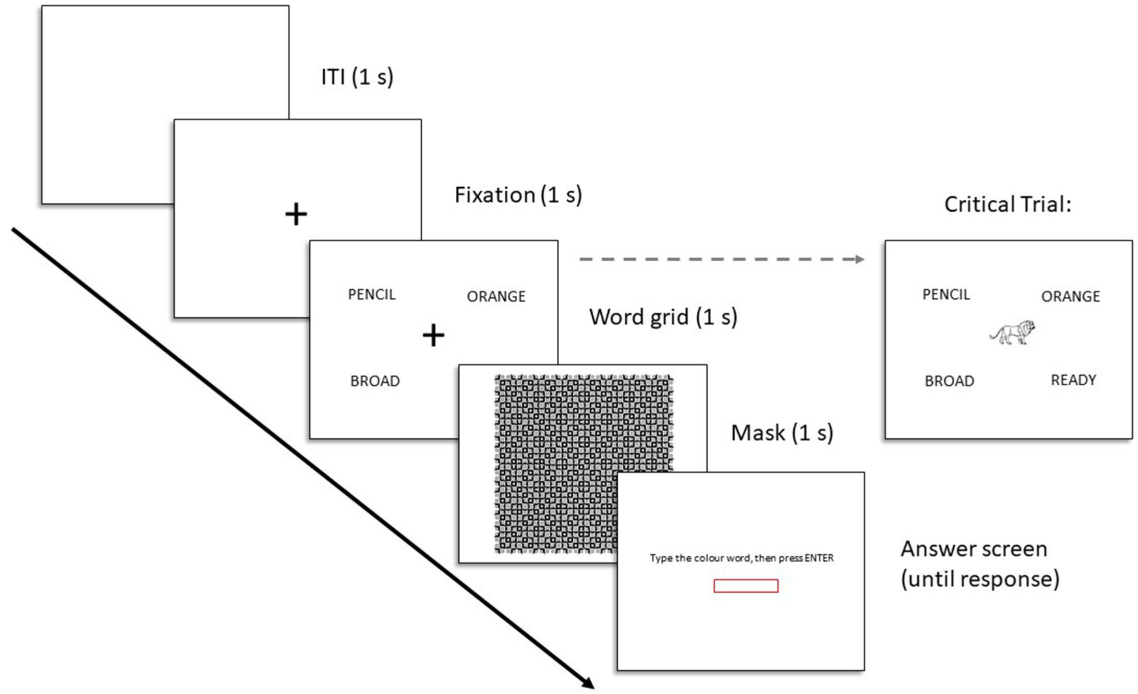


FIGURE 4
Trial structure for Experiments 2 and 3.

Australia ($n=5$), Mexico ($n=5$), Israel ($n=3$), and Chile ($n=1$). An additional 7 participants were sampled but dropped, as they did not notice the color word on two or more real trials, with one of these being the critical trial. Approximately equal numbers of participants were assigned to the mammal ($n=56$), insect ($n=62$), and inanimate ($n=61$) conditions.

Noticing rates in this replication were highly similar to the initial experiment. Mammals were noticed at rate of 53%, which was a significantly higher rate than that of the insects at 32%, $z=2.23$, $p=0.026$, Cohen's $h=0.42$, and the inanimates at 23%, $z=3.31$, $p<0.001$, Cohen's $h=0.62$. However, with this sample the identification rate for the insects was not significantly higher than that of the inanimates, $z=1.15$, $p>0.25$.

For the guessing data, there were sample sizes of $n=26$ in the mammal condition, $n=42$ in the insect condition, and $n=47$ in the inanimate condition. Mammals were identified as living things 73% of the time, which was significantly greater than chance, goodness of fit $\chi^2(1)=5.54$, $p=0.019$, and inanimates were identified as non-living things 81% of the time, which was also significantly greater than chance, $\chi^2(1)=17.89$, $p<0.001$. Replicating the initial experiment, insects were identified as livings things only 41% of the time, which was not different than chance, $\chi^2(1)=1.52$, $p>0.21$.

## 3.4. Omnibus analysis

Despite some differences in sample characteristics, we also ran an omnibus analysis pooling the initial and replication experiments together, since larger sample sizes are directly proportional to determining which percentage differences will be deemed significant with proportional data. In this case. In this case, the total sample sizes were $n=113$ in the mammal condition, $n=116$ in the insect condition, and $n=122$ in the inanimate condition.

Overall, mammals were noticed at 57%, which was a significantly higher rate than that of insects at 35%, test of two proportions $z=3.29$, $p=0.001$, Cohen's $h=0.45$, and inanimates at 22%, $z=5.52$, $p<0.001$, Cohen's $h=0.74$. In this analysis, the identification rate for the insects was also significantly higher than that of the inanimates, $z=2.32$, $p=0.020$, Cohen's $h=0.29$. Note that the effect size for the mammal advantage over inanimates was larger than the effect size insect advantage. Table 1 displays the noticing rates for all entities from the omnibus analysis.

For the guessing data, sample sizes were $n=48$ in the mammal condition, $n=75$ in the insect condition, and $n=95$ in the inanimate condition. For these participants, those who were shown a mammal guessed correctly that it was a living thing 69% of the time, which was significantly greater than chance, goodness of fit $\chi^2(1)=7.37$, $p=0.007$, and those who were shown an inanimate guessed correctly that it was a non-living thing 75% of the time, which was also significantly greater than chance, $\chi^2(1)=23.25$, $p<0.001$. However, those who were shown an insect guessed correctly only 49% of the time, which was no different than chance, $\chi^2(1)=0.01$, $p>0.90$.

## 3.5. Discussion

The data from Experiment 2 show a clear divide in participants' ability to spontaneously detect an unexpected image of a mammal, an insect, or an inanimate, and replicate the basic finding from Experiment 1. Participants noticed the mammal about half of the time, while they only noticed insects about a third of the time, and inaninmates about a fifth of the time. With the very large sample from

TABLE 1 Detection rates for each entity in Experiments 2 and 3.

| Experiment 2: Omnibus data | | | Experiment 3 | |
|---|---|---|---|---|
| Mammal | Insect | Inanimate | Bird | Herpetofauna |
| Fox: 69% | Bee: 50% | Helicopter: 33% | Eagle: 70% | Frog: 62% |
| Camel: 58% | Beetle: 46% | Flag: 24% | Chicken: 62% | Snake: 52% |
| Raccoon: 55% | Grasshopper: 33% | Kite: 17% | Ostrich: 55% | Turtle: 46% |
| Deer: 50% | Ant: 32% | Whistle: 17% | Owl: 38% | Alligator: 43% |
| Lion: 47% | Butterfly: 13% | Sled: 16% | | |

the omnibus analysis, we have a reliable estimate that insect detection was about 13% higher than inanimate detection, which is a considerably smaller advantage in comparison to the mammals at 35%. This decreased strength of an animate advantage for insects is likely correlated with the neural representation of insects in the brain, which are represented in a network overlapping with inanimate tools (Sha et al., 2015). This distinction may be driven via the perceived animacy of the entity, which is higher for mammals in comparison to insects (Connolly et al., 2012). An alternative explanation might be that the mammals were rated more highly, on average, in danger or usefulness than the insects and inanimates (e.g., Wurm, 2007). We do not have measures of these dimensions for these image stimuli, but future researchers could incorporate this into their design.

These findings also generally replicate the results of Calvillo and Jackson (2014), who compared animates to inanimates in a very similar inattentional blindness task, but with an additional manipulation of working memory. While they showed a clear advantage for animates, half of their animate stimuli were humans, and another 40% were non-human mammals. Our rates of mammal detection are in the ballpark for their data, but we used a different number of words in the word task (4 vs. their 3 and 6), so the rates are not easily comparable. It may of interest in future studies to compare humans and non-human mammals to each other, in order to see whether humans have an advantage in detection (e.g., Bonatti et al., 2002).

It is also noteworthy that broad categories of living vs. non-living things were accessible to our participants in a partially conscious manner, but the former of these was only possible when the target was a mammal, and not when it was an insect. These results are reminiscent of Wu et al. (2015), who showed that the visual system can rapidly assess the animate vs. inanimate distinction in as short as 120 ms. However, they largely used mammals for their animate stimuli, and our results indicate that insects are not rapidly assessed in the same fashion. We will save further discussion of this issue for the general discussion.

## 3.6. Experiment 3

In Experiment 3, we investigated the detection rates of birds and herpetofauna (reptiles and amphibians) using the same methods as Experiment 2, in order to get further clarity on potential hierarchical differences in animate status. Possidónio et al. (2019) found that people rate birds as having both a higher capacity for cognition and being more similar to humans in comparison to herpetofauna. We hypothesized that an animate's ability to capture attention would be related to observer's perception of its animate status, and thus predicted that birds would be recognized more readily than the herpetofauna, but that both would be recognized at rates lower than the mammals from Experiment 1.

## 3.7. Method

### 3.7.1. Participants

Participants for Experiment 3 were 173 University of Regina undergraduate students (29 male, 2 non-binary), who earned partial course credit for their participation. We aimed for a higher sample size in this experiment than in Experiment 2, as we did not anticipate as large a percentage difference between these animal types as between mammals and insects. Self-reported race of this sample was: White ($n = 106$), South-East Asian ($n = 21$), Black ($n = 17$), South Asian ($n = 17$), Indigenous ($n = 6$), mixed ($n = 3$), West Asian ($n = 2$), and Arab ($n = 1$). An additional 11 people were sampled but dropped, due to missing the color word on 2/3 real trials, with one of these being the critical trial. Approximately equal numbers of participants were assigned to the bird ($n = 86$) and herpetofauna ($n = 87$) conditions.

### 3.7.2. Stimuli

The target stimuli for this experiment included four birds (chicken, eagle, ostrich, owl) and four herpetofauna (alligator, frog, snake, turtle), taken from Snodgrass and Vanderwort (1980). We only selected 4 animals for each group as we were again constrained by the available animals in the set, and there were only 3 reptiles and 1 frog. Four birds were selected following this, and as in Experiment 2, all of these entities had roughly similar scores for familiarity and image agreement, in the middle range for both measures (Snodgrass and Vanderwort, 1980). They can be viewed at https://osf.io/b2zpf/?view_only=4129537fbef14d018b52e90e6a4c55c1.

### 3.7.3. Design and procedure

The design and procedure was identical to Experiment 2.

## 3.8. Results

Spontaneous responses were evaluated by the first and fourth author, in the same manner as Experiment 2. One participant reported "a cat in sunglasses" for the owl, and this was not counted.

Overall, birds were noticed about half of the time, at 56%, which was not statistically different from the rate of the herpetofauna at 51%, $z = 0.69$, $p > 0.49$. Although statistical comparisons are not appropriate given that these are different samples, it is clear that these detection rates are extremely similar to the mammal detection rates in Experiment 2. Table 1 displays the noticing rates for each animal.

The sample sizes for those participants who failed to spontaneously identify the image were $n = 38$ in the bird condition and $n = 43$ in the herpetofauna condition. For these participants, those who were shown a bird guessed correctly that it was a living thing only 50% of the time, which was right at chance, and those who were shown herpetofauna correctly guessed that it was a living thing only 44% of the time, which was no different than chance, $\chi^2(1) = 0.58$, $p > 0.45$.

## 3.9. Discussion

The results of Experiment 3 further indicate a role for perceived animacy in animate detection. On the one hand—and in contrast to our prediction—the spontaneous detection rates for the birds and herpetofauna were quite similar to the mammals from Experiment 2, with about half of participants noticing these animals. This shows a fairly clear advantage for these three animal types over inanimates. On the other hand—and in partial support of our prediction—those who failed to notice the image explicitly were no better than chance at guessing whether these animals were living things, while those who viewed mammals in Experiment 2 could do so. This result supports our hypothesis that an animate's ability to capture attention may be related to its perceived animacy, with birds and herpetofauna ranking slightly lower than mammals, but above insects, on such a hierarchy.

# 4. General discussion

The present findings demonstrate that not all animals are equal in the mind's eye, unlike what New et al. (2007) originally theorized. Experiment 1 demonstrated that mammals are detected more rapidly than a variety of non-mammals, who did not appear to have an animate advantage at all. This result does not support the idea animate monitoring is applied broadly across all animates. Experiments 2 and 3 painted a more nuanced picture of this initial finding: mammals were very clearly noticed more easily in the context of inattentional blindness in comparison to insects, who were noticed at a much lower, but somewhat higher, rates above inanimates. Birds, reptiles, and amphibians appear to also be noticed at higher rates than inanimates, but even so appear to be processed in a slightly more fragile manner in comparison to mammals, sharing poor implicit recognition as living things along with insects. Taken together, these results suggest that animate monitoring does not operate in an all-or-nothing fashion, and instead may operate in a more graded fashion, potentially as it relates to the animate status of the entities in the observer's attentional window.

To be clear, the present results certainly do not argue for a rejection of New et al. (2007) original hypothesis. Instead, they suggest there may be limits on (1) how broadly the effect applies across the animal kingdom, and (2) the relative power of a specific animate's ability to eschew an animate advantage according to context. For example, in Experiment 1 a variety of non-mammals were not detected more quickly than a variety of inanimate objects, while in Experiments 2 and 3 non-mammals were noticed at higher rates than the inanimates (though just barely for insects). This relative difference in the mammal advantage may have been due to task differences. In Experiment 1, attention to the entities was assessed in a competitive fashion (participants were searching for one of either an animate or an inanimate), while in Experiments 2 and 3 the entities themselves were used to evoke observers' attention. Perhaps because non-mammals are less robustly attended to, they cannot outcompete inanimates in more competitive attentional situations, but can more readily stand out from inanimates when they are encountered unexpectedly. It is also possible that this difference is the result of the differences in the stimuli themselves: relatively unfamiliar, real images of animals were used in Experiment 1, while relatively recognizable line drawings of much more familiar animals (to mostly White Westerners) were used in Experiments 2 and 3. In these latter experiments, this may have given the non-mammals an advantage that they may not normally have in real life, while mammals do not require this advantage to be easily recognized and found. Further research that can directly manipulate the roles of familiarity and type of attention in animate detection could help to tease apart possible explanations for this difference.

Overall, these findings are broadly consonant with neuroimaging research on animals and inanimate tools, which has revealed a continuum of perceived animacy in LOC (Connolly et al., 2012). Sha et al. (2015) demonstrated that there is no neural distinction between animate and inanimate objects in ventral vision. While activation for highly familiar mammals (cats, dogs, and humans) was clearly distinguished from activation for inanimate tools in LOC, activation for less familiar mammals and birds were less clearly differentiated in this respect, and activation for fish and invertebrates clearly overlapped with that for tools. Sha et al. argued that the representation of an animacy continuum in the brain may still be the result of evolutionary pressures, but these shaped visual perception on the basis of the agentive capacity of an animal in the environment, and/or the similarity of the animate in relation to humans. This idea differs from a broad interpretation of animate monitoring that New et al. (2007) originally hypothesized. Future research on differences between types of animates should increase the number of types as well, in order to get more detail on this possible hierarchy. For example, the present research included Mollusca only in Experiment 1, and did not examine fish in any of the experiments.

What is the ultimate nature of these differences, mechanistically? Let us first consider the possible role of surface visual features. In Experiment 1, although our targets did not differ in average luminosity or contrast, they may have differed in mid-level features such as curvilinearity, and this may also have been the case for the line drawings in Experiments 2 and 3. Such a difference would most likely have been present for the mammal/inanimate comparison, and less likely for the mammal/non-mammal comparison. However, considering curvilinearity on its own, a higher degree should have made search more difficult, as Long et al. (2017) found that inanimate texforms were found faster than animate texforms. Further, the results of He and Cheung (2019) indicated that when certain mid-level features are controlled for (i.e., overall shape) an animate advantage still persists. Further investigation into the nature of these broad visual features will certainly be fruitful, but such features do not seem readily poised to explain the present results. On the other hand, visual features *must* be playing some role in the findings, as the different types of animates all share common morphological features which are strongly correlated within their respective categories. Perhaps the parts of mammals are more rapidly encoded or integrated into a whole than the parts of birds and herpetofauna, and perhaps the parts of insects are the most slowly encoded or integrated of all the animates. For insects in particular, perhaps these processes are slower than the rapid assessment of whether an entity is inanimate (Crouzet et al., 2012), and this is why insects were not implicitly recognized as being living things by our participants in Experiment 2.

Another possibility is that non-mammals are not encoded in as robust a manner in working memory as mammals during encoding. This could be related to perceptual differences in parts or overall form, as discussed in the preceding paragraph, but it may also be due to semantic differences between the entities (e.g., mammals as better representatives for the concept "animate"). In Experiment 1, while the presentation stage was seemingly long enough (3 s) to promote sufficient encoding of the entities, there may be differences in the strength or survivability of the representations for mammals in comparison to non-mammals and inanimates (especially in the context of more complex search). It is also possible that encoding of the mammals was superior to the other entities during the search stage itself. Differences in how easily the entities could be recognized in Experiments 2 and 3, both consciously and unconsciously, may also be related to how robustly they can intrude on attention and working memory in the midst of another task. Hagen and Laeng (2017) provided evidence that animates do not appear to be attended to preferentially compared to inanimates in an RSVP task, but do appear to be reported more accurately in such tasks. They similarly argued that this finding is most consistent with either post-attentive perceptual processing, or survivability in short-term memory. The present findings suggest that such processes most readily apply only to mammals, however, and apply to other animals less so as we move down a potential hierarchy of animate status.

More broadly, these findings argue for a theoretical shift in our conceptualization of animate monitoring, but the exact nature of the shift requires further study and investigation. One possibility is that a more accurate term for the effect might be "agentive monitoring." Evolution may have forged an attentional system in humans that prioritizes entities in the environment that have a relatively high capacity to notice or react the observer themselves. Mammals would subsequently receive the greater prioritization than birds, reptiles, or insects. Perhaps stronger "agentivity" in an animate boost processing more than just animacy itself. Blunt and VanArsdall (2021) found animate imagery and featural animacy had additive effects on word memory; Perhaps these concepts relate to the different animates used in the present research. A distinct possibility is that evolution may have shaped a system which prioritizes humans, and that mammals get prioritized over birds in that they are perceived to be more similar to humans, and thus get a little boost in processing as a result (see Ritchie et al., 2021, for a related discussion of the organization of LOC). A third possibility is that animate monitoring in adults is the result of experience, and not evolution. Mammals may have more robust representations in working memory because they are encountered more frequently than non-mammals, or interacted with in a deeper way over the course of development.

We thus see two important directions for future research on animate monitoring. First, there is a need for a more systematic investigations pitting the detection of various classes of animals against one another. This will provide more clarity on the precise nature of a potential hierarchy of animals, which may map onto a perceived animacy/agentive continuum. Second, there is a need for further explorations into the nature of animate monitoring in early development (i.e., Loucks et al., 2020). Such investigations are critical in distinguishing between evolutionary vs. experiential origins of animate monitoring observed in adulthood.

## 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.

## Ethics statement

The studies involving human participants were reviewed and approved by the University of Regina Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

JL, BR, RG, and SF contributed to conception and design of the studies. BR developed stimuli and designed the methodology for Experiment 1. JL performed the statistical analysis and wrote the first draft of the manuscript. BR, RG, and SF wrote sections of the manuscript. JL and BR contributed to manuscript revision. All authors read and approved the submitted version.

## Funding

This work was funded by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (awarded to JL). Experiment 1 was also developed while BR held a MITACS Globalink Internship to work in JL's lab.

## Acknowledgments

We are grateful to Shae Sackman, who programmed Experiment 1 in PsychoPy.

## 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.

## Publisher's note

## References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J. K. (2020). Gorilla in our midst: an online behavioral experiment builder. *Behav. Res. Methods* 52, 388–407. doi: 10.3758/s13428-019-01237-x

Blunt, J. R., and VanArsdall, J. E. (2021). Animacy and animate imagery improve retention in the method of loci among novice users. *Mem. Cogn.* 49, 1360–1369. doi: 10.3758/s13421-021-01175-0

Bonatti, L., Frot, E., Zangl, R., and Mehler, J. (2002). The human first hypothesis: identification of conspecifics and individuation of objects in the young infant. *Cogn. Psychol.* 44, 388–426. doi: 10.1006/cogp.2002.0779

Calvillo, D. P., and Jackson, R. E. (2014). Animacy, perceptual load, and inattentional blindness. *Psychon. Bull. Rev.* 21, 670–675. doi: 10.3758/s13423-013-0543-8

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., et al. (2012). The representation of biological classes in the human brain. *J. Neurosci.* 32, 2608–2618. doi: 10.1523/JNEUROSCI.5547-11.2012

Crouzet, S. M., Joubert, O. R., Thorpe, S. J., and Fabre-Thorpe, M. (2012). Animal detection precedes access to scene category. *PLoS One* 7:e51471. doi: 10.1371/journal.pone.0051471

Eddy, T. J., Gallup, G. G. J., and Povinelli, D. J. (1993). Attribution of cognitive states to animals: anthropomorphism in comparative perspective. *J. Soc. Issues* 49, 87–101. doi: 10.1111/j.1540-4560.1993.tb00910.x

Guerrero, G., and Calvillo, D. P. (2016). Animacy increases second target reporting in a rapid serial visual presentation task. *Psychon. Bull. Rev.* 23, 1832–1838. doi: 10.3758/s13423-016-1040-7

Hagen, T., Espeseth, T., and Laeng, B. (2018). Chasing animals with split attention: are animals prioritized in visual tracking? *i-Perception* 9, 1–35. doi: 10.1177/2041669518795932

Hagen, T., and Laeng, B. (2016). The change detection advantage for animals: an effect of ancestral priorities or progeny of experimental design? *i-Perception* 7, 1–17. doi: 10.1177/2041669516651366

Hagen, T., and Laeng, B. (2017). Animals do not induce or reduce attentional blinking, but they are reported more accurately in a rapid serial visual presentation task. *i-Perception* 8, 1–35. doi: 10.1177/2041669517735542

He, C., and Cheung, O. S. (2019). Category selectivity for animals and man-made objects: beyond low-and mid-level visual features. *J. Vis.* 19:22. doi: 10.1167/19.12.22

Jackson, R. E., and Calvillo, D. P. (2013). Evolutionary relevance facilitates visual information processing. *Evol. Psychol.* 11:147470491301100. doi: 10.1177/147470491301100506

Kellert, S. R. (1993). Values and perceptions of invertebrates. *Conserv. Biol.* 7, 845–855.

Knight, A. J. (2008). "Bats, snakes and spiders, oh my!" how aesthetic and negativistic attitudes, and other concepts predict support for species protection. *J. Environ. Psychol.* 28, 94–103. doi: 10.1016/j.jenvp.2007.10.001

Lipp, O. V., Derakshan, N., Waters, A. M., and Logies, S. (2004). Snakes and cats in the flower bed: fast detection is not specific to pictures of fear-relevant animals. *Emotion* 4, 233–250. doi: 10.1037/1528-3542.4.3.233

LoBue, V., and DeLoache, J. S. (2008). Detecting the snake in the grass: attention to fear-relevant stimuli by adults and young children. *Psychol. Sci.* 19, 284–289. doi: 10.1111/j.1467-9280.2008.02081.x

Long, B., Störmer, V. S., and Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *J. Vis.* 17:20. doi: 10.1167/17.6.20

Loucks, J., Verrett, K., and Reise, B. (2020). Animates engender robust memory representations in adults and young children. *Cognition* 201:104284. doi: 10.1016/j.cognition.2020.104284

Macé, M. J.-M., Joubert, O. R., Nespoulous, J.-L., and Fabre-Thorpe, M. (2009). The time-course of visual categorizations: you spot the animal faster than the bird. *PLoS One* 4:e5927. doi: 10.1371/journal.pone.0005927

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017). Remembering the living: episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

New, J. J., and German, T. C. (2015). Spiders at the cocktail party: an ancestral threat that surmounts inattentional blindness. *Evol. Hum. Behav.* 36, 165–173. doi: 10.1016/j.evolhumbehav.2014.08.004

Nguyen, H. B., and van Buren, B. (in press). May the force be against you: better visual sensitivity to speed changes opposite to gravity. *J. Exp. Psychol. Hum. Percept. Perform.*

Possidónio, C., Graça, J., Piazza, J., and Prada, M. (2019). Animal images database: validation of 120 images for human-animal studies. *Animals* 9:475. doi: 10.3390/ani9080475

Pratt, J., Radulescu, P. V., Guo, R. M., and Abrams, R. A. (2010). It's alive!: animate motion captures visual attention. *Psychol. Sci.* 21, 1724–1730. doi: 10.1177/0956797610387440

Ritchie, J. B., Zeman, A. A., Bosmans, J., Sun, S., Verhaegen, K., and Op de Beeck, H. P. (2021). Untangling the animacy organization of occipitotemporal cortex. *J. Neurosci.* 41, 7103–7119. doi: 10.1523/JNEUROSCI.2628-20.2021

Sha, L., Haxby, J. V., Abdi, H., Guntupalli, J. S., Oosterhof, N. N., Halchenko, Y. O., et al. (2015). The Animacy continuum in the human ventral vision pathway. *J. Cogn. Neurosci.* 27, 665–678. doi: 10.1162/jocn_a_00733

Snodgrass, J. G., and Vanderwort, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215.

Tisdell, C., Wilson, C., and Swarna Nantha, H. (2006). Public choice of species for the 'ark': phylogenetic similarity and preferred wildlife species for survival. *J. Nat. Conserv.* 14, 97–105. doi: 10.1016/j.jnc.2005.11.001

van Buren, B., and Scholl, B. J. (2017). Minds in motion in memory: enhanced spatial memory driven by the perceived animacy of simple shapes. *Cognition* 163, 87–92. doi: 10.1016/j.cognition.2017.02.006

Wichmann, F. A., Drewes, J., Rosas, P., and Gegenfurtner, K. R. (2010). Animal detection in natural scenes: critical features revisited. *J. Vis.* 10, 1–27. doi: 10.1167/10.4.6

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., and Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behav. Res. Methods* 42, 671–684. doi: 10.3758/BRM.42.3.671

Wu, C.-T., Crouzet, S. M., Thorpe, S. J., and Fabre-Thorpe, M. (2015). At 120 msec you can spot the animal but you don't yet know it's a dog. *J. Cogn. Neurosci.* 27, 141–149. doi: 10.1162/jocn_a_00701

Wurm, L. H. (2007). Danger and usefulness: an alternative framework for understanding rapid evaluation effects in perception? *Psychon. Bull. Rev.* 14, 1218–1225. doi: 10.3758/BF03193116

Yang, J., Wang, A., Yan, M., Zhu, Z., Chen, C., and Wang, Y. (2012). Distinct processing for pictures of animals and objects: evidence from eye movements. *Emotion* 12, 540–551. doi: 10.1037/a0026848

Zachariou, V., Del Giacco, A. C., Ungerleider, L. G., and Yue, X. (2018). Bottom-up processing of curvilinear visual features is sufficient for animate/inanimate object categorization. *J. Vis.* 18:3. doi: 10.1167/18.12.3

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 05 May 2023
DOI 10.3389/fpsyg.2023.1145884

# A cautionary note on the studies using the picture-word interference paradigm: the unwelcome consequences of the random use of "in/animates"

Ana Rita Sá-Leite[1,2]*, Montserrat Comesaña[3], Carlos Acuña-Fariña[4] and Isabel Fraga[1]

[1]Cognitive Processes and Behavior Research Group, Department of Social Psychology, Basic Psychology, and Methodology, University of Santiago de Compostela, Santiago de Compostela, Spain, [2]Institut für Romanische Sprachen und Literaturen, Goethe University Frankfurt, Frankfurt, Germany, [3]Psycholinguistics Research Line, CIPsi, School of Psychology, University of Minho, Braga, Portugal, [4]Cognitive Processes and Behavior Research Group, Department of English and German, University of Santiago de Compostela, Santiago de Compostela, Spain

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University, United States

REVIEWED BY
Patrick Bonin,
Université de Bourgogne, France
Ruth Corps,
Max Planck Institute for Psycholinguistics, Netherlands
Giacomo Spinelli,
University of Milano-Bicocca, Italy
Jens Bölte,
University of Münster, Germany

*CORRESPONDENCE
Ana Rita Sá-Leite
✉ anarita.saleite.dias@usc.es

RECEIVED 16 January 2023
ACCEPTED 13 April 2023
PUBLISHED 05 May 2023

CITATION
Sá-Leite AR, Comesaña M, Acuña-Fariña C and Fraga I (2023) A cautionary note on the studies using the picture-word interference paradigm: the unwelcome consequences of the random use of "in/animates".
*Front. Psychol.* 14:1145884.
doi: 10.3389/fpsyg.2023.1145884

The picture-word interference (PWI) paradigm allows us to delve into the process of lexical access in language production with great precision. It creates situations of interference between target pictures and superimposed distractor words that participants must consciously ignore to name the pictures. Yet, although the PWI paradigm has offered numerous insights at all levels of lexical representation, in this work we expose an extended lack of control regarding the variable animacy. Animacy has been shown to have a great impact on cognition, especially when it comes to the mechanisms of attention, which are highly biased toward animate entities to the detriment of inanimate objects. Furthermore, animate nouns have been shown to be semantically richer and prioritized during lexical access, with effects observable in multiple psycholinguistic tasks. Indeed, not only does the performance on a PWI task directly depend on the different stages of lexical access to nouns, but also attention has a fundamental role in it, as participants must focus on targets and ignore interfering distractors. We conducted a systematic review with the terms "picture-word interference paradigm" and "animacy" in the databases PsycInfo and Psychology Database. The search revealed that only 12 from a total of 193 PWI studies controlled for animacy, and only one considered it as a factor in the design. The remaining studies included animate and inanimate stimuli in their materials randomly, sometimes in a very disproportionate amount across conditions. We speculate about the possible impact of this uncontrolled variable mixing on many types of effects within the framework of multiple theories, namely the Animate Monitoring Hypothesis, the WEAVER++ model, and the Independent Network Model in an attempt to fuel the theoretical debate on this issue as well as the empirical research to turn speculations into knowledge.

KEYWORDS

**animacy, picture-word interference paradigm, lexical access, animate nouns, inanimate nouns, Animate Monitoring Hypothesis, language production**

# 1. Introduction

The picture-word interference (PWI) paradigm has served as a window for the study of lexical access at the level of semantics, grammar, and ortho-phonology. It is a variant of the Stroop task in which the classic Stroop effect (Stroop, 1935) is caused by the simultaneous or quasi-simultaneous presentation of images and distractor words that share the linguistic aspects under study (Lupker, 1979; Dell'Acqua et al., 2007; Shao et al., 2015; Starreveld and La Heij, 2017). When confronted with this type of paradigm, participants have to name pictures aloud using either a noun or a short noun phrase whilst ignoring a distractor word, usually a noun, that is either superimposed over the picture or presented auditorily (see Figure 1). The sharing of certain characteristics between target and distractor is expected to affect the response times of the participants. The effects that have been mostly explored are probably the ones concerning semantics and ortho-phonology. Indeed, when both nouns are from the same semantic field, interference is usually obtained (Cutting and Ferreira, 1999). For instance, the picture of an "apple" is generally named faster when paired with the distractor "table" than when paired with the distractor "orange". Yet, when the semantic relationship is associative, this is, when nouns tend to happen together in speech, such as "dog" and "bone", facilitation is obtained (Sailor et al., 2009; Geng et al., 2013). As for the so-called phonological facilitation effect, the sharing of the initial or final syllable or last letters/phonemes when stress patterns are controlled facilitates picture naming (e.g., Meyer and Schriefers, 1991; Melinger and Abdel Rahman, 2004; Ayora et al., 2011; Wilshire et al., 2016).

The PWI paradigm hence offers a versatile experimental option in which the interference created by the reading of a written noun during language production gives interesting insights about the way lexical access occurs. For instance, among many other contributions, it offers evidence about how neighboring lexical entries compete during the selection of a certain noun (Alario and Martín, 2010), how grammatical gender is accessed depending on the presence of an agreement context (Cubelli et al., 2005), or how cross-linguistic influence between languages occurs, including the possible interaction between spoken and sign languages (Giezen and Emmorey, 2016). The complementary use of other measuring techniques, namely electroencephalography or fMRI, further extends the evidence obtained with this paradigm by providing information about the temporal and neural organization of lexical encoding (Abel et al., 2009; Bürki et al., 2016). More recently, the PWI paradigm has been used as a resource to understand how lexical access is affected in its different levels in the context of normative aging (Lorenz et al., 2018) and a range of clinical conditions, namely second language impairment (de Hoog et al., 2015), apraxia of speech (Mailend and Maas, 2013), or aphasia (Hashimoto and Thompson, 2010).

However, the most ingenious element of the PWI paradigm, i.e., the use of language comprehension to study language production in its oral and written form (Bonin and Fayol, 2000; Bürki et al., 2019), involves a complex process whose outcomes can be misleading. In the PWI paradigm trials consist of a target that is both a picture and a noun, along with a written or oral distractor noun. This means that during the design of a PWI task, authors have to take into account multiple variables from three different stimuli (that may belong to different modalities, e.g., visual target and its associated noun plus orally presented distractor). Acknowledging this situation is critical because the outcome of a PWI task can be influenced by a great range of uncontrolled variables, not only of a psycholinguistic nature (e.g., the visual complexity of the images, the tone of the oral distractor). Indeed, such degree of complexity can be troublesome and has actually been regarded as a source of possible disruption in the observation of effects (for more detail, see the systematic/meta-analytic reviews of Bürki et al., 2020, and Sá-Leite et al., 2022). In this sense, one variable has been recently pointed out as possibly relevant: animacy (Sá-Leite et al., 2021).

Animacy may be understood as a gradient feature, a continuum in which humans are at one extreme and are followed by other categories such as mammals, other animals, plants, and objects (Dahl, 2000; The Animacy Hierarchy, Aissen, 2003). Across this continuum we can locate a cutoff point whose limits are often vague, but whose presence stablishes a cognitively relevant dichotomy between animate entities (e.g., elephants, jellyfish) and inanimate objects (e.g., tree, table).[1] Such a dichotomy has shown to have clear cognitive repercussions at attentional, memory-related, and psycholinguistic levels (Rakison and Poulin-Dubois, 2001; Nairne et al., 2017), with multiple studies having shown that there are different brain regions specialized in the processing of either animates or inanimates (e.g., Perani et al., 1995; Mahon et al., 2009; Proklova et al., 2016). The impact of this dichotomy can be observed even in toddlers, since it has shown to be a central organizing principle of children's cognitive experiences (e.g., Rostad et al., 2012).

Coming back to the PWI paradigm, different cognitive processes in which animacy may have an impact are involved in the resolution of the task, namely, the degree of attention given to target and distractor, or the number of semantic features associated with the target and distractor nouns to be accessed. Surprisingly, only Sá-Leite et al. (2021) seem to have considered animacy as a potential intervening factor in the PWI paradigm, specifically when considering the area of grammatical gender encoding. More precisely, the authors analyzed a typical PWI effect, the gender congruency effect, through the scope of animacy. The gender congruency effect consists in modulations on the response times of the participants depending on the gender congruency between the target and distractor nouns. Many authors have combined nouns of different gender (e.g., masculine and feminine) as target-distractor

---

[1] Following the definition proposed by Rakison and Poulin-Dubois (2001), animate entities are self-propelled, their line of trajectory is smooth, can cause action at a distance, their pattern of interaction is contingent and tend to be the agents of action (e.g., "elephant", "gardener", "fish"), whilst inanimate objects' motion is caused by external agents, their motion is usually irregular, they do not cause action at a distance by themselves, only by contact, their pattern of interaction is noncontingent and they are the recipient of actions rather than their agent (e.g., "tree", "table", "car"; for more details on this definition see Rakison and Poulin-Dubois, 2001).

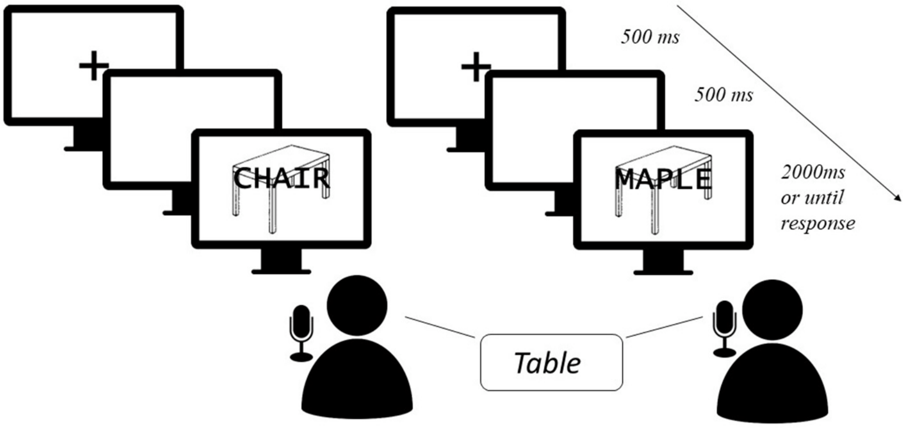


FIGURE 1
Example of a typical PWI task. In this example, distractors are presented written over the targets, rather than orally. Presentation of both targets and distractors is simultaneous, but different stimulus onset asynchronies have been tested in which the distractor can be presented either before or after the target (and, differently from a prime word, it is maintained on the screen along with the target). In the condition to the left, there is a semantic relationship between target "table" and distractor "chair"; in the condition to the right, there is phonological overlap between target "table" and distractor "maple". The image of a table was taken from the International Picture Naming Project (IPNP) database (Szekely et al., 2004).

pairs to check whether response times are affected depending on the activation and selection of one gender node or another. Yet, the outcome of these experiments is often mixed, with effects of facilitation being found in both directions (for both gender congruent pairs and incongruent pairs) and with many factors affecting the outcome (for a meta-analytic review, see Sá-Leite et al., 2022). When trying to better understand the gender congruency effect, Sá-Leite et al. (2021) manipulated the number of animate target pictures within the stimuli list and discovered that the effect was only present for the list featuring exclusively inanimate targets. The mere presence of 25% of animate targets prevented registering a significant effect, and the integrative analysis of all stimuli from all lists showed an effect of gender congruency restricted to inanimate targets which was smaller than the effect obtained when only the stimuli from the list with exclusively inanimate targets were considered. This led the authors to wonder what the effect of animacy might be in the activation of gender and to alert other authors regarding the overall role of animacy in the PWI paradigm.

In the present work, we discuss the possibility of animacy having an impact on the general outcome of a PWI paradigm across the different effects under study. Note that the nature of this work is hence speculative and intends nothing more than to nurture a theoretically motivated debate among researchers and hopefully inspire future studies that might turn speculation into possible evidence. With this aim, we first assess the cognitive impact of animacy on the mechanisms of attention, as well as the possible consequences that such an impact can have for the outcomes of a PWI paradigm, and then we do the same regarding the role of animacy in language processing. Afterwards, we present the reader with a systematic review in which we assess the animate status of targets and distractors within the PWI paradigm across all studies. As we will see, animacy has been almost completely ignored either as a confounding or as an independent variable.

## 1.1. The impact of animacy on attention

The most important theoretical framework on the link between animacy and attention was developed by New et al. (2007) under the name of "Animate Monitoring Hypothesis". The authors conducted a series of change-detection tasks in which both animate and inanimate stimuli were included in pictures of naturalistic scenes that suffered changes. More specifically, participants were rapidly presented with pairs of similar naturalistic scenes (250 ms each), but the second scene suffered changes regarding the presence or absence of animate and inanimate stimuli in relation to the first scene. The results showed that participants were faster and more likely to detect changes in animate than inanimate stimuli. The authors explained these results as a matter of ancestral priorities: the experience of humans living during millennia in hunter–gatherer environments would have derived in the ontogenetical development of an attentional advantage for animacy.

These ideas were supported by numerous studies (New et al., 2010; Yang et al., 2012; Altman et al., 2016; but see Hagen and Laeng, 2016; He and Cheung, 2019), even with toddlers (Hofrichter et al., 2021). Altman et al. (2016) study is especially interesting because they conducted change-detection tasks but analyzed in more detail not only the performance on specific stimuli but the influence of the presence of these stimuli on the detection of changes in others. The results not only showed the typical animacy advantage, but also showed that the detection of changes in inanimate stimuli is hampered by the presence of animate stimuli in the scene, but not vice-versa. This was true even when these animate stimuli remain unchanged and camouflaged. Similar outcomes were obtained in other paradigms. Visual search tasks also showed that animate entities are located faster than inanimate objects (Jackson and Calvillo, 2013). In particular, Calvillo and Hawkins (2016) observed that both threatening and non-threatening animate entities were more frequently detected than their inanimate object counterparts. Likewise, Guerrero and

Calvillo (2016) conducted an attentional blink task with animate and inanimate stimuli. Attentional blink refers to the phenomena by which participants fail to detect the second target in a task in which two target items are presented very closely in time (∼500 ms) in a series of rapid presentations. Their results were clear: animate targets were detected significantly more times than inanimate targets, and hence they were less prone to experience the phenomenon of attentional blindness. Ro et al. (2007) conducted multiple experiments in which participants searched for a green frame among blue frames. More specifically, they were asked to make speeded categorical decisions on stimuli presented within the green target frame (e.g., "was it food?"). Their results showed clearly that animate stimuli were attended preferentially (Ro et al., 2007). Animate stimuli are also detected more frequently than inanimate items in situations of both low and high perceptual load (Calvillo and Jackson, 2014), and animate motion is detected more quickly than inanimate motion (Pratt et al., 2010), even for newborn infants (Di Giorgio et al., 2017, 2021).

In sum, it seems undeniable that there is an attentional advantage when it comes to animate stimuli, whose presence seems to negatively affect the perception of inanimate stimuli. This can have important consequences for the outcomes of a PWI experiment, especially considering the inclusion of animate stimuli as target pictures. Thus, not only would the mechanisms of attention prioritize these stimuli over others, this is, over the distractors, but also the perception of a distractor would be especially hampered by the mere presence of an animate target (e.g., "elephant"). This could mean that the "distracting" role of the distractors is at least partially attenuated when animate targets are included. Since their potential to interfere decreases, competition between animate targets and inanimate distractors ("elephant" - "pencil") would produce smaller effect sizes than in the case of inanimate targets and distractors ("house" - "pencil"). On the other hand, when distractors are animate their ability to interfere should increase. On a pure attentional basis, an effect size should hence be greater for purely animate target-distractor pairs ("elephant" - "king") than for animate target/inanimate distractor pairs ("elephant" - "pencil"). Similarly, the increased ability to interfere of an animate distractor ("king") should create even greater competition with an inanimate target ("house"), powering even more the effects of competition in comparison to animate target/animate distractor pairs ("elephant"/"king") and inanimate target/inanimate distractor pairs ("house"/"pencil"). In any case, understanding the effect of animate nouns on the outcomes of PWI experiments is a topic that must also be addressed from the perspective of Psycholinguistics, as assessed in our next section.

## 1.2. How animacy impacts lexical access

Evidence suggests that animate words are somehow privileged during lexical access. For instance, animate targets are consistently named faster than inanimate targets (Laws and Neve, 1999; Laws et al., 2002). Even though we could think that this advantage at naming tasks could be explained by the attentional bias we discussed in the previous section, evidence shows that the existing differences in the performance of participants when considering animate vs. inanimate nouns cannot be exclusively explained on the basis of such a bias (see Xiao et al., 2016). In this sense, the reasons behind this advantage are usually related to the semantic content of animate nouns, as they are considered semantically "richer" than inanimates. This has been explained in multiple complementary ways that are backed up by numerous studies. Among these explanations, a commonly cited one is the theory that animate nouns present greater overlap among them in terms of semantic features, by which is meant that animates are overall more similar to each other than inanimates (e.g., Cree and McRae, 2003; Zannino et al., 2006; Davis et al., 2014; Xiao et al., 2016). Indeed, animates form categories of words that are semantically closer than those of inanimates and whose activation shows highly similar brain patterns (Xiao et al., 2016). Other studies suggest that animate nouns have more sensorimotor features than inanimate ones (Hargreaves et al., 2012a; Bonin et al., 2014; Heard et al., 2019). This is because animate nouns are related to more sensory and/or perceptual experiences than inanimates (Bonin et al., 2014). Indeed, words associated with more sensorimotor features have been found to be better recalled and recognized as well as processed faster as a function of their lexico-semantics (Hargreaves et al., 2012b; Hoffman et al., 2013). To be precise, animate nouns have shown to be consistently better recalled and recognized than inanimate ones (Nairne et al., 2013, 2017; VanArsdall et al., 2013, 2014). Regarding lexico-semantic encoding, an advantage of animate nouns over inanimate nouns has been found in semantic/animacy categorization tasks and lexical decision tasks as well (e.g., Becker et al., 1997; Radanović et al., 2016; Bonin et al., 2019).

Although many studies observe the advantage of animate nouns over inanimate nouns and obtain evidence regarding the semantic richness of animacy, another proposal offers an interesting and empirically supported view on the mechanism behind this advantage, i.e., why being semantically richer (greater overlap, greater number of sensorimotor features) translates into faster response times or better accuracy. It is based on the concept of *lexical accessibility*. If we define language production as an incremental process by which speakers can begin to generate speech once minimal input is made available (and hence word class, number, gender, phonological form, orthographic form, etc. are encoded incrementally and in parallel), how each piece of linguistic information is processed depends on its own relative accessibility. In this sense, information that is retrieved easily is given priority over information that is retrieved less easily. The relative ease of information retrieval depends on the baseline levels of activation of the information to be encoded. For instance, the relative accessibility of the elements of a syntactic structure has been shown to depend on whether or not these have been activated earlier through previous production or comprehension (Branigan et al., 2000). In this regard, animacy has been recognized as one of the factors that impact the relative accessibility of conceptual information (i.e., conceptual accessibility: the number of pathways available for retrieval, so that the more the pathways to the lexical concept, the faster its retrieval; Bock and Warren, 1985). Concepts that refer to animates would therefore be faster retrieved for production than those that refer to inanimates. This would be related to a semantic dimension they call *predicability* (Bock, 1987), this is, the number of conceptual relations an entity can establish. Animates can establish many more conceptual relations than inanimates. For instance, a dog can be born, bought, adopted, abandoned, it can die, sleep, communicate, bark, attack, run, walk, sit, break things, get hurt, be scared, etc., whilst the number of conceptual relations of a table or even a plant are much more reduced. Thus, animates not only tend to have more semantic overlap between each other and more sensorimotor features, but they also tend to be more predicable than inanimates and hence to enter in more syntactic relations (a tendency that can be broken for certain examples, e.g., bacteria, which is animate vs. doll, which is inanimate). This entails a higher conceptual accessibility. In sum, we could say that animates have a rich semantic content that contributes to an increase of their baseline activation level relative to inanimates. Therefore, they have a higher conceptual accessibility because they are prioritized by our system and are hence retrieved faster than inanimate nouns.

To better understand the consequences of the semantic peculiarities of animate nouns in lexical access and hence, in the PWI paradigm, let us first introduce the typical structure of lexical access as proposed by most models of language processing, including the influential Word-form Encoding by Activation and VERification model ++ (WEAVER++) (Roelofs, 1992, 1993; Levelt et al., 1999). In simple terms, three types of information must be encoded when accessing a noun: conceptual information related to meaning, grammatical-syntactic information, and form-based information (see Figure 2). All these three types of information are organized following three levels of lexical representation formed by nodes. Thus, for instance, the noun "table" is defined by a specific set of semantic features (e.g., "furniture", "four legs" "wood", "place to eat/work", etc.) represented by nodes at the conceptual level of representation. When producing "table", all these nodes are activated in the speaker's lexicon. This activation then spreads to the other levels of representation, namely, the grammatical-syntactic level in which features such as word class (noun) and number (singular) are activated and selected, and the orthographic-phonological level in which the phonological representation of the word is encoded (e.g., /'teɪb(ə)l/). This forms a pattern of activation that specifically represents the word to be produced, in this case, "table". Importantly, when the word "table" is not to be produced, it remains at a basal level of activation [lower level of activation than the required for production (or recognition) to occur]. This baseline level of activation can be higher the more we use the word (the basal level of activation depends on the frequency of use of a word - it is higher for "table" than for instance "cacophony") or depending on other factors impacting conceptual accessibility (e.g., animacy).

Now, note that in light of the literature that we have just reviewed on animacy, the semantic particularities of animate nouns may have direct repercussions on the effects obtained with the PWI paradigm, mainly those of semantic nature. Indeed, it is well known that the higher the semantic overlap between the two stimuli, the stronger the competition for selection between both lexical entries (as both are highly activated and reinforce each-other), and the greater the semantic relatedness effect. In this sense, whilst inanimate nouns do not share many semantic

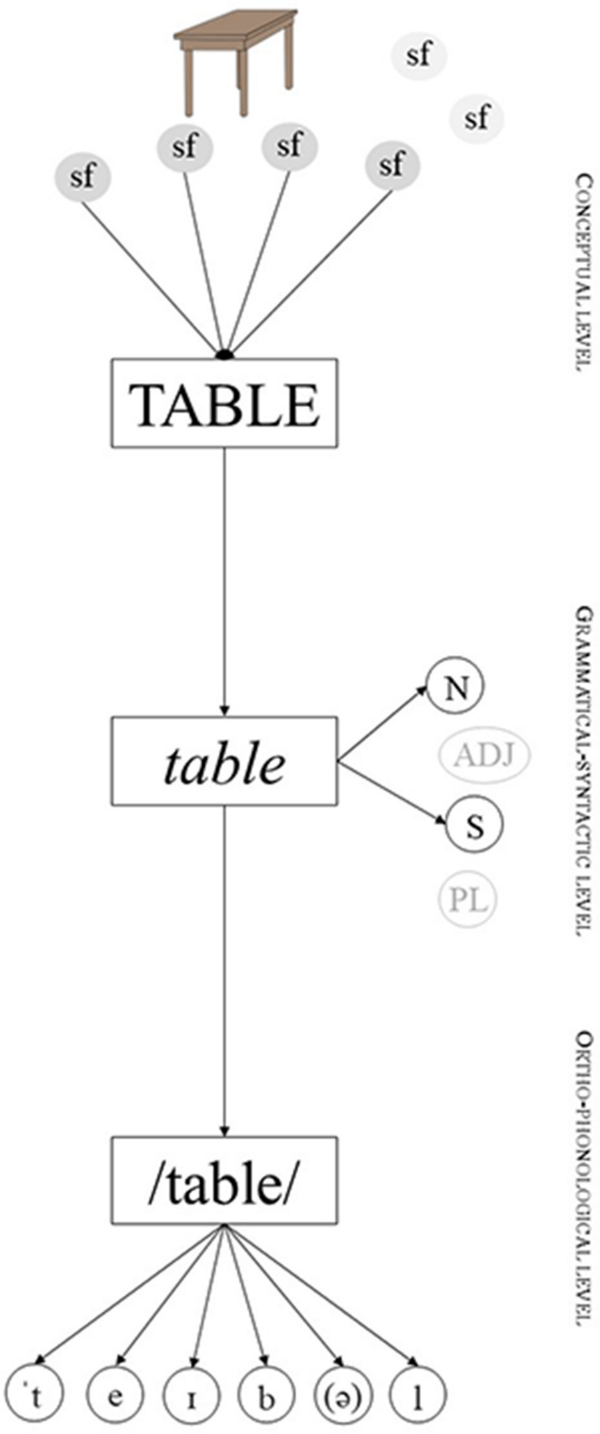


FIGURE 2
Simplified structure of lexical access during the production of the noun "table". Semantic features with lighter background at the conceptual level are meant to represent other features that are not related to "TABLE" and hence are not active. The same applies to lighter features at the grammatical-syntactic level. Arrows represent the flowing of activation. Sf, Semantic Feature; *N*, Noun; Adj, Adjective; S, Singular; PL, Plural. Based on the architecture proposed by WEAVER++. Adapted from Levelt et al. (1999).

features simply because of being inanimate (e.g., "car" and "pencil" are both inanimate but highly different), animates not only tend to have a higher number of semantic and sensorimotor features than inanimate nouns, but they also tend to share a vast amount of these features. Take, for instance, "gorilla" and "zebra": both are alive, both are animals, both are mammals, both have eyes, both have teeth, both are viviparous, both are hairy, both live in the outside, both are vegetarian, both have a heart, both feel pain, both have a nervous system, and so on. From this, we may speculate that the strongest effect of semantic competition within a PWI paradigm would be obtained when both target and distractor are animates. Importantly, the semantic overlap that animates *naturally* have could have general consequences for the outcome of a PWI paradigm, as facilitative and competitive effects of other types may behave differently for pure animate target-distractor pairs, in which the primary source of interaction between both entries is of semantic competition. Consequently, not only semantic relatedness effects should be analyzed through the scope of animacy, that is, taking into consideration possible differing size effects for semantically similar animates pairs in comparison to semantically similar inanimate pairs, but also pure animate pairs should be considered with caution when studying other type of effects.

The fact that animate nouns are semantically richer ultimately means that the number of semantic features to be processed at the conceptual stage of lexical access is greater than that of inanimate nouns. This points to the idea that our system may devote a great number of cognitive resources to animate nouns than inanimate nouns, particularly when considering the semantic level of lexical encoding. When animate nouns are being comprehended and produced, more resources would have to be devoted to process the semantic information of animates – perhaps to the detriment of other levels of encoding, something that has been called semantic prioritization (Sá-Leite et al., 2021). We could hence think that semantic prioritization can affect the amount and distribution of cognitive resources across the other levels of lexical encoding. The grammatical level of lexical encoding might be in a particularly fragile position. This is because it would be especially prone to suffer the possible consequences of high amounts of semantic information needing to be processed while the speed of lexical access still has to be increased for the sake of animacy itself. Indeed, the WEAVER++ model highlights the idea that grammatical information is selected (fully encoded) only when necessary – this leaves the door open to the idea that grammatical information does not have to be *always* selected. Another classic model of language production, the Independent Network model (Caramazza, 1997), remarks that grammatical encoding can be skipped as it is not a compulsory intermediate step between semantics and word-form encoding. Thus, on these views, language production can occur with information flowing directly from the semantic to the ortho-phonological level. In short, the idea that grammatical information can be skipped under certain conditions is not new and fits in with the evidence and theories on animacy. Indeed, it would seem as if our cognitive system devoted a higher number of cognitive resources to process the greater amount of semantic information of animate nouns, but we were still faster processing them in comparison to inanimate nouns. If grammatical processing can be skipped for specific reasons, animacy might perhaps be one of those reasons, so that semantics can be prioritized whilst maintaining lexical access especially fast. This theory is in line with the results of Sá-Leite et al. (2021), who systematically failed to observe effects based on gender processing for animate nouns, as if grammatical gender were not being encoded in these cases. Since grammatical gender is not an indispensable characteristic to be encoded when an agreement context is not present [which is the case of Sá-Leite et al. (2021) study, whose participants only produced bare nouns (i.e., with no adjectives or determiners whose form co-changes with the form of the noun)], once the system is overflown by the processing of animacy but is still forced to prioritize the processing of animate words, it would seem as if it dropped grammatical gender from the processing stage. Yet, this is a speculative hypothesis that should be further tested experimentally. As suggested by a Reviewer, one way of testing this would be designing an experimental situation in which the number of resources available was manipulated (e.g., a concurrent task manipulation). If the gender congruency effect requires a certain amount of available resources in order to emerge, the effect should disappear if the task is made more difficult with a concurrent task draining some of those resources away.

Finally, facilitative effects based on orthography and phonology could theoretically suffer variations due to the presence of animate stimuli as well. Note that facilitation here means that the distractor is speeding up the processing of the target, probably by contributing to the activation of the shared word-form attributes. Thus, if the target is animate ("baboon") and the ortho-phonologically related distractor is inanimate ("typhoon"), the facilitative effect produced by the inanimate distractor could be particularly small. This is because the target is already being processed quite fast and it is perhaps maintained at a high basal level of activation by our system. In fact, maybe we should consider the possibility of a ceiling effect for animate targets. The size of the phonological effect would hence decrease in comparison to pairs formed by an inanimate target. On the contrary, the combination of an inanimate target ("vanilla") and an animate distractor ("gorilla") would increase the size of the effect due to an accentuated facilitative role by the animate distractor, which would be highly and quickly activated. Now, if both target and distractor are animate, there is a potential confound with the strong effect of semantic relatedness we mentioned before, and hence we are not sure of how facilitative effects of phonological overlap would behave in this scenario.

## 1.3. Summing up the interference of animacy on the PWI

Taking into consideration the impact of animacy on the human attentional processes, in general, any effect of interaction between target and distractor could be influenced by animacy in the following way: animate targets will hamper the perception and interfering/facilitative role of inanimate distractors, diminishing the observed effects; animate distractors will have an increased interfering/facilitative role when paired with inanimate targets, increasing the observed effects. Purely animate pairs would hold stronger effects than animate target-inanimate distractor pairs.

On the other hand, we propose that the semantic particularities of animate nouns would have a main role on the outcome of a PWI paradigm, overruling any attentional bias in the case of studies exploring the semantic relatedness effect, grammatical effects, or to a certain degree, orthographic and phonological facilitation effects. Regarding the semantic relatedness effect, the strongest competition should occur between targets and distractors that are animate due to the semantic prioritization of both nouns and to the high degree of semantic similarity. Regarding grammatical effects, if the grammatical aspects at issue are skippable (such as gender with no agreement context), effects on their basis may not even be observed when one of the stimuli is animate due to semantic prioritization. Regarding orthographic and phonological facilitation effects, the main point to have into consideration is the fact that the overall degree and speed of activation of animate nouns is higher than that of inanimate nouns. Ultimately, this could mean that animate targets will benefit less from the presence of an ortho-phonologically similar distractor regardless of its animacy status; contrariwise, an inanimate target will benefit to a higher extent from the presence of an ortho-phonological similar animate distractor in comparison to a similar inanimate distractor.

Further speculating about how attentional and semantic factors interact to predict the outcome of a PWI task is out of the reach of a theoretical paper such as this one. The same applies when trying to understand how the semantic factors of animacy would affect other linguistic effects, such as the semantic association effect (Brooks et al., 2014), the word-frequency effect (Mulatti et al., 2015), or the compound effect (Lorenz and Zwitserlood, 2016).

## 2. The present study

It seems clear that animacy is an important factor in the organization, structuring, and functioning of our cognitive system, with important attentional repercussions as well as consequences at different levels of language processing. It is therefore possible that animate stimuli pose a source of disruption in the outcomes of experiments done with the PWI paradigm, and they might have an interesting role if considered within the experimental design, especially for effects of a semantic nature. Yet, the question remains: how many studies using this paradigm have controlled or considered animacy? In the next section, we will present a systematic review of this matter in detail.

### 2.1. Systematic search

We conducted a search with the keyword "picture-word interference paradigm" by itself as well as combined with the keyword "animacy" in the databases PsycInfo and Psychology Database. The whole process of systematic search is summarized in the PRISMA graph presented in Figure 3. Our search cast a total of 326 results. After removing duplicates with the software RefWorks$^{®}$ ($n = 66$) a total of 260 studies remained. We checked for availability of the full text of all the studies. When we lacked the permission to access the full text online, authors were contacted mainly through ResearchGate (e.g., Bürki and Madec, 2022). We could not find or obtain upon request the full text of one of the studies (Collina et al., 2014). All the remaining 259 studies were inspected and the next criteria for inclusion were applied:

a. The study makes an experimental contribution (e.g., Mahon and Caramazza, 2009; Sá-Leite et al., 2019, 2020; Fuhrmeister and Bürki, 2022; i.e., it is not a commentary, a theoretical proposal, a systematic review, or a meta-analysis).
b. The study includes at least one PWI task which is not a variation of the classic task (e.g., using a post-cue naming paradigm,[2] Hocking et al., 2010; Mädebach et al., 2018; using a picture-sound interference paradigm).[3]
c. The study uses nouns or noun phrases that include nouns as either targets or distractors or both (e.g., they do not use verbs as both targets and distractors; Lüttmann et al., 2011).
d. The study is written in English (Yu and Shu, 2003; e.g., a few of the studies were exclusively written in Chinese: see Qingfang and Yufang, 2004).

After applying the criteria for inclusion, 114 studies were disregarded. The inspection of the reference list of each one of them ($n = 145$) allowed us to obtain 29 new studies not contemplated in the initial search that complied with the criteria of inclusion (check the Supplementary Materials for the full list of additional studies). By considering a previous systematic review on the Stroop task and the PWI paradigm by MacLeod (1991), we obtained a total of 19 PWI studies that fitted our criteria for inclusion and were published before 1990. A total of 193 studies were kept in the final sample.

### 2.2. Inspection of animacy

All 193 works were inspected independently by two evaluators with knowledge of gender processing and animacy. A description of the works was made according to: (a) the effect being explored; (b) whether or not animacy is explicitly mentioned and considered; (c) whether or not animacy is considered as a potential confounding factor; (d) examples of animate stimuli target-distractor pairs. To do so, both evaluators first assessed whether or not the paper considered animacy theoretically in the Introduction; then, regardless of whether the study did or did not mention animacy in that section, they assessed whether the paper considered animacy in the Method, namely in the control of the materials, the design, or the results. To conclude, the Discussion was assessed in case the

---

[2] Within this variation of the PWI paradigm, both target and distractor are pictures, and the target picture is cued subsequently to its presentation along with the distractor picture. For instance, the target could be presented in green tones, and the distractor in blue tones. A subsequent cue (e.g., a green dot) tells the participant which picture to name.

[3] In the picture-sound interference paradigm, the distractor is a sound rather than a written word. For instance, a semantic interference effect within this paradigm would arise from presenting the picture of a dog with the sound of a dog barking in comparison to the sound of a car engine.

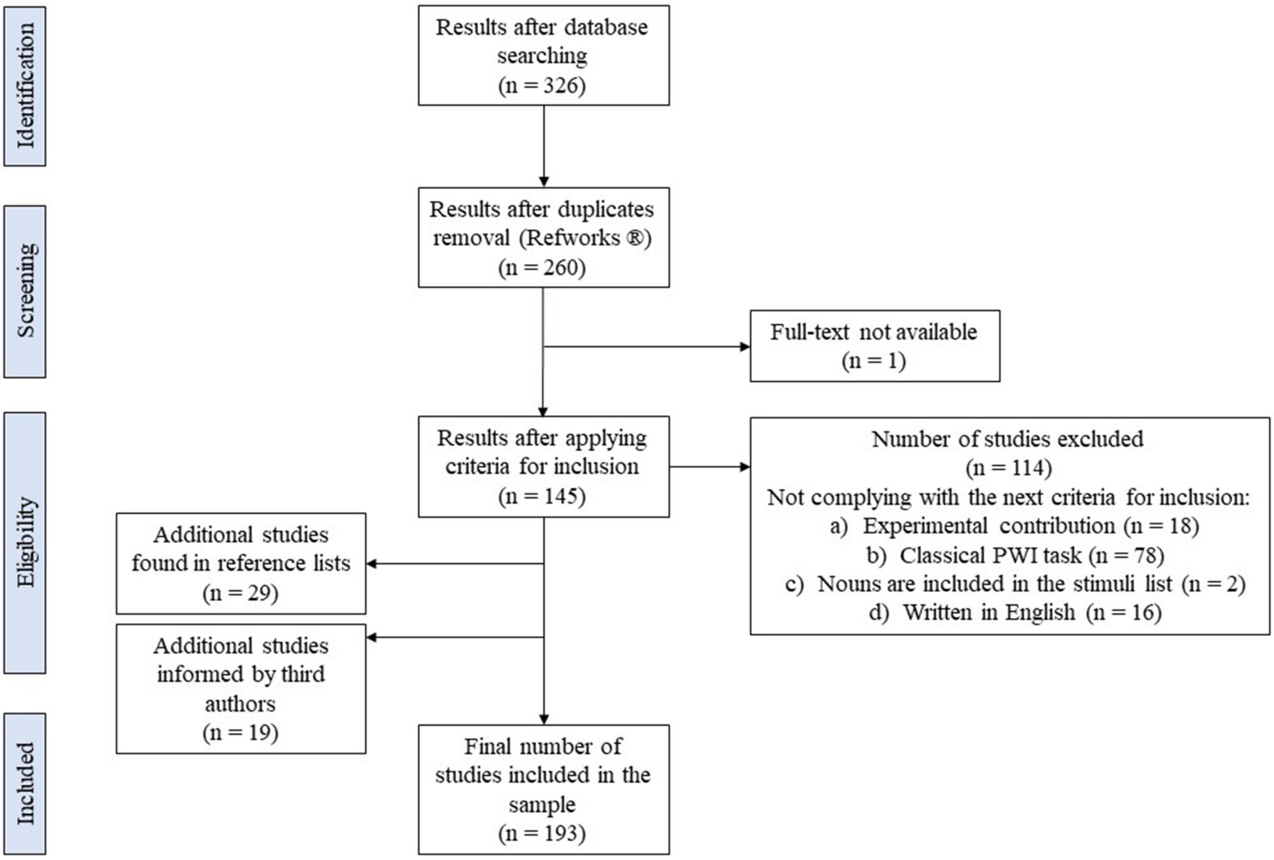


FIGURE 3
Structure of the search according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart of the literature search (Moher et al., 2009).

authors decided to consider it at the end as a *post-hoc* explanation of the results, a limitation, or a future research step. Finally, they inspected the stimuli list when available (either within the paper or as an online appendix). When the stimuli list was unavailable, the evaluators checked the examples provided in the Method section. The assessments of both evaluators were compared. When any of the information did not match, the work in question was checked again. The details of every study are collected in Table 1 available at https://doi.org/10.17605/OSF.IO/CJR37.

## 2.3. Summary and description of the studies

The close inspection revealed that from a total of 193 works, only 12 did not mix animate and inanimate stimuli randomly within the target and distractor pairs and the experimental conditions (Ehri, 1976; Guttentag and Haith, 1979; Schnur et al., 2006; Foucart et al., 2010; Muehlhaus et al., 2013; Hwang and Kaiser, 2014; Dank and Deutsch, 2015; DiBattista, 2015; Shin, 2016; Bürki et al., 2019; Deutsch and Dank, 2019; Sá-Leite et al., 2021). Additionally, among these 12, only three of them explicitly stated that they controlled animacy (Foucart et al., 2010; Shin, 2016; Bürki et al., 2019; i.e., "only inanimate stimuli were used") and only one included animacy as a factor to check its impact on the effect sizes (Sá-Leite et al., 2021). Importantly, only Shin (2016) and Sá-Leite et al. (2021) explicitly mention and discuss animacy theoretically as a potential factor affecting the results. Four studies controlled or manipulated animacy as a factor, but this was not due to animacy itself but because animacy is at the core of certain grammatical cut-offs that happened to be the object of study (Hwang and Kaiser, 2014; Dank and Deutsch, 2015; Deutsch and Dank, 2019; e.g., natural vs. grammatical gender). Guttentag and Haith (1979) decided to only use animate nouns and distractors to study the memory capacity of their participants using the PWI paradigm – they did not mention animacy, though. Ehri (1976), who studied the general mechanisms of attentional interference through the PWI paradigm, was cautious when deciding to use only pairs of semantically related targets and distractors, so that animals were only paired with animals, to avoid any type of confounding effect. Finally, among the 12 studies there is one Thesis (DiBattista, 2015) which considered the impact of animacy theoretically on certain effects obtained with other types of tasks, but not within the PWI paradigm itself – however, the author included only inanimate stimuli in the experiment featuring the PWI paradigm. None of the studies inspected the possible role of the different degrees of animacy (i.e., the animacy hierarchy).

The remaining 181 studies failed to control (explicitly or not), consider, or even mention animacy. All these studies hence feature uncontrolled animate/inanimate target-distractor pairs e.g., when studying phonological overlap and grammatical gender, target "leopard" with distractor "brother-in-law" and target "leopard" with distractor "clay" seen as comparable as target "pear" with distractors "tie" and "beaver", Bürki et al., 2016). Around 50% of them explore effects of a semantic nature (e.g., semantic relatedness, semantic association, etc.) and still fail to take animacy into account, even though animacy has a direct impact on the semantic richness and the conceptual availability of the noun, as well as on the number of shared features between nouns of the same semantic category. For instance, Rosinski (1977) studies the semantic relatedness effect and considers two categories, animals and household objects, but fails to take the opportunity to check whether the category animals holds stronger effects of semantic interference than that of household objects. Overall, situations arise in which the authors assume the size of the effect of semantic relatedness to be the same for pairs such as "chair$_{(\text{target})}$-sofa$_{(\text{distractor})}$" and "frog-cat" (Collina et al., 2013). They also compare semantically related pairs to unrelated pairs even if animacy is probably undersizing/oversizing the interaction between target and distractor in the unrelated condition. For instance, in one study the pair formed by "frog" and "pen" was compared to the pair formed by "chair" and "child" (Collina et al., 2013). Certainly, despite the fact that both pairs are from the unrelated condition, the attention given to "frog" is probably higher than that given to "chair" and the interference from "child" is probably far higher than the interference from "pen". In this line, the intrusion of animates of different degrees of animacy within the condition of semantic unrelatedness is quite widespread. We can find semantically unrelated pairs such as "pear-sheep" or "mouse-brush" that are put at the same level as "dog-truck" or "bench-wolf" (Melinger and Abdel Rahman, 2004; Janssen et al., 2008; Jerger et al., 2013; Krott et al., 2019; Jescheniak et al., 2020). In terms of comparisons of effect sizes across semantic conditions, another interesting example is that of De Zubicaray et al. (2013). The authors compare the size effects between conditions of semantic relatedness and semantic association. Yet, they do this without taking into account animacy, which derives in situations in which the *semantic association* is made between an animate and an inanimate, such as "baby" and "pram", and compared to the *semantic relatedness* between two animates such as "baby" and "priest". They thus do not ponder the possibility that rather than differences between types of semantic relations, they may be also observing differences due to the animacy of the distractor. The same applies when they compare the semantic association between an inanimate (e.g., "cave") and an animate ("bat") to the semantic relatedness of two inanimate nouns ("cave" and "sea").

Other types of effects we highlighted as especially prone to suffer from interference due to animacy were of a grammatical nature. From the 193 studies, 40 studied some type of grammatical effect (grammatical class effect, case status effect, classifier and gender congruency effect, countability congruency effect), and 6 belong to the 10 that did not mix animate and inanimate stimuli in an uncontrolled manner (still, 3 of them happened to control animacy due to its role as a cut-off point for their object of study, and not due to animacy itself). Therefore, among these studies, we observe situations in which the authors compare conditions with a different number of animate stimuli (gender congruency 13, gender incongruency 7, Schiller and Caramazza, 2003; gender incongruency + semantic relatedness, 4; gender incongruency + semantic relatedness, 0), as well as many random pairs, such as "ax-emperor" belonging to one condition but "ax-rhythm" to the opposite. A curious example is the interesting study by Fieder et al. (2018), which explores the processing of count and mass nouns but assumes that the incongruent (in terms of countability) pairs "kings-yogurt" and "nuns-sand", to be the same as "pedals-vinegar" (Fieder et al., 2018).

Regarding effects of orthographic and phonological facilitation, besides the random inclusion of mixed pairs, it is interesting to see the use of animate pure pairs without considering the fact that there is a great semantic overlap in that case and hence the effect of phonological facilitation is probably interacting with an effect of semantic interference. For instance, pairs such as "pig" (target) and "rabbit" (Costa et al., 2003), "penguin" and "farmer" ("penguin" was paired with "pizza" in the phonologically unrelated condition; Ayora et al., 2011), "dog" and "goat" ("dog" was paired with "dot" in the related condition, Roelofs and Verhoef (2006), and so on. In addition to the studies exploring semantic, orthographic, phonological, and grammatical effects, the rest of the literature also presents many examples of mixed animate/inanimate pairs. For instance, Dhooge and Hartsuiker (2011) explore frequency effects, and hence compare the times to name low and high frequency pictures depending on whether they are paired with either low or high frequency distractors. Yet, they do not control for animacy, which means that they have combinations such as low frequency animate pictures ("fox") paired with high frequency animate distractors ("king"), and high frequency inanimate pictures ("stone") also paired with the same high frequency animate distractors ("king"). In the case exemplified between parentheses, "fox" may be more protected against the interference generated by "king" than inanimate nouns of low frequency, and thus the comparison with "stone" is not as precise as it should. Also, "king", as a human animate noun, would be an especially interfering distractor in comparison to inanimate nouns. In this line, Geng et al. (2014) also assess the naming times depending on whether the targets are paired with high or low frequency distractors. However, by not considering animacy, the authors create situations in which the high frequency distractor for "drum" is "woman", but for "hat" it is "air" and for "pig" it is "name"; all this whilst the low frequency distractor for "drum" is "bacon", for "hat" it is "owl" and for "pig" it is "bale" (among many other examples). Likewise, Schnur et al. (2006) asked their participants to use short sentences to name pictures in which different people were performing different actions. The authors were particularly interested in the effect of phonological relatedness that could emerge between the verbs used in the target sentences and the distractor nouns. However, even though all their pictures depicted humans and all their distractors were inanimate nouns (e.g., dam, dish, jug, rust...), a possible undersizing of the expected effect due to the animacy of the targets was not discussed. Furthermore, some studies used pseudowords as distractors, and assumed the potential interfering role of these strings of letters to be the same both when paired with animate ("farmer", "mouse") and inanimate targets ("house", "needle", e.g., Oppermann et al., 2008; Brooks et al., 2015). It is also interesting to see how certain studies exploring the perceptive and attentional mechanisms of humans by manipulating the type, position, and other characteristics of the distractor, also disregarded animacy and did not consider differential effects depending on the animacy of targets and distractors (e.g., Underwood, 1976; Toma and Tsao, 1985).

Finally, out of the 181 studies that did not consider or control animacy, 11 involve major works (doctoral theses) and, importantly, 17 explore clinical or aging populations, whose results can have important repercussions regarding our understanding of these conditions. Indeed, it should not be a problem if populations are compared but still confronted with the same set of stimuli (e.g., deaf children, children with Specific Language Impairment, and hard-of-hearing children, de Hoog et al., 2015). Yet, by unknowingly obtaining underestimated or overestimated effects we may be missing precision when judging the specific capabilities of each group - also, semantic processing may be somehow especially impaired in certain clinical populations, which may have a special impact on animacy. In this sense, a relevant example is that of Durfee (2019), who assesses language impairment after a stroke through the size of the effects of semantic relatedness and phonological overlap without considering animacy. This can therefore lead to quite imprecise results, especially if both effects are compared to determine the affection of each of the language abilities.

## 3. Discussion

In the present study, we discussed the possible role of animacy in the outcomes of experiments done using a specific task: the PWI paradigm. Indeed, the PWI paradigm is a vastly used task for the assessment of language processing at the different levels of lexical encoding. However, it involves a complex interaction between comprehension and production processes in which three stimuli of different types (target picture, target noun, written or oral distractor noun), can be sources for confounding variables. We theorized that animacy might be especially relevant both as a possible confounding variable and an independent variable for the outcome of a PWI paradigm because it has great repercussions on (a) the monitorization of attention and hence on the degree of attention given to the target picture and the distractor noun, and (b) on language processing itself, by determining the amount and overlapping of semantic information to be processed and the distribution and number of resources implied in each stage of lexical access. Yet, our systematic review of the literature using the PWI task has shown that animacy has been mostly neglected when it comes to both the control of the materials and its direct study as an independent variable. Of a total of 193 studies reviewed, only 12 have managed to control for animacy. Three of them have done so explicitly, four of them have done it indirectly because animacy is at the base of the cutoff points for the grammatical structures under study, two of them have done it indirectly by controlling the category of "animals", one of them actually included animacy as a design factor, and the other two happened to use only inanimate stimuli. The remaining have mixed together animate and inanimate stimuli from different points of the animacy continuum, without any regard for the experimental conditions. Among them, a few are of clinical orientation and thus establish conclusions on the language capabilities of populations with clinical conditions affecting language and lexical access.

The apparent absence of animacy in such a vast portion of the literature using the PWI paradigm comes as rather surprising to us. This is because, as hypothesized in the Introduction, the inclusion of animate stimuli can overestimate or underestimate the obtained effects within the PWI paradigm, and can also give interesting insights regarding lexical access in terms of semantic processing, the mandatory processing of grammatical features, or the distribution of resources during the different stages of lexical encoding. In attentional terms, animacy may have a role on the general outcome of a PWI paradigm by maximizing or minimizing the interfering role of the distractors. From the point of view of language processing, animacy is quite interesting as a factor *per se* since its semantic peculiarities might affect specific effects differentially and exploring it may give researchers insights about the way cognitive resources are distributed across the different stages of lexical access. In this sense, regarding the semantic interference effect, the competition between pure animate target-distractor pairs might be especially strong in comparison to semantically similar inanimate pairs due to a greater number of semantic and sensorimotor features and a greater overlap between them. Yet, none of the reviewed studies has considered this. On the other hand, semantic prioritization may somehow affect how cognitive resources are distributed at the other levels of lexical encoding. Of interest is the impact that animacy may have at the level of grammatical encoding. This is because, in line with previous models of lexical access, grammatical encoding has been said to be skippable (Caramazza, 1997; Levelt et al., 1999), thus putting the effects of a grammatical nature in a particularly fragile position. More exactly, if cognitive resources are directed to the conceptual level of encoding and lexical access has to be quick for the sake of animacy, the skipping of grammatical encoding, when possible, might be a useful way of effectively distributing and preserving cognitive resources while speeding up word processing. Finally, effects of orthographic and phonological facilitation may also be affected by the degree and speed of activation of animate nouns. More specifically, the processing of an animate target noun would perhaps not benefit so much from the presence of an ortho-phonologically similar distractor, but an inanimate noun would benefit to a higher extent from the presence of an ortho-phonological similar animate distractor than of an inanimate distractor. Importantly, in lexical terms, the semantic characteristics of animates may affect other specific effects of different lexical nature, especially when including pure animate target-distractor pairs to study other effects such as those of orthographic and phonological facilitation, in which the semantic interference effect coming from the overlap of animate characteristics in the pure animate target-distractor pairs should not be ignored. Still, these are all mere speculations raised to create debate among researchers and which necessarily would have to be put under test. Should authors test any of these ideas, they would inevitably also have to carefully consider whether the results of PWI experiments inform us on the deployment of attentional mechanisms or on semantic prioritization, or both (and, if both, when and how). A disentanglement between both type of impacts (attentional vs. linguistic) could be better explored with additional techniques, mainly fMRI, which could show the differential activation of areas related to attention and linguistic processing among the different types of stimuli. Electroencephalographic techniques would also be interesting to have an idea of the time-course of lexical access, for instance to detect effects of semantic interference in pure animate pairs when studying effects of orthographic and phonological nature.

In sum, we hope that this work captures the attention of researchers when it comes to animacy, as we believe there is enough empirical evidence to think that animacy might have the potential to be a fruitful variable for the PWI paradigm. Of course, in terms of experimental control, we are aware that neglecting animacy as a confounding variable in some cases probably does not have a great impact on the results of a study and the conclusions to be drawn. Some imprecision from mixing animate and inanimate stimuli in an uncontrolled manner might arise, but this imprecision still does not change the final results. After all, for manipulations in which the same stimuli are used in different conditions, the same amount of animates are present of each condition, and the research interest is on the impact of those conditions on the interference effect. Still, we believe that the present work makes a point that is important in the current state-of-the-art: highlighting that most PWI studies are ignoring a variable that has enough theoretical foundation to be considered of high interest for the PWI task due to the characteristics of the paradigm itself.

## Author contributions

AS-L: idea, theoretical background, writing, discussion of the results, systematic review, and funding. MC: revision and systematic review. IF: revision and funding. CA-F: revision, editing, and funding. All authors contributed to the article and approved the submitted version.

## Funding

This project was funded by the Consellería de Cultura, Educación e Ordenación Universitaria, Xunta de Galicia (ED481B-2022-041), by Axencia Galega de Innovación and Consellería de Economía, Industria e Innovación, Xunta de Galicia (ED431B 2019/2020 and ED431B 2022/19), and also by the Ministerio de Ciencia e Innovación (PID2019-110583GB-I00). This study was partially conducted at the Psychology Research Centre (CIPsi), University of Minho and supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (UID/01662/2020), and grant PD/BD/52396-2013.

## 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.

## Publisher's note

## References

Abel, S., Dressel, K., Bitzer, R., Kümmerer, D., Mader, I., Weiller, C., et al. (2009). The separation of processing stages in a lexical interference fMRI-paradigm. *NeuroImage* 44, 1113–1124. doi: 10.1016/j.neuroimage.2008.10.018

Aissen, J. (2003). Differential object marking: iconicity vs. economy. *Nat. Lang. Ling. Theor.* 21, 435–483. doi: 10.1023/A:1024109008573

Alario, F. X., and Martín, F. M. (2010). On the origin of the "cumulative semantic inhibition" effect. *Memory Cognit.* 38, 57–66. doi: 10.3758/MC.38.1.57

Altman, M. N., Khislavsky, A. L., Coverdale, M. E., and Gilger, J. W. (2016). Adaptive attention: how preference for animacy impacts change detection. *Evol. Hum. Behav.* 37, 303–314. doi: 10.1016/j.evolhumbehav.2016.01.006

Ayora, P., Peressotti, F., Alario, F. X., Mulatti, C., Pluchino, P., Job, R., et al. (2011). What phonological facilitation tells about semantic interference: a dual-task study. *Front. Psychol.* 2, 57. doi: 10.3389/fpsyg.2011.00057

Becker, S., Moscovitch, M., Behrmann, M., and Joordens, S. (1997). Long-term semantic priming: a computational account and empirical evidence. Journal of experimental psychology. *Learn. Mem. Cognit.* 23, 1059–1082. doi: 10.1037/0278-7393.23.5.1059

Bock, J. K., and Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21, 47–67. doi: 10.1016/0010-0277(85)90023-X

Bock, K. (1987). An effect of the accessibility of word forms on sentence structures. *J. Memory Lang.* 26, 119–137. doi: 10.1016/0749-596X(87)90120-3

Bonin, P., and Fayol, M. (2000). Writing words from pictures: what representations are activated and when? *Memory Cognit.* 28, 677–689. doi: 10.3758/bf03201257

Bonin, P., Gelin, M., and Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Memory Cogn.* 42, 370–382. doi: 10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Dioux, V., and Méot, A. (2019). "It is alive!" Evidence for animacy effects in semantic categorization and lexical decision. *Appl. Psycholing.* 40, 965–985. doi: 10.1017/S0142716419000092

Branigan, H. P., Pickering, M. J., Stewart, A. J., and McLean, J. F. (2000). Syntactic priming in spoken production: linguistic and temporal interference. *Memory Cognit.* 28, 1297–1302. doi: 10.3758/BF03211830

Brooks, P. J., Seiger-Gardner, L., Obeid, R., and MacWhinney, B. (2015). Phonological priming with nonwords in children with and without specific language impairment. *J. Speech Lang. Hearing Res.* 58, 1210–1223. doi: 10.1044/2015_JSLHR-L-14-0212

Brooks, P. J., Seiger-Gardner, L., and Sailor, K. (2014). Contrasting effects of associates and coordinates in children with and without language impairment: a picture-word interference study. *Appl. Psycholing.* 35, 515–545. doi: 10.1017/S0142716412000495

Bürki, A., Besana, T., Degiorgi, G., Gilbert, R., and Alario, F. X. (2019). Representation and selection of determiners with phonological variants. *J. Exp. Psychol. Learn. Memory Cognit.* 45, 1287–1315. doi: 10.1037/xlm0000643

Bürki, A., Elbuy, S., Madec, S., and Vasishth, S. (2020). What did we learn from forty years of research on semantic interference? A Bayesian meta-analysis. *J. Memory Lang.* 114, 104125. doi: 10.1016/j.jml.2020.104125

Bürki, A., and Madec, S. (2022). Picture-word interference in language production studies: Exploring the roles of attention and processing times. *J. Exp. Psychol. Learn. Memory Cognit.* 48, 1019–1046. doi: 10.1037/xlm0001098

Bürki, A., Sadat, J., Dubarry, A. S., and Alario, F. X. (2016). Sequential processing during noun phrase production. *Cognition* 146, 90–99. doi: 10.1016/j.cognition.2015.09.002

Calvillo, D. P., and Hawkins, W. C. (2016). Animate objects are detected more frequently than inanimate objects in inattentional blindness tasks independently of threat. *J. Gen. Psychol.* 143, 101–115. doi: 10.1080/00221309.2016.1163249

Calvillo, D. P., and Jackson, R. E. (2014). Animacy, perceptual load, and inattentional blindness. *Psycho. Bullet. Rev.* 21, 670–675. doi: 10.3758/s13423-013-0543-8

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognit. Neuropsychol.* 14, 177–208. doi: 10.1080/026432997381664

Collina, S., and Simone, D., F. (2014). Polarity effects and the role of strategies in the picture-word interference paradigm. *Riv. Psicoling. Applicata* 14, 55–71.

Collina, S., Tabossi, P., and Simone, F. (2013). Word production and the picture-word interference paradigm: the role of learning. *J. Psycholing. Res.* 42, 461–473. doi: 10.1007/s10936-012-9229-z

Costa, A., Colome, À., Gómez, O., and Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Biling. Lang. Cognit.* 6, 167–179. doi: 10.1017/S1366728903001111

Cree, G. S., and McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J. Exp. Psychol. General* 132, 163–201. doi: 10.1037/0096-3445.132.2.163

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., and Job, R. (2005). Grammatical gender is selected in bare noun production: evidence from the picture–word interference paradigm. *J. Memory Lang.* 53, 42–59. doi: 10.1016/j.jml.2005.02.007

Cutting, J. C., and Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *J. Exp. Psychol. Learning Mem. Cognit.* 25, 318–344. doi: 10.1037/0278-7393.25.2.318

Dahl, Ö. (2000). *Animacy and the Notion of Semantic Gender in Grammar and Cognition*. Berlin: De Gruyter Mouton, 99–116.

Dank, M., and Deutsch, A. (2015). Morphological structure governs the process of accessing grammatical gender in the course of production. *The Mental Lexicon* 10, 186–220. doi: 10.1075/ml.10.2.02dan

Davis, T., Xue, G., Love, B. C., Preston, A. R., and Poldrack, R. A. (2014). Global neural pattern similarity as a common basis for categorization and recognition memory. *J. Neurosci. Off. J. Soc. Neurosci.* 34, 7472–7484. doi: 10.1523/JNEUROSCI.3376-13.2014

de Hoog, D., Langereis, B. E., van Weerdenburg, M. C., Knoors, M. H., and Verhoeven, L. (2015). Lexical access in children with hearing loss or specific language impairment, using the cross-modal picture–word interference paradigm. *Res. Dev. Disab.* 37, 81–94. doi: 10.1016/j.ridd.2014.11.007

De Zubicaray, D., Hansen, G. I. S., and McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *J. Exp. Psychol. General* 142, 131–142. doi: 10.1037/a0028717

Dell'Acqua, R., Job, R., Peressotti, F., and Pascali, A. (2007). The picture-word interference effect is not a Stroop effect. *Psych. Bullet. Rev.* 14, 717–722. doi: 10.3758/BF03196827

Deutsch, A., and Dank, M. (2019). Morphological structure mediates the notional meaning of gender marking: evidence from the gender-congruency effect in Hebrew speech production. *Q. J. Exp. Psychol.* 72, 389–402. doi: 10.1177/1747021818757942

Dhooge, E., and Hartsuiker, R. J. (2011). The distractor frequency effect in a delayed picture-word interference task: further evidence for a late locus of distractor exclusion. *Psych. Bullet. Rev.* 18, 116–22. doi: 10.3758/s13423-010-0026-0

Di Giorgio, D., Lunghi, E., and Simion, M. F., and Vallortigara, G. (2017). Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Dev. Sci.* 20, 12394. doi: 10.1111/desc.12394

Di Giorgio, D., Lunghi, E., and Vallortigara, M. G., and Simion, F. (2021). Newborns' sensitivity to speed changes as a building block for animacy perception. *Sci. Rep.* 11, 542. doi: 10.1038/s41598-020-79451-3

DiBattista, A. (2015). *Grammatical Encoding in the Noun Phrase: Effects of Constituent Similarity and Grammatical Class Available from APA PsycInfo®*. Available online at: https://www.proquest.com/dissertations-theses/grammatical-encoding-noun-phrase-effects/docview/1707078555/se-2?accountid=14542 (accessed May 2022).

Durfee, A. Z. (2019). *Are Input and Output Language Networks Linked?: Evidence from the Verification Task Paradigm and Its Role in Assessing Language Impairment After Stroke (Ph.D.)*. Available from Psychology Database. Available online at: https://www.proquest.com/dissertations-theses/are-input-output-language-networks-linked/docview/2478605065/se-2?accountid=14542

Ehri, L. C. (1976). Do words really interfere in naming pictures? *Child Dev.* 47, 502–505. doi: 10.2307/1128808

Fieder, N., Nickels, L., Krajenbrink, T., and Biedermann, B. (2018). Garlic and ginger are not like apples and oranges: Effects of mass/count information on the production of noun phrases in English. *Q. J. Exp. Psychol.* 71, 717–748. doi: 10.1080/17470218.2016.1276203

Foucart, A., Branigan, H. P., and Bard, E. G. (2010). Determiner selection in romance languages: evidence from French. *J. Exp. Psychol. Learning Mem. Cognit.* 36, 1414–1421. doi: 10.1037/a0020432

Fuhrmeister, P., and Bürki, A. (2022). Distributional properties of semantic interference in picture naming: Bayesian meta-analyses. *Psych. Bullet. Rev.* 29, 635–647. doi: 10.3758/s13423-021-02016-6

Geng, J., Kirchgessner, M., and Schnur, T. (2013). The mechanism underlying lexical selection: evidence from the picture-picture interference paradigm. *Q. J. Exp. Psychol.* 66, 261–276. doi: 10.1080/17470218.2012.705861

Geng, J., Schnur, T. T., and Janssen, N. (2014). Relative speed of processing affects interference in Stroop and picture–word interference paradigms: evidence from the distractor frequency effect. *Lang. Cognit. Neurosci.* 29, 1100–1114. doi: 10.1080/01690965.2013.846473

Giezen, M. R., and Emmorey, K. (2016). Language co-activation and lexical selection in bimodal bilinguals: evidence from picture-word interference. *Bilingualism* 19, 264–276. doi: 10.1017/S1366728915000097

Guerrero, G., and Calvillo, D. P. (2016). Animacy increases second target reporting in a rapid serial visual presentation task. *Psych. Bullet. Rev.* 23, 1832–1838. doi: 10.3758/s13423-016-1040-7

Guttentag, R. E., and Haith, M. M. (1979). A developmental study of automatic word processing in a picture classification task. *Child Dev.* 50, 894–896. doi: 10.2307/1128962

Hagen, T., and Laeng, B. (2016). The change detection advantage for animals: an effect of ancestral priorities or progeny of experimental design? *I-Perception* 7, 366. doi: 10.1177/2041669516651366

Hargreaves, I. S., Leonard, G. A., Pexman, P. M., Pittman, D. J., Siakaluk, P. D., Goodyear, B. G., et al. (2012a). The neural correlates of the body-object interaction effect in semantic processing. *Front. Hum. Neurosci.* 6, 22. doi: 10.3389/fnhum.2012.00022

Hargreaves, I. S., Pexman, P. M., Johnson, J. C., and Zdrazilova, L. (2012b). Richer concepts are better remembered: number of features effects in free recall. *Front. Hum. Neurosci.* 6, 73. doi: 10.3389/fnhum.2012.00073

Hashimoto, N., and Thompson, C. K. (2010). The use of the picture-word interference paradigm to examine naming abilities in aphasic individuals. *Aphasiology* 24, 580–611. doi: 10.1080/02687030902777567

He, C., and Cheung, O. S. (2019). Category selectivity for animals and man-made objects: Beyond low- and mid-level visual features. *J. Vision* 19, 22. doi: 10.1167/19.12.22

Heard, A., Madan, C. R., Protzner, A. B., and Pexman, P. M. (2019). Getting a grip on sensorimotor effects in lexical-semantic processing. *Behav. Res. Methods* 51, 1–13. doi: 10.3758/s13428-018-1072-1

Hocking, J., McMahon, K. L., and Zubicaray, G. I. (2010). Semantic interference in object naming: an fMRI study of the postcue naming paradigm. *NeuroImage* 50, 796–801. doi: 10.1016/j.neuroimage.2009.12.067

Hoffman, P., Lambon Ralph, M. A., and Rogers, T. T. (2013). Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words. *Behav. Res. Methods* 45, 718–730. doi: 10.3758/s13428-012-0278-x

Hofrichter, R., Mueller, M. E., and Rutherford, M. D. (2021). Children's perception of animacy: social attributions to moving figures. *Perception* 50, 387–398. doi: 10.1177/03010066211010142

Hwang, H., and Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *J. Exp. Psychol. Learning Memory Cognit.* 40, 1363–1376. doi: 10.1037/a0036797

Jackson, R. E., and Calvillo, D. P. (2013). Evolutionary relevance facilitates visual information processing. *Evol. Psychol.* 11, 1011–1026. doi: 10.1177/147470491301100506

Janssen, N., Schirm, W., Mahon, B. Z., and Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *J. Exp. Psychol. Learn. Memory Cognit.* 34, 249–256. doi: 10.1037/0278-7393.34.1.249

Jerger, S., Damian, M. F., Mills, C., Bartlett, J., Tye-Murray, N., Abdi, H., et al. (2013). Effect of perceptual load on semantic access by speech in children. *J. Speech Lang. Hearing Res.* 56, 388–403. doi: 10.1044/1092-4388(2012/11-0186)

Jescheniak, J. D., Wöhner, S., Bethcke, H. S., and Beaupain, M. C. (2020). Semantic interference in the picture-word interference task: is there a pre-lexical, conceptual contribution to the effect? *Psychonomic Bullet. Rev.* 27, 373–378. doi: 10.3758/s13423-019-01667-w

Krott, A., Medaglia, M. T., and Porcaro, C. (2019). Early and late effects of semantic distractors on electroencephalographic responses during overt picture naming. *Front. Psychol.* 10, 696. doi: 10.3389/fpsyg.2019.00696

Laws, K. R., Leeson, V. C., and Gale, T. M. (2002). The effect of 'masking' on picture naming. *J. Devoted Study Nerv. Syst. Behav.* 38, 137–147. doi: 10.1016/S0010-9452(08)70646-4

Laws, K. R., and Neve, C. (1999). A 'normal' category-specific advantage for naming living things. *Neuropsychologia* 37, 1263–1269. doi: 10.1016/S0028-3932(99)00018-4

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75. doi: 10.1017/S0140525X99001776

Lorenz, A., Regel, S., Zwitserlood, P., and Rahman, R. A. (2018). Age-related effects in compound production: Intact lexical representations but more effortful encoding. *Acta Psychol.* 191, 289–309. doi: 10.1016/j.actpsy.2018.09.001

Lorenz, A., and Zwitserlood, P. (2016). Semantically transparent and opaque compounds in German noun-phrase production: evidence for morphemes in speaking. *Front. Psychol.* 7, 1943. doi: 10.3389/fpsyg.2016.01943

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory Cognit.* 7, 485–495. doi: 10.3758/BF03198265

Lüttmann, H., Zwitserlood, P., and Bölte, J. (2011). Sharing morphemes without sharing meaning: production and comprehension of german verbs in the context of morphological relatives. *Can. J. Exp. Psychol.* 65, 173–91. doi: 10.1037/a0023794

MacLeod, C. M. (1991). Half a century of research on the stroop effect: an integrative review. *Psychol. Bullet.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163

Mädebach, A., Kieseler, M., and Jescheniak, J. D. (2018). Localizing semantic interference fromdistractor sounds in picture naming: a dual-task study. *Psych. Bullet. Rev.* 25, 1909–1916. doi: 10.3758/s13423-017-1386-5

Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., and Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron* 63, 397–405. doi: 10.1016/j.neuron.2009.07.012

Mahon, B. Z., and Caramazza, A. (2009). Why does lexical selection have to be so hard? Comment on Abdel Rahman and Melinger's swinging lexical network proposal. *Lang. Cognit. Proc.* 24, 735–748. doi: 10.1080/01690960802597276

Mailend, M., and Maas, E. (2013). Speech motor programming in apraxia of speech: Evidence from a delayed picture-word interference task. *Am. J. Speech Lang. Pathol.* 22, S380–S396. doi: 10.1044/1058-0360(2013/12-0101)

Melinger, A., and Abdel Rahman, R. (2004). Investigating the interplay between semantic and phonological distractor effects in picture naming. *Brain Lang.* 90, 213–220. doi: 10.1016/S0093-934X(03)00434-6

Meyer, A. S., and Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *J. Exp. Psychol. Learn. Memory Cognit.* 17, 1146–1160. doi: 10.1037/0278-7393.17.6.1146

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Group PRISMA preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6, e1000097. doi: 10.1371/journal.pmed.1000097

Muehlhaus, J., Heim, S., Sachs, O., Schneider, F., Habel, U., Sass, K., et al. (2013). Is the motor or the garage more important to the car? the difference between semantic associations in single word and sentence production. *J. Psycholing. Res.* 42, 37–49. doi: 10.1007/s10936-012-9209-3

Mulatti, C., Ceccherini, L., and Coltheart, M. (2015). What can we learn about visual attention to multiple words from the word-word interference task? *Mem. Cognit.* 43, 121–132. doi: 10.3758/s13421-014-0450-x

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017). Remembering the living: Episodic memory is tuned to animacy. *Curr. Directions Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., and Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory: the mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

New, J. J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Nat. Acad. Sci.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

New, J. J., Schultz, R. T., Wolf, J., Niehaus, J. L., Klin, A., German, T. C., et al. (2010). The scope of social attention deficits in autism: prioritized orienting to people and animals in static natural scenes. *Neuropsychologia* 48, 51–59. doi: 10.1016/j.neuropsychologia.2009.08.008

Oppermann, F., Jescheniak, J. D., and Schriefers, H. (2008). Conceptual coherence affects phonological activation of context objects during object naming. *J. Exp. Psychol. Learning Mem. Cognit.* 34, 587–601. doi: 10.1037/0278-7393.34.3.587

Perani, D., Cappa, S. F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., et al. (1995). Different neural systems for the recognition of animals and man-made tools. *Neuroreport* 6, 1637–1641. doi: 10.1097/00001756-199508000-00012

Pratt, J., Radulescu, P. V., Guo, R. M., and Abrams, R. A. (2010). It's alive! animate motion captures visual attention. *Psychol. Sci.* 21, 1724–1730. doi: 10.1177/0956797610387440

Proklova, D., Kaiser, D., and Peelen, M. V. (2016). Disentangling representations of object shape and object category in human visual cortex: the animate-inanimate distinction. *J. Cognit. Neurosci.* 28, 680–692. doi: 10.1162/jocn_a_00924

Qingfang, Z., and Yufang, Y. (2004). The time course of semantic, orthographic and phonological activation in Chinese word production. *Acta Psychol. Sinica* 36, 1–8.

Radanović, J., Westbury, C., and Milin, P. (2016). Quantifying semantic animacy: how much are words alive? *Appl. Psycholing.* 37, 1477–1499. doi: 10.1017/S0142716416000096

Rakison, D. H., and Poulin-Dubois, D. (2001). Developmental origin of the animate–inanimate distinction. *Psychol. Bullet.* 127, 209–228. doi: 10.1037/0033-2909.127.2.209

Ro, T., Friggel, A., and Lavie, N. (2007). Attentional biases for faces and body parts. *Visual Cognit.* 15, 322–348. doi: 10.1080/13506280600590434

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: retrieval of verbs. *Cognition* 47, 59–87. doi: 10.1016/0010-0277(93)90062-Z

Roelofs, A., and Verhoef, K. (2006). Modeling the control of phonological encoding in bilingual speakers. *Bilingualism* 9, 167–176. doi: 10.1017/S1366728906002513

Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Dev.* 48, 643–647. doi: 10.2307/1128667

Rostad, K., Yott, J., and Poulin-Dubois, D. (2012). Development of categorization in infancy: advancing forward to the animate/inanimate level. *Infant Behav. Dev.* 35, 584–595. doi: 10.1016/j.infbeh.2012.05.005

Sailor, K., Brooks, P. J., Bruening, P. R., Seiger-Gardner, L., and Guterman, M. (2009). Exploring the time course of semantic interference and associative priming in the picture–word interference task. *Q. J. Exp. Psychol.* 62, 789–801. doi: 10.1080/17470210802254383

Sá-Leite, A. R., Fraga, I., and Comesaña, M. (2019). Grammatical gender processing in bilinguals: an analytic review. *Psych. Bullet. Rev.* 26, 1148–1173. doi: 10.3758/s13423-019-01596-8

Sá-Leite, A. R., Haro, J., Comesaña, M., and Fraga, I. (2021). Of beavers and tables: the role of animacy in the processing of grammatical gender within a picture-word interference task. *Front. Psychol.* 12, 661175. doi: 10.3389/fpsyg.2021.661175

Sá-Leite, A. R., Luna, K., Fraga, I., and Comesaña, M. (2020). The gender congruency effect across languages in bilinguals: a meta-analysis. *Psych. Bullet. Rev.* 27, 677–693. doi: 10.3758/s13423-019-01702-w

Sá-Leite, A. R., Luna, K., Tomaz, Â., Fraga, I., and Comesaña, M. (2022). The mechanisms underlying grammatical gender selection in language production: a meta-analysis of the gender congruency effect. *Cognition* 224, 105060. doi: 10.1016/j.cognition.2022.105060

Schiller, N. O., and Caramazza, A. (2003). Grammatical feature selection in noun phrase production: evidence from German and Dutch. *J. Memory Lang.* 48, 169–194. doi: 10.1016/S0749-596X(02)00508-9

Schnur, T. T., Costa, A., and Caramazza, A. (2006). Planning at the phonological level during sentence production. *J. Psycholing. Res.* 35, 189–213. doi: 10.1007/s10936-005-9011-6

Shao, Z., Roelofs, A., Martin, R. C., and Meyer, A. S. (2015). Selective inhibition and naming performance in semantic blocking, picture-word interference, and color–word Stroop tasks. *J. Exp. Psychol.* 41, 1806–1820. doi: 10.1037/a0039363

Shin, S. (2016). *Word Frequency Effects on Lexical Selection: Evidence From a Picture-Word Interference (PWI) Task (Ph.D.). Available from Psychology Database. (1846543022).* Available online at: https://www.proquest.com/dissertations-theses/word-frequency-effects-on-lexical-selection/docview/1846543022/se-2?accountid=14542 (accessed May, 2022).

Starreveld, P. A., and La Heij, W. (2017). Picture-word interference is a stroop effect: a theoretical analysis and new empirical findings. *Psychon. Bullet. Rev.* 24, 721–733. doi: 10.3758/s13423-016-1167-6

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *J. Mem. Lang.* 51, 247–250. doi: 10.1016/j.jml.2004.03.002

Toma, R. J., and Tsao, Y. C. (1985). Interference effects in the picture-word stroop task. *Percep. Motor Skills* 61, 223–228. doi: 10.2466/pms.1985.61.1.223

Underwood, G. (1976). Semantic interference from unattended printed words. *Br. J. Psychol.* 67, 327–338. doi: 10.1111/j.2044-8295.1976.tb01519.x

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Blunt, J. R. (2013). Adaptive memory: animacy processing produces mnemonic advantages. *Exp. Psychol.* 60, 172–178. doi: 10.1027/1618-3169/a000186

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Cogdill, M. (2014). Adaptive memory: animacy effects persist in paired-associate learning. *Memory* 23, 657–663. doi: 10.1080/09658211.2014.916304

Wilshire, C., and Singh, S., and Tattersall, C. (2016). Serial order in word form retrieval: New insights from the auditory picture–word interference task. *Psych. Bullet. Rev.* 23, 299–305. doi: 10.3758/s13423-015-0882-8

Xiao, X., Dong, Q., Chen, C., and Xue, G. (2016). Neural pattern similarity underlies the mnemonic advantages for living words. *Cortex* 79, 99–111. doi: 10.1016/j.cortex.2016.03.016

Yang, J., Wang, A., Yan, M., Zhu, Z., Chen, C., Wang, Y., et al. (2012). Distinct processing for pictures of animals and objects: evidence from eye movements. *Emotion* 12, 540–551. doi: 10.1037/a0026848

Yu, L., and Shu, H. (2003). Phonological processing mechanism in Chinese speech production. *Psychol. Sci.* 26, 818–822.

Zannino, G. D., Perri, R., Pasqualetti, P., Caltagirone, C., and Carlesimo, G. A. (2006). Analysis of the semantic representations of living and nonliving concepts: a normative study. *Cognit. Neuropsychol.* 23, 515–540. doi: 10.1080/02643290542000067

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 10 May 2023
DOI 10.3389/fpsyg.2023.1141540

EDITED BY
Josefa N. S. Pandeirada,
University of Aveiro, Portugal

REVIEWED BY
Janell Rebecca Blunt,
Purdue University, United States
Stephanie Kazanas,
Tennessee Technological University,
United States

*CORRESPONDENCE
Aurélia Bugaïska
✉ aurelia.bugaiska@u-bourgogne.fr

SPECIALTY SECTION
This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

RECEIVED 10 January 2023
ACCEPTED 31 March 2023
PUBLISHED 10 May 2023

CITATION
Bugaiska A, Bonin P and Witt A (2023) Do young children, like young adults, remember animates better than inanimates?
*Front. Psychol.* 14:1141540.
doi: 10.3389/fpsyg.2023.1141540

# Do young children, like young adults, remember animates better than inanimates?

Aurélia Bugaiska*, Patrick Bonin and Arnaud Witt

LEAD-CNRS UMR 5022, University of Burgundy, Dijon, France

It has repeatedly been shown in adults that animates are remembered better than inanimates. According to the adaptive view of human memory this is due to the fact that animates are generally more important for survival than inanimates. Animacy enhances not only the quantity but also the quality of remembering. The effect is primarily driven by recollection. Virtually all studies have been conducted in adults, and we believe that the investigation of animacy effects in children is also highly relevant. The present study therefore tested the animacy effect on recollection in young (6–7 years, $M = 6.6$ years) and older children (10–12 years, $M = 10.83$ years) using the Remember/Know paradigm. As found in adults, an animacy effect on memory was found, but only in older children, and specifically in the "remember" responses, suggesting, once again, its episodic nature.

KEYWORDS

**animacy effect, children, recollection, episodic memory, Remember/Know**

## Introduction

Adaptive memory was first described by Nairne et al. (2008), Nairne (2010, 2015), Nairne and Pandeirada (2016). The authors postulated that human memory has evolved as a result of pressures faced by our ancestors in the distant past. According to this theory, memory is enhanced when the information is relevant to fitness and survival. A number of studies have provided evidence supporting this view (see Nairne et al., 2017a for a comprehensive review), including animacy effects in memory (e.g., Nairne et al., 2013; Bonin et al., 2014; for a review: Nairne et al., 2017b).

The animacy effect concerns the observation that animate entities (e.g., *snake*, *cow*, and *woman*) are remembered better than inanimate entities (e.g., *mountain*, *bottle*, and *car*). Importantly, it was by adopting an evolutionary lens to the study of (episodic) memory that Nairne et al. (2013) first demonstrated the importance of the mnemonic dimension of animacy. Because animates have a stronger fitness value than inanimates (i.e., they can be predators, prey or potential sexual partners), they predicted and then empirically demonstrated that animates have a memory advantage over inanimates. Animacy effects in memory have been found by different research teams world-wide, first in the United States (e.g., VanArsdall et al., 2013), followed by researchers in Europe [e.g., France (Bonin et al., 2014), Germany (Meinhardt et al., 2018)], and also in China (e.g., Li et al., 2016). The memory benefit of animacy has been found with different types of stimuli: words (Nairne et al., 2013; Bonin et al., 2014), non-words linked to animate vs. inanimate properties (VanArsdall et al., 2013), and pictures (Bonin et al., 2014). Importantly, these effects have been observed in both recall rates and in recognition accuracy. Of particular interest here is that animacy effects have been found in studies using the Remember/Know paradigm (Gardiner, 1988), in which participants indicate whether they specifically remember (R) contextual details of items they recognize (e.g., a feeling, a location), or whether they just know (K) that they have seen the items. Regarding animacy, it has been found that participants give more "R" responses to animate than inanimate items, whereas "K" responses do not differ reliably between the two types of item (Bonin et al., 2014; Bugaiska et al., 2016). This pattern of findings strongly suggests that animacy effects in memory are episodic in nature, in that episodic memory is characterized by the remembering of contextual information in (young) adults (e.g., Rawlinson and Kelley, 2021; Komar et al., 2022).

To the best of our knowledge, only one study has examined whether young children, like adults, show enhanced retention of animacy-related information (Aslan and John, 2016). In Aslan and John's (2016) study, kindergarten and elementary school children (4–11 years) were tested. As in Nairne et al.'s (2013) study with adults, the children were presented with non-words paired either with properties characteristic of humans (e.g., speaks French) or animals (e.g., has fur), or with properties characteristic of inanimate things (e.g., has a lid). For each non-word (e.g., BULA, LAFE), children were asked to give a quick "living" or "non-living" response (forced choice), and after a retention interval of 3 min, they had to recognize the non-words. Non-words paired with human or animal characteristics were recognized better than those paired with inanimate properties; in other words, an "animacy effect" in memory was found in children. The advantage of animate over inanimate non-words was identical across age groups, suggesting developmental invariance of the benefit over the tested age range. The authors concluded that young children's memory is tuned to process and retain information related to animacy. Hence, Aslan and John's (2016) findings provide further support for the evolutionary view of memory put forward by Nairne et al. (2008). However, their study did not distinguish between recollection and familiarity. As proposed in the literature, retrieval using a recognition task relies on two distinct processes: recollection and familiarity. As stated above, adults have been shown to recognize animates better than inanimates (Bonin et al., 2014; Bugaiska et al., 2016; Rawlinson and Kelley, 2021; Komar et al., 2022), and importantly, the animacy effect has been observed on "Remember" but not on "Know" responses (Bonin et al., 2014). This pattern of findings supports the hypothesis that the animacy effect in memory is episodic in nature. Animacy enhances not only the quantity but also the quality of remembering; in other words, the effect is primarily driven by recollection. We believe that identifying the nature of animacy effects in young children is an important issue, which was not addressed by Aslan and John (2016). Therefore, the aims of the present study were to establish whether animacy effects are replicable in children, and more importantly, to determine whether these effects are episodic in nature, as found in adults (e.g., Bonin et al., 2014). The Remember/Know paradigm has rarely been used in studies with young children, but the available evidence suggests that the proportion of "Remember" responses made by young children (8–10 years) is smaller than that made by older children (11–13 and 14–16 years) and young adults (17–19 years); by contrast, there is no age-related difference in the proportion of "Know" responses (Billingsley et al., 2002). A recent study by Canada et al. (2022) found that other aspects of children's cognitive development might enhance episodic memory performance, especially during middle childhood (e.g., 6–8 years; for a review see Schneider and Ornstein, 2019). This supports the view that middle childhood is a transitional period for the development of episodic memory and attention (Diaz et al., 2018).

To recap, the aim of this research was to study animacy effects in memory, and more specifically in recollection, in young children. Unlike Aslan and John's (2016) study, which used non-words linked to animate and inanimate properties, we investigated the quality of retrieval of animate and inanimate words in order to investigate whether animacy effects emerge at a relatively young age, as found by Aslan and John (2016), and more importantly, whether these effects in children are episodic in nature.

# Materials and methods

## Participants

A total of 42 children from two age groups (younger and older elementary school children) took part in this study. One of the children in the younger group was excluded because her/his false alarm rates exceeded hit rates in all animate and inanimate conditions. The final sample thus comprised 20 younger elementary school children (6–7 years, $M = 6.6$ years), and 21 older elementary school children (10–12 years, $M = 10.83$ years). Prior to the study, we conducted a power analysis using G*Power (Faul et al., 2007) for sample size estimation based on the data of Aslan and John (2016). Their sample ($N = 90$) was divided into three age groups, but the authors did not provide data analyses for each age group because the animacy condition factor did not interact with age, $F(4, 174) < 1$. Our estimation is therefore based on the main effect of the animacy condition as a within-subject factor (3: human, animal and inanimate). The authors reported a main effect of this condition, $F(2, 174) = 12.9$, $MSE = 0.03$, $p < 0.001$, $\eta = 0.129$, with higher recognition of items related to humans (56.7%) and animals (57.4%) than inanimates (45.7%, $ps < 0.001$), while there was no difference between the two animate conditions (human vs. animal, ($p = 0.769$). The partial eta-squared effect size was $\eta = 0.129$. For a group assessed across three observations, with effect size specification as in GPower 3.0, an alpha of.05 and a power of.80, the minimum sample size needed with this effect size is $N = 13$. In the present experiment, we compared two animacy conditions (animate vs. inanimate) rather than three (human, animal and inanimate), as no difference was observed between the two animate categories in Aslan and John's (2016) study. With this effect size, the minimum sample size for two repeated measures is $N = 16$. We rounded this figure up to $N = 20$ per age group, which is adequate to test the study hypothesis.

## Material and design

The participants performed a recognition memory test using the Remember/Know/Guess method. This study was carried out in the context of a research agreement (agreement no. 0482-2021) between the laboratory, the university, the French national center for scientific research (CNRS) and the academic inspectorate ("Inspection Académique de Côte d'Or"). We conducted this study in accordance with the 1964 Declaration of Helsinki and we obtained written parental consent for each child. All participants were tested individually.

### Stimuli for encoding

For the R/K/G paradigm, the material consisted of 24 nouns selected from the databases of Snodgrass and Vanderwart (1980) and Bonin et al. (2003). Each word referred to either an animate or an inanimate object.[1] The words included 12 animate and 12 inanimate items, matched for surface variables (number of letters and bigram frequency), lexical variables (book frequency, subtitle frequency, age-of-acquisition, number of orthographic neighbors, and orthographic uniqueness point), and semantic variables (conceptual familiarity, imageability, image variability, concreteness, and emotional valence).[2] Regarding age of acquisition, we selected words expected to be acquired by the children in our sample. The statistical characteristics of the controlled variables can be found in the **Supplementary material**. For the recognition task, we included twelve additional ("new") words (6 animate and 6 inanimate), which matched the objective word frequency of the initial experimental words ("old").

## Procedure

The children were tested individually, seated comfortably in a quiet room.

### Encoding task

They were fitted with headphones so that they could hear the words perfectly. A word was presented every three seconds, and the children were asked to repeat each one out loud to ensure that they had heard it correctly. They were not instructed to learn the words, so encoding was incidental. Two lists were created, each with the same 24 words (12 animate and 12 animate) in a different order, so that half the children were presented with the words in one order and half with the words in the other order.

### Distractor task

After the encoding task, the participants were given 2 min to perform the Cancelation subtest of the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) (Wechsler, 2014). This test was used as an interference task.

### Recognition task

They then performed the recognition task, in which all 36 words (24 targets and 6 animate and 6 inanimate fillers) were presented orally *via* headphones. The children had as much time as they wanted to respond. For each word, they were asked if they recognized it from the previously presented list. If so, they were instructed to give a remember (R), know (K), or guess (G) response: an R-response if they had a specific recollection of the learning sequence (e.g., it brought to mind a particular association, image, or some other personal experience, or because they recalled something about its appearance or position); a K-response if they were sure they recognized the word but had no conscious recollection of learning it); a G-response if they were not sure whether they had already seen the word or not. To ensure that the instructions were understood, they were asked the following question: *"Do you remember hearing this word before?"* If they answered yes, they were asked: *"Did you think of anything in particular when you heard this word, or did you think of nothing? For example, if you had heard the word bike, you might have thought of your bike, or of a cartoon with a bike in it, or a family bike ride."* To ensure that the children had followed the instructions correctly, they were asked to explain two of their Remember and two of their Know judgments after the recognition phase.

### Control task

Finally, the participants were given a naming test, in which they had to name pictures corresponding to the words they had heard during the encoding phase. This task was included in addition to the age-of-acquisition control variable to ensure that all the children knew the words shown at encoding. All the children correctly named the pictures.

# Results

## Analysis of Remember/Know paradigm

The hits minus false alarms and standard errors for overall recognition, Remember and Know responses are presented in **Figure 1**.[3]

To test the effect of animacy and age on overall recognition, Remember responses and Know responses, we conducted a 2 × 2 ANOVA with Animacy as a within-subject factor and Age as a between-subjects factor on these measures.

### Overall recognition

The younger children recognized fewer words from the previously presented list than the older children, $F(1,39) = 25.69$, $p < 0.001$, $\eta^2_p = 0.40$. The main effect of Type of words was not significant, $F(1,39) < 0.001$, $p = 0.99$. Finally, as shown in **Figure 1**, the interaction between Age and Type of words was not significant, $F(1,39) = 0.33$, $p = 0.57$.

### Remember responses

The ANOVA of R-responses revealed a reliable main effect of Age, $F(1,39) = 6.10$, $p = 0.02$, $\eta^2_p = 0.14$, the younger children recollecting fewer words than the older children. A main effect of Type of words emerged, $F(1,39) = 5.72$, $p = 0.02$, $\eta^2_p = 0.13$, showing that animate words gave rise to more recollective experience than inanimate words. Finally, the interaction between

---

[1] We used a broad definition of animacy as in previous literature on animacy effects in memory (e.g., Gelin et al., 2019). Thus, we considered animate words to refer to living things that can move on their own (e.g., a cow) and inanimate words to refer to non-living things that cannot move on their own (e.g., a car). The authors used this definition to code the words as animates versus inanimates; full agreement for all the words was reached. Lists of words can be found in the **Supplementary material**.

[2] We are aware that the norms listed here are based on adults. Unfortunately, norms on these dimensions are not available for children in French.

[3] A table summarizing the proportions of correct and false alarm responses for general recognition, R-responses and K-responses as a function of age and word type can be found in Table 2 in the **Supplementary material**.

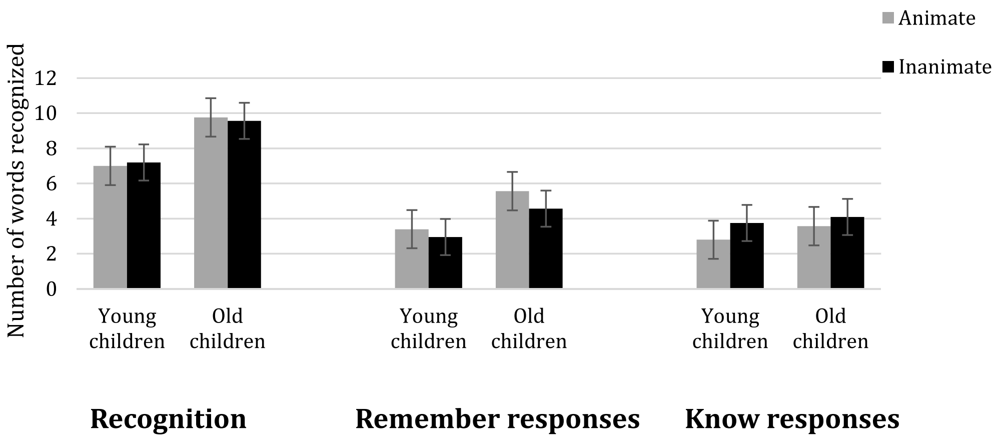


FIGURE 1
Mean number of correctly recognized words (Hits-FA) as a function of age (young vs. older children) and type of words (animate vs. inanimate) for recognition, remember responses and know responses. Error bars represent standard errors of the mean.

Age and Type of words was not significant, $F(1,39) = 0.82$, $p = 0.37$. Nevertheless, as shown in **Figure 1**, the animacy effect was greater in older than in younger children. A Paired sample *t*-test were conducted to determine the effect of animacy on the remember responses in the older and younger children. Results showed that older children recollected animate words better than inanimate words, $t(20) = 2.14$, $p = 0.0451$ ($M = 5.57$ and $M = 4.57$). However, the difference between animate and inanimate words for younger children was not significant, $t(19) = 1.18$, $p = 0.25$ ($M = 3.4$ and $M = 2.95$).

### Know responses

The analysis of K-responses (**Figure 1**) revealed that there was no reliable effect of age, $F(1,39) = 0.72$, $p = 0.40$. The main effect of Type of words was significant, $F(1,39) = 5.7$, $p = 0.02$, $\eta^2_p = 0.13$, with more Know responses for inanimate than animate words. Finally, there was no reliable interaction between Age and Type of words, $F(1,39) = 0.48$, $p = 0.49$.

## Discussion

Previous studies have established that animacy effects in memory are found on R-responses (an index of recollection) but not on K-responses (Bonin et al., 2014; Bugaiska et al., 2016; Rawlinson and Kelley, 2021; Komar et al., 2022), suggesting that the animacy effect is episodic in nature. However, these studies only involved young adults. We believe that it is worth investigating whether this is also the case for children. To the best of our knowledge, this is the first study to investigate the role of animacy in children's recollection, and to this end we asked a simple question: Do younger and older children remember animates better than inanimates, in the same way as young adults? Using a remember-know procedure with a sample of children aged 6–12 years, the present findings do not provide a clear-cut answer to this question, some conclusions varying with age group. First of all, there was no reliable animacy effect on overall recognition for either age group. However, when the recognition performance of older children was divided between "recollection" and "familiarity," we found an animacy effect on recollection but not on know responses. This suggests that the animacy effect in older children is due to an increase in recollection. Importantly, our findings are in line with those of Aslan and John (2016), but extend them by suggesting that the animacy effect is underpinned by episodic memory processes from the age of 10–11 years. It is particularly noteworthy that recollection processes are involved for animate but not inanimate words from an early age. It seems that inanimate words are not encoded with contextual details, and therefore this type of information is not helpful when they have to be remembered. Finally, our findings in older children are in line with the literature on young adults (Bonin et al., 2014; Bugaiska et al., 2016; Rawlinson and Kelley, 2021; Komar et al., 2022), showing that the animacy effect in memory is related to the quality of remembering, but has no effect on knowing. Overall, the current findings are consistent with a functional-evolutionary view of human memory, which posits that our memory systems have been tuned by natural selection due to pressures faced by our hunter-gatherer ancestors in the distant past. In particular, it has been suggested that recurrent interactions with animates exerted strong evolutionary pressure on humans, leading to the development of memory systems that prioritize the processing and remembering of animates.

Regarding the younger children, there was no reliable animacy effect for either overall recognition or recollection. This is at odds with the results of Aslan and John (2016), who found that the benefit of animate non-words was identical across age groups, suggesting developmental invariance of the benefit over the age range tested (i.e., 4–11 years). They concluded that young children's memory is "tuned" to process and retain animacy from a very early age (4 years). While we cannot provide a satisfactory explanation for this discrepancy between our findings and theirs, we suggest that it could be linked to the way the stimuli were presented. In Aslan and John's (2016) study, the children were asked to respond rapidly about the animacy status of non-words (e.g., BULA, LAFE) based on properties that referred to humans (e.g., speaks French) or animals (e.g., has fur) or to inanimate entities (e.g., has a lid). In our study, we did not ask the children to pay attention at encoding to the status (animate versus non-animate) or to certain semantic characteristics of the words, but simply to read them aloud. It is possible that this difference in the protocol was significant, because information about animacy had to be inferred from reading the words. As previously suggested, imagery skills could explain the animacy effect on memory performance (Blunt and VanArsdall, 2021). Immature imagery skills could therefore explain the lack of animacy effect in young children. The ability of children to form internal representations including movements is indeed still a matter of debate. From a Piagetian perspective, mental representations develop with age and are constrained by the characteristics of the stages of cognitive development. According to Piaget and Inhelder (1966), children under the age of 7 to 8 years are not able to represent movements, limiting mental representations to static states. The concrete operational stage would provide the framework within which transformations or movements can be represented. However, conflicting results suggest that 4- to 5-year-old children use kinetic imagery to solve mental rotation tasks (e.g., Marmor, 1972). A major difference between Marmor's study and those of Piaget and Inhelder is that the children were instructed to use kinetic imagery to solve the rotation task in the former but not in the latter. If young children do not perceive the relevance of using kinetic imagery during a mental rotation task, it is very likely that they will not spontaneously represent animates in motion when nothing invites them to do so, as in the present study, while they are able to form kinetic imagery of animates when the procedure draws attention to their "animated" characteristics, as in Aslan and John's (2016) study. Further studies should address this issue by contrasting the effect of these two procedures on the emergence of the animacy effect in young children.

Turning to recollection, the difference in the effect of animacy between young and older children is not totally unexpected. As mentioned in the Introduction, it has been suggested that other aspects of children's cognitive development may enhance episodic memory performance, notably during middle childhood (e.g., 7 years; for review see Schneider and Ornstein, 2019). A study conducted with 6- to 8-year-old children suggested that middle childhood is a transitional period for the development of episodic memory and attention (Diaz et al., 2018). Similarly, it is likely that environmental changes interact with the development of multiple cognitive processes and contribute to improvements during childhood.

To examine further the animacy effect in episodic memory in very young children, it would be interesting to repeat this experiment, but asking the children to read the words and say whether they refer to animate or inanimate entities. It would also be interesting to replicate Aslan and John's (2016) study with non-words by adding a recognition task using the Remember/Know paradigm to examine the extent to which animacy effects in young children rely on recollection. It is possible that the animacy effect in episodic memory does not emerge till a later age and that it is related to the development of episodic memory. In that case, animacy effects could be used as an index of episodic memory functioning/maturation in children.

To conclude, do children, like young adults, remember animates better than inanimates? The answer is "yes" for older children, for whom the animacy effect relies on recollection. As found with young adults, the animacy effect in memory in older children (10–12 years, $M = 10.83$ years) is episodic in nature. For younger children, the tentative answer is "no," but further studies are clearly needed to gain a better understanding of when (and how) this memory effect emerges in young children.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by the research agreement (agreement n°: 0482- 2021) between the laboratory, the university, the French National Center for Scientific Research (CNRS) and the academic inspectorate ("Inspection Académique de Côte-d'Or"). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## Author contributions

AB, AW, and PB contributed to the conception and design of the study and performed the statistical analysis, and contributed to the manuscript writing. AB and AW performed the investigation. AB wrote the first draft of the manuscript. All authors approved the submitted version.

## Funding

This research was supported by a BQR (Bonus Qualité Recherche) grant from the Conseil Régional de Bourgogne Franche-Comté.

## 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.

## Publisher's note

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1141540/full#supplementary-material

# References

Aslan, A., and John, T. (2016). The development of adaptive memory: Young children show enhanced retention of animacy-related information. *J. Exp. Child Psychol.* 152, 343–350. doi: 10.1016/j.jecp.2016.07.007

Billingsley, R. L., Smith, M. L., and McAndrews, M. P. (2002). Developmental patterns in priming and familiarity in explicit recollection. *J. Exp. Child Psychol.* 82, 251–277. doi: 10.1016/S0022-0965(02)00007-3

Blunt, J. R., and VanArsdall, J. E. (2021). Animacy and animate imagery improve retention in the method of loci among novice users. *Mem. Cogn.* 49, 1360–1369. doi: 10.3758/s13421-021-01175-0

Bonin, P., Gelin, M., and Bugaiska, A. (2014). Animates are better remembered than inanimates: Further evidence from word and picture stimuli. *Mem. Cogn.* 42, 370–382. doi: 10.3758/s13421-013-0368-8

Bonin, P., Peereman, R., Malardier, N., Méot, A., and Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behav. Res. Methods Instrum. Comput.* 35, 158–167. doi: 10.3758/bf03195507

Bugaiska, A., Méot, A., and Bonin, P. (2016). Do healthy elders, like young adults, remember animates better than inanimates? An adaptive view. *Exp. Aging Res.* 42, 447–459. doi: 10.1080/0361073X.2016.1224631

Canada, K. L., Hancock, G. R., and Riggins, T. (2022). Developmental changes in episodic memory across early- to mid-childhood: Insights from a latent longitudinal approach. *Memory* 30, 248–261. doi: 10.1080/09658211.2021.2006233

Diaz, A., Blankenship, T. L., and Bell, M. A. (2018). Episodic memory in middle childhood: Age, brain electrical activity, and self-reported attention. *Cogn. Dev.* 47, 63–70. doi: 10.1016/j.cogdev.2018.03.003

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Mem. Cogn.* 16, 309–313. doi: 10.3758/BF03197041

Gelin, M., Bugaiska, A., Méot, A., Vinter, A., and Bonin, P. (2019). Animacy effects in episodic memory: Do imagery processes really play a role? *Memory* 27, 209–223. doi: 10.1080/09658211.2018.1498108

Komar, G. F., Mieth, L., Buchner, A., and Bell, R. (2022). Animacy enhances recollection but not familiarity: Convergent evidence from the remember-know-guess paradigm and the process-dissociation procedure. *Mem. Cogn.* 51, 143–159. doi: 10.3758/s13421-022-01339-6

Li, P., Jia, X., Li, X., and Li, W. (2016). The effect of animacy on metamemory. *Mem. Cogn.* 44, 696–705. doi: 10.3758/s13421-016-0598-7

Marmor, J. (1972). Holistic conception, and points of mild issue. *J. Individ. Psychol.* 28, 153–154.

Meinhardt, M. J., Bell, R., Buchner, A., and Röer, J. P. (2018). Adaptive memory: Is the animacy effect on memory due to emotional arousal? *Psychon. Bull. Rev.* 25, 1399–1404. doi: 10.3758/s13423-018-1485-y

Nairne, J. S. (2010). "Adaptive Memory: Evolutionary constraints on remembering," in *The psychology of learning and motivation*, Vol. 53, ed. B. H. Ross (Burlington, NJ: Academic Press), 1–32.

Nairne, J. S. (2015). "Adaptive memory: Novel findings acquired through forward engineering," in *Remembering: Attributions, processes, and control in human memory*, eds D. S. Lindsay, C. M. Kelley, A. P. Yonelinas, and H. L. Roediger (New York, NY: Psychology Press). doi: 10.1118/1.3661998

Nairne, J. S., and Pandeirada, J. N. S. (2016). Adaptive memory: The evolutionary significance of survival processing. *Perspect. Psychol. Sci.* 11, 496–511. doi: 10.1177/1745691616635613

Nairne, J. S., Pandeirada, J. N. S., and Fernandes, N. L. (2017a). "Adaptive memory," in *Learning and memory: A comprehensive reference*, 2nd Edn, Vol. 2, ed. H. B. John (Oxford: Elsevier), 279–293. doi: 10.1016/B978-0-12-809324-5.21060-2

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017b). Remembering the living: Episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Nairne, J. S., Pandeirada, J. N., and Thompson, S. R. (2008). Adaptive memory: The comparative value of survival processing. *Psychol. Sci.* 19, 176–180. doi: 10.1111/j.1467-9280.2008.02064.x

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory: The mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

Piaget, J., and Inhelder, B. (1966). *L'image mentale chez l'enfant. [The mental image in the child]*. Paris: Presses Universitaires de France.

Rawlinson, H. C., and Kelley, C. M. (2021). In search of the proximal cause of the animacy effect on memory: Attentional resource allocation and semantic representations. *Mem. Cogn.* 49, 1137–1152. doi: 10.3758/s13421-021-01154-5

Schneider, W., and Ornstein, P. A. (2019). Determinants of memory development in childhood and adolescence. *Int. J. Psychol.* 54, 307–315. doi: 10.1002/ijop.12503

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol.* 6, 174–215. doi: 10.1037//0278-7393.6.2.174

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Blunt, J. R. (2013). Adaptive memory: Animacy processing produces mnemonic advantages. *Exp. Psychol.* 60, 172–178. doi: 10.1027/1618-3169/a000186

Wechsler, D. (2014). *Wechsler intelligence scale for children (WISC-V)*, 5th Edn. Bloomington, MN: Pearson.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 11 May 2023
DOI 10.3389/fpsyg.2023.1164038

OPEN ACCESS

EDITED BY
Niharika Singh,
University of Allahabad, India

REVIEWED BY
Patrick Bonin,
Université de Bourgogne, France
Raoul Bell,
Heinrich Heine University of Düsseldorf,
Germany

*CORRESPONDENCE
Michael J. Serra
✉ michael.serra@ttu.edu

RECEIVED 11 February 2023
ACCEPTED 20 April 2023
PUBLISHED 11 May 2023

CITATION
Serra MJ and DeYoung CM (2023) The animacy advantage in memory occurs under self-paced study conditions, but participants' metacognitive beliefs can deter it.
*Front. Psychol.* 14:1164038.
doi: 10.3389/fpsyg.2023.1164038

# The animacy advantage in memory occurs under self-paced study conditions, but participants' metacognitive beliefs can deter it

Michael J. Serra* and Carlee M. DeYoung

Department of Psychological Sciences, Texas Tech University, Lubbock, TX, United States

**Introduction:** Animacy distinguishes living (animate) things from non-living (inanimate) things. People tend to devote attention and processing to living over nonliving things, resulting in a privileged status for animate concepts in human cognition. For example, people tend to remember more animate than inanimate items, a phenomenon known as the "animacy effect" or "animacy advantage." To date, however, the exact cause(s) of this effect is unknown.

**Methods:** We examined the animacy advantage in free-recall performance under computer-paced versus self-paced study conditions and using three different sets of animate and inanimate stimuli (Experiments 1 and 2). We also measured participants' metacognitive beliefs (expectations) about the task before it began (Experiment 2).

**Results:** We consistently obtained an animacy advantage in free-recall, regardless of whether participants studied the materials under computer-paced or self-paced conditions. Those in self-paced conditions spent less time studying items than did those in computer-paced conditions, but overall levels of recall and the occurrence of the animacy advantage were equivalent by study method. Importantly, participants devoted equivalent study time to animate and inanimate items in self-paced conditions, so the animacy advantage in those conditions cannot be attributed to study time differences. In Experiment 2, participants who believed that inanimate items were more memorable instead showed equivalent recall and study time for animate and inanimate items, suggesting that they engaged in equivalent processing of animate and inanimate items. All three sets of materials reliably produced an animacy advantage, but the effect was consistently larger for one set than the other two, indicating some contribution of item-level properties to the effect.

**Discussion:** Overall, the results suggest that participants do not purposely allocate greater processing to animate over inanimate items, even when study is self-paced. Rather, animate items seem to naturally trigger greater richness of encoding than do inanimate items and are then better remembered, although under some conditions participants might engage in deeper processing of inanimate items which can reduce or eliminate the animacy advantage. We suggest that researchers might conceptualize mechanisms for the effect as either centering on intrinsic, item-level properties of the items or centering on extrinsic, processing-based differences between animate and inanimate items.

KEYWORDS

**animacy advantage, animacy effect, adaptive memory, free recall (memory), self-paced study, metacognitive beliefs**

## 1. Introduction

Animacy refers to the qualities that distinguish animate (living) things from inanimate (non-living) things, or the qualities that make something seem alive (VanArsdall and Blunt, 2022). Examples of living things are animals (hedgehog) and humans (dancer), while nonliving things include natural objects (rock) and man-made objects (plate). People tend to preferentially devote attention and processing to living over nonliving things in their environment and in their thoughts, and this tendency in turn affects many other aspects of human cognition, including attention (New et al., 2007; Altman et al., 2016; Bugaiska et al., 2019), perception (Scholl and Tremoulet, 2000), language (Vihman and Nelson, 2019), numerical processing (Zanini et al., 2020), metacognitive monitoring (Li et al., 2016; DeYoung and Serra, 2021), and memory (Popp and Serra, 2016; Nairne et al., 2017). Most relevant, in many memory tasks, people tend to remember more animate than inanimate items, a phenomenon known as *the animacy effect* or *the animacy advantage*. In the present experiments, we compared the occurrence of the animacy advantage for free-recall performance under computer-paced versus self-paced study conditions and using three different sets of stimuli.

The animacy advantage in memory can occur in recognition tasks (e.g., VanArsdall et al., 2013; Bonin et al., 2014; Leding, 2020) and cued-recall tasks (e.g., VanArsdall et al., 2015; Popp and Serra, 2016; DeYoung and Serra, 2021; but see Popp and Serra, 2016; Kazanas et al., 2020; Serra and DeYoung, 2023), but to date researchers have most often examined the effect in the context of free-recall (*cf.* Nairne et al., 2013, 2017; Bonin et al., 2015; Li et al., 2016; Popp and Serra, 2016, 2018; Gelin et al., 2017, 2019; VanArsdall et al., 2017; Leding, 2018, 2019; Meinhardt et al., 2018, 2020; Félix et al., 2019; Serra, 2021). In free-recall tasks, participants usually study a list of words, one at a time, and then try to recall the words from memory without hints or assistance. Different researchers have found the effect with different sets of words, different numbers of words per list, and different numbers of study trials. It occurs with both pure lists (only animate words versus only inanimate words) and mixed lists (both animate and inanimate words) of to-be-remembered words (Popp and Serra, 2016), with and without the inclusion of buffer words or a distractor task (e.g., Nairne et al., 2013; Popp and Serra, 2016), and across the serial-positions of the list (Serra, 2021).

The discovery of and initial accounts of the animacy advantage in memory stemmed from concepts in evolutionary psychology (*cf.* Nairne et al., 2013, 2017; VanArsdall et al., 2013, 2015). Specifically, the tendency of our attention and memory systems to prioritize animate over inanimate things in our environment might have stemmed from fitness pressures faced by our early ancestors to quickly detect threats in the environment (predators, competitors) or to remember where sources of food or shelter were located. These tendencies might still exist today because they are deeply ingrained in our evolutionary heritage. Such ideas are quite viable: it is apparent that physical memory mechanisms can be conserved through evolution (Alberini, 1999), and different mouse strains exhibit differences in learning and memory performance (Wehner and Silva, 1996; Crawley et al., 1997). Genetic analyses are important and support adaptive accounts, but it is difficult to directly test the assumptions of such "ultimate" accounts using behavioral methods. Instead, behavioral researchers focus on testing "proximate" mechanisms for the effects of animacy on memory, testing mechanisms that can produce the effects now in modern humans, regardless of the ultimate origins of such tendencies.

Researchers have examined several potential proximate mechanisms that could cause or contribute to the animacy advantage in memory, but the exact cause(s) of the effect in free-recall performance is not yet apparent. The effect *does not* seem to occur because, compared to inanimate words, animate words are more threatening (Leding, 2019, 2020), more arousing (Meinhardt et al., 2018; Popp and Serra, 2018; Leding, 2019), more easily categorizable (Gelin et al., 2017; VanArsdall et al., 2017; Serra, 2021), or more likely to invoke mental imagery (Gelin et al., 2019; Blunt and VanArsdall, 2021; but see Bonin et al., 2015). Despite ample evidence that animate items attract attention compared to inanimate items (e.g., Altman et al., 2016; Bugaiska et al., 2019), research has not consistently found a relationship between attention capture and the animacy advantage in free-recall performance (*cf.* Bonin et al., 2015; Leding, 2019; Rawlinson and Kelley, 2021). It does not seem that most participants know about the animacy advantage (DeYoung and Serra, 2021; but see Li et al., 2016) or actively cause it to occur by purposely allocating greater processing to animate over inanimate items (Serra, 2021), but they *can* alter how they process animate and inanimate items to cause the effect to be larger or smaller if the task instructions lead them to do so (*cf.* DeYoung and Serra, 2021; Shull et al., n.d.). Some currently viable accounts of the animacy advantage in memory suggest that the effect could occur because animate items naturally trigger greater richness of encoding than do inanimate items, perhaps because animate items activate more related information (Meinhardt et al., 2020; Bonin et al., 2022) or have more semantic features (Rawlinson and Kelley, 2021). Of course, more than one factor could simultaneously contribute to the effect (Rawlinson and Kelley, 2021; Leding, 2022). For example, Meinhardt et al. (2020) suggested that animate items might first capture attention compared to inanimate items, which then leads to the preferential and deeper processing of animate items over inanimate items, leading to the memory difference.

The purpose of the present experiments was to examine the animacy advantage in free-recall performance using some novel methodological conditions for this topic. Most prior demonstrations of the animacy advantage in free-recall performance have involved the experimenter-paced (i.e., computer-paced) presentation of the stimuli during encoding. In both Experiments 1 and 2, we compared the effect under computer-paced and self-paced study conditions (between-participants). Given that processing-based differences between animate and inanimate items seem to contribute to the occurrence of the animacy advantage under computer-paced study conditions (i.e., Meinhardt et al., 2020; Rawlinson and Kelley, 2021; Bonin et al., 2022), self-paced study conditions could allow a greater opportunity to observe such differences. In Experiment 2, we measured participants' metacognitive beliefs about the animacy advantage just before they began the task to allow for a more fine-grained consideration of how people study the items.

Most often, participants devote greater study time to more difficult items than easier items (e.g., Nelson and Leonesio, 1988; Thiede and Dunlosky, 1999; Metcalfe, 2002). Such effects can be enhanced after participants gain experience with the items or task. For example, participants who gained experience with the effects of serial position on free-recall performance under computer-paced conditions later devoted greater study time to items in the middle of the list (those that were least likely to be learned) under self-paced study conditions (Murphy et al., 2022). Therefore, one possibility is that participants might devote more study time to inanimate over animate items, leading to a reduced memory advantage for animate over inanimate items, or perhaps even no difference or an *inanimate* memory advantage (*cf.* DeYoung and Serra, 2021; Shull et al., n.d.). A related possibility is that, as participants gain experience with the animacy advantage, they might begin to devote greater study time to inanimate than animate items across study-test trials. In Experiment 2, participants might devote more study time to the type of item they believe is *more difficult* to remember (i.e., to counteract expected differences by animacy), regardless of whether that belief reflects an animate or inanimate advantage.

In contrast, in some situations, participants devote more study time to items they perceive as easier to learn (Metcalfe, 2002; Serra and Dunlosky, 2010; Magreehan et al., 2016) or that have greater actual or perceived value (Dunlosky and Thiede, 1998; Murphy et al., 2023). As such, another possibility is that participants might devote more study time to animate over inanimate items, leading to a large memory advantage for animate over inanimate items (perhaps even larger than occurs under computer-paced conditions). In Experiment 2, specifically, participants might devote more study time to the type of item they believe is *easier* to remember, potentially causing the animacy advantage to become larger than normal if they believe that animate items are more memorable than inanimate items (*cf.* Shull et al., n.d.). In contrast, participants who believe that inanimate items are more memorable than animate items might devote study time preferentially to inanimate over animate items, reducing or even reversing the typical animacy advantage.

Of course, a third possibility is that participants in the self-paced conditions will devote study time equally to animate and inanimate items and the animacy advantage will occur anyway as in computer-paced conditions, which would be consistent with prior evidence that participants do not seem to purposely produce the animacy advantage by devoting greater or differential processing to animate over inanimate items (DeYoung and Serra, 2021; Serra, 2021). Such an outcome would also support the idea that item-level or processing-based differences between animate and inanimate items contribute to the animacy advantage, but likely in an unconscious way (DeYoung and Serra, 2021; Rawlinson and Kelley, 2021; Serra, 2021).

Finally, we also used three different sets of materials from earlier studies (i.e., Nairne et al., 2013; Popp and Serra, 2016, 2018) in the present experiments. By directly comparing the size of the animacy advantage for different sets of materials within the same experiments and samples, we can consider whether item-level properties of animate and inanimate items contribute to the effect in a bottom-up way, likely outside of participants' awareness or control (*cf.* DeYoung and Serra, 2021; see Serra and DeYoung, 2023, for a similar examination using paired-associates materials). To our knowledge, these are the first studies to directly compare the size of the animacy advantage in free-recall performance for different sets of materials within the same sample or experiment.

## 2. Experiment 1

### 2.1. Materials and method

#### 2.1.1. Participants

The participants were 210 undergraduate college students from the psychology participant pool at Texas Tech University. They participated for class credit.

We used PANGEA ("Power ANalysis for GEneral Anova designs") ver. 0.2 (Westfall, 2016) to perform power analysis for the present design. Several prior studies have demonstrated effects of animacy on free-recall performance yielding an effect size of *Cohen's* $d = 0.4$ or greater when using a single study-test trial of free-recall. Our studies, however, examined this effect over three study-test trials of the same items [as in Nairne et al. (2013)], which adds statistical power. Using just this aspect of the design (animacy by trial) and assuming an effect size of *Cohen's* $d = 0.4$ or greater for the main effect of animacy on recall, we would need 35 participants per group to have sufficient power (above 0.8) to detect such an effect within any single group. Therefore, this is the per-group sample size we used in both experiments (35 participants per group multiplied by six groups = 210 participants). That said, the present experiments also included two between-participants variables of interest: study method and list source. With 35 participants in each of the six groups, we would be sufficiently powered (above 0.8) to detect an overall effect size of animacy or trial on memory of *Cohen's* $d = 0.17$ or higher. More important, we would be sufficiently powered (above 0.8) to detect an effect size of list on memory of *Cohen's* $d = 0.25$ or higher, and an effect size of study method on memory of *Cohen's* $d = 0.20$ or higher.

We neglected to obtain demographics data for the participants in the present two experiments. That said, our participant pool is typically about 70% female and 30% male, of a mean age around 19 years old, and approximately 70% white or Caucasian, 20% Hispanic or Latin, and the rest identifying as another race or ethnic group (or as more than one race or ethnic group). Our samples likely had similar demographics.

#### 2.1.2. Materials

The study materials were three lists of animate words (e.g., duck, soldier, turtle) and inanimate words (e.g., hat, rake, violin) from three previously published papers on this topic: a list of 12 animate and 12 inanimate words from Nairne et al. (2013), a list of 84 animate and 84 inanimate words from Popp and Serra (2016), and a list of 40 animate and 40 inanimate words from Popp and Serra (2018). Some words appeared on more than one of the lists, but we assigned each participant to only study words from one list, so no word ever appeared more than once per participant. The word lists all appear in their entirety in the original papers, including values for the factors on which those authors balanced the animate and inanimate word lists. For ease of access and direct comparison, however, we provide summaries of these lists' properties in Table 1.

We also identified four words that were not on any of the word lists to serve as primacy and recency buffers. Specifically, for all participants, the words "goose" and "fork" were always the first two words presented on every trial, and the words "spoon" and "deer" were always the last two words presented on every trial, but we did not count participants' recall of these words (*cf.* Nairne et al., 2013). Prior research indicates that the animacy advantage in free recall is persistent across serial position (Serra, 2021), so excluding these words from analysis is not likely to alter the occurrence of the effect over the rest of the list.

The materials also included a custom computer program that presented all items for study and recorded participants' free-recall performance, as well as performed the consent process, and provided the instructions at the start of the task and a debriefing at the end. We created this program using LiveCode Ltd. (2019).

**TABLE 1** Attributes of the word lists.

| | Animate words | | Inanimate words | |
|---|---|---|---|---|
| | M | SD | M | SD |
| Nairne et al. (2013) List | | | | |
| Age of Acquisition | 2.8 | 1.0 | 2.7 | 0.8 |
| Category Size | 22.3 | 5.9 | 23.2 | 6.0 |
| Category Typicality | 0.2 | 0.2 | 0.2 | 0.2 |
| Concreteness | 593 | 29 | 592 | 17 |
| Familiarity | 504 | 70 | 507 | 31 |
| Imagery | 589 | 37 | 578 | 30 |
| Kučera-Francis Freq. | 21.7 | 23 | 16.5 | 16 |
| Meaningfulness | 448 | 56 | 438 | 32 |
| Number of Letters | 5.3 | 1.8 | 5.0 | 1.4 |
| Relatedness | 0.1 | 0.1 | 0.1 | 0.1 |
| Popp and Serra (2016) List | | | | |
| Concreteness | 6.3 | 0.3 | 6.3 | 0.2 |
| Google Frequency | $126.8 \times 10^6$ | $200.8 \times 10^6$ | $135.2 \times 10^6$ | $118.8 \times 10^6$ |
| Imagery | 6.1 | 0.5 | 6.1 | 0.3 |
| Number of Letters | 6.1 | 1.9 | 6.3 | 1.7 |
| Popp and Serra (2018) List | | | | |
| Age of Acquisition | 7.0 | 2.4 | 7.4 | 2.7 |
| Arousal | 4.0 | 0.7 | 4.0 | 0.8 |
| Concreteness | 4.6 | 0.3 | 4.6 | 0.2 |
| Dominance | 5.4 | 0.5 | 5.4 | 0.6 |
| Google Frequency | $187.8 \times 10^6$ | $324.7 \times 10^6$ | $215.6 \times 10^6$ | $267.4 \times 10^6$ |
| Number of Letters | 6.7 | 1.6 | 6.6 | 1.4 |
| Valence | 5.6 | 0.9 | 5.3 | 0.9 |

### 2.1.3. Design

The study involved a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) × 2 (study method: computer-paced vs. self-paced) mixed design. Word list and study method were between-participants factors; animacy and trial were within-participants factors.

### 2.1.4. Procedure

We based the procedure closely on that of Nairne et al. (2013), Study 2, see also the experiments in VanArsdall et al. (2017). For the conditions that studied the words from Nairne et al. (2013) with computer-paced study, the procedure was therefore nearly identical to the procedure in that original study except for the choice of the specific primacy and recency buffer items (the recall of which we did not score, as in that study).

Prior to the start of the session, we randomly assigned each participant to study words from one of the three lists and either under computer-paced or self-paced study conditions, with the restriction that we eventually assigned the same number of participants (35) to each of the six groups. When the procedure started, the computer program randomly chose 10 animate words and 10 inanimate words from the designated word list to serve as the participant's items for all three study trials. Participants in the same list condition were therefore unlikely to study the same exact subset of words as each other, although obviously there was less room for variance with the shorter Nairne et al. (2013) list than with the longer lists from Popp and Serra (2016) and Popp and Serra (2018). The program then conducted the informed-consent process and provided written instructions to the participants. The instructions explained that participants would study two-dozen items for a free-recall test, over three study-test trials, but made no mention of the animacy of the words. The instructions noted that participants did not need to recall the words in the order studied. The instructions informed participants in the computed-paced conditions that each word would appear on screen "for a few seconds," but informed participants in the self-paced conditions that they would control the study time of each item by clicking an on-screen icon to proceed to the next item.

The program then began the first study phase. Regardless of word list condition, the program first presented the two primacy buffer words, with a 250 millisecond interitem interval. The program then presented the participants' actual 20-word list, one at a time in a fully randomized order, with a 250 millisecond interitem interval. Finally, the program presented the two recency buffer words, one at a time, with a 250 millisecond interitem interval. For participants in the computer-paced conditions, buffer words and target words appeared on screen for study for five seconds each. For participants in the self-paced conditions, participants controlled how long a buffer or target word appeared on screen for study by clicking an on-screen icon to proceed to the next item. The computer program recorded the self-paced study time for each non-buffer item.

After studying all the words, participants completed a 60 s distracter task as in Nairne et al. (2013): the program showed the participant a random whole-number digit from 1 to 8 and the participant clicked an on-screen button to indicate whether the number shown was an odd or even number. After the distracter task, participants attempted to recall the words they previously studied by typing them into a field on the computer screen. The computer program displayed words already entered to the participants but did not provide any feedback regarding correctness. When participants felt they could not recall any more words, they clicked an icon on the screen to continue.

The procedure then repeated for two more trials, using a new random ordering for the 20 critical words on the subsequent study phases (but maintaining the same primacy and recency buffer items). After completing the third test, participants read a debriefing on the computer screen and then the researcher dismissed them.

## 2.2. Results

The data for Experiment 1 is available at https://osf.io/6kndh/.

### 2.2.1. Self-paced study time

We calculated the mean study time (measured in milliseconds then converted to seconds) for those participants in the self-paced conditions for animate and inanimate items on the three study trials (Table 2; Figure 1). We analyzed study time with a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) mixed ANOVA. Self-paced study time did not differ by animacy, $F(1,102) = 0.103$, $MSE = 1.645$, $p = 0.749$, $\eta_p^2 < 0.01$, or by list, $F(2,102) = 0.207$, $MSE = 21.382$, $p = 0.813$, $\eta_p^2 < 0.01$. Polynomial contrasts indicated that self-paced study time decreased both linearly, $F(1,102) = 80.322$, $MSE = 18.060$, $p < 0.001$, $\eta_p^2 = 0.44$, and quadratically, $F(1,102) = 37.149$, $MSE = 6.095$, $p < 0.001$, $\eta_p^2 = 0.27$, across the study trials. None of the interactions were significant.

We also considered whether study time differed for the computer-paced and self-paced conditions using one-sample $t$-tests. Compared to a fixed study time of five seconds per item on all trials for the computer-paced conditions, participants in the self-paced conditions spent an equivalent amount of time studying both animate and inanimate items on the first study trial (both $ps > 0.7$) but spent significantly less time studying both animate and inanimate items on the second and third study trials (all $ps < 0.001$). These values would remain significant after a Bonferroni correction ($\alpha = 0.008$).

### 2.2.2. Free-recall performance

We scored recall using a strict criterion as either correct or incorrect. We did not score participants' recall of the buffer words. We calculated the mean percentage of animate and inanimate words that participants correctly recalled on each trial (Table 2; Figure 1). We analyzed recall with a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) × 2 (study method: computer-paced vs. self-paced) mixed ANOVA. Although participants in the self-paced conditions spent less time studying the items on later trials than did those in the computer-paced conditions, overall levels of recall did not differ based on study method, $F(1,204) = 0.187$, $MSE = 2078.589$, $p = 0.666$, $\eta_p^2 < 0.01$. Recall also did not differ by list, $F(2,204) = 0.121$, $MSE = 2078.589$, $p = 0.886$, $\eta_p^2 < 0.01$. Participants recalled more animate than inanimate items, $F(1,204) = 119.469$, $MSE = 235.188$, $p < 0.001$, $\eta_p^2 = 0.37$. Animacy interacted with list, $F(2,204) = 6.150$, $MSE = 235.188$, $p = 0.003$, $\eta_p^2 = 0.06$. Follow-up comparisons indicated that participants recalled more animate than inanimate items for all three lists (all $ps < 0.001$); the effect size was largest for the Nairne et al. (2013) list (*Cohen's* $d = 1.08$) and smaller in comparison for the Popp and Serra (2016) list (*Cohen's* $d = 0.65$) and the Popp and Serra (2018) list (*Cohen's* $d = 0.54$). Polynomial contrasts indicated that recall increased both linearly, $F(1,204) = 726.191$, $MSE = 241.009$, $p < 0.001$, $\eta_p^2 = 0.78$, and quadratically, $F(1,204) = 25.674$, $MSE = 102.484$, $p < 0.001$, $\eta_p^2 = 0.11$, across the study trials. Unexpectedly, trial interacted with both study method, $F(2,408) = 11.877$, $MSE = 171.747$, $p < 0.001$, $\eta_p^2 = 0.06$, and list, $F(4,408) = 4.521$, $MSE = 171.747$, $p = 0.001$, $\eta_p^2 = 0.04$; both interactions suggest that gains in memory across trials varied somewhat by the study method and the materials studied. No other interactions were significant.

## 2.3. Discussion

As in past studies, the animacy advantage in free-recall performance occurred for participants in the computer-paced conditions. More important, it also occurred for those in the self-paced conditions, who allocated study time equally to animate and inanimate items. Participants in the self-paced conditions studied the items for less time overall than did those in the computer-paced conditions (and reduced their study time across the trials), but they achieved a comparable level of overall recall compared to participants in the computer-paced conditions. Although past research indicates that some extrinsic or processing differences between animate and inanimate items likely contribute to the animacy advantage (Meinhardt et al., 2020; Rawlinson and Kelley, 2021; Shull et al., n.d.), the present results are consistent with the prior conclusion that under typical settings participants do not seem to be producing this effect purposely, such as by intentionally devoting greater processing effort or depth of processing to animate over inanimate items (*cf.* DeYoung and Serra, 2021; Serra, 2021; Shull et al., n.d.). The self-paced study conditions in the present experiment presented an obvious opportunity for participants to devote greater study time to animate over inanimate items if they chose to do so, but that did not occur. There might still, however, be some conditions under which participants allocate study time differently to animate and

TABLE 2 Mean study time and mean free-recall performance in experiment 1.

| List and trial | Computer-paced conditions | | | | Self-paced conditions | | | |
|---|---|---|---|---|---|---|---|---|
| | Inanimate | | Animate | | Inanimate | | Animate | |
| | M | SD | M | SD | M | SD | M | SD |
| Study time (in seconds) | | | | | | | | |
| Nairne et al. (2013) List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 5.23 | 5.04 | 5.54 | 4.41 |
| Trial 2 | 5 | – | 5 | – | 2.07 | 2.14 | 1.80 | 1.30 |
| Trial 3 | 5 | – | 5 | – | 1.71 | 1.22 | 1.51 | 1.05 |
| Popp and Serra (2016) List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 5.07 | 4.38 | 5.47 | 5.33 |
| Trial 2 | 5 | – | 5 | – | 1.86 | 1.44 | 1.87 | 1.73 |
| Trial 3 | 5 | – | 5 | – | 1.11 | 0.68 | 1.23 | 0.91 |
| Popp and Serra (2018) List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 4.77 | 5.05 | 4.51 | 3.77 |
| Trial 2 | 5 | – | 5 | – | 2.24 | 1.68 | 1.96 | 1.13 |
| Trial 3 | 5 | – | 5 | – | 1.42 | 1.09 | 1.30 | 0.69 |
| Free-recall performance (% correct) | | | | | | | | |
| Nairne et al. (2013) List | | | | | | | | |
| Trial 1 | 24.05 | 17.36 | 38.10 | 15.56 | 27.38 | 21.54 | 39.05 | 23.46 |
| Trial 2 | 43.81 | 20.74 | 58.57 | 20.76 | 40.71 | 24.73 | 57.86 | 21.57 |
| Trial 3 | 56.19 | 20.45 | 68.76 | 21.96 | 52.86 | 23.65 | 64.05 | 23.20 |
| Popp and Serra (2016) List | | | | | | | | |
| Trial 1 | 24.76 | 17.33 | 36.43 | 16.05 | 35.71 | 27.31 | 42.14 | 27.19 |
| Trial 2 | 45.48 | 20.14 | 54.29 | 20.15 | 46.90 | 28.66 | 53.81 | 27.74 |
| Trial 3 | 56.90 | 24.63 | 64.05 | 20.69 | 56.43 | 29.43 | 59.29 | 29.83 |
| Popp and Serra (2018) List | | | | | | | | |
| Trial 1 | 25.00 | 9.26 | 32.86 | 14.98 | 25.48 | 14.28 | 32.62 | 20.45 |
| Trial 2 | 45.24 | 18.45 | 54.29 | 20.55 | 43.81 | 21.71 | 48.81 | 22.25 |
| Trial 3 | 63.57 | 24.18 | 70.24 | 19.42 | 54.29 | 21.52 | 62.38 | 24.37 |

Values are the mean study time (in seconds) of words of each type on each trial for participants in the self-paced conditions (those in the computer-paced conditions studied each item for 5 s each) and the mean percentage of words of each type that participants correctly recalled on each trial for all conditions.

inanimate items; we explore this possibility further in Experiment 2.

Direct comparison of the occurrence of the animacy advantage in free-recall performance for the three different word lists indicates that the effect was larger for the Nairne et al. (2013) list than for the other two lists (Popp and Serra, 2016, 2018). As the researchers who created those lists balanced the animate and inanimate items on different factors, it is possible that all three lists contain embedded confounding variables that could moderate the size of the animacy advantage in recall. This does not mean that the entirety of the animacy advantage stems from imbedded confounds between animate and inanimate items, but it does indicate that differences in intrinsic properties between these items can contribute to the effect, especially if left unchecked (*cf.* Popp and Serra, 2018).

## 3. Experiment 2

Participants in the self-paced conditions in Experiment 1 did not devote study time differently to animate versus inanimate items, so the occurrence of the animacy advantage in those experiments cannot be explained by differential study time (although other extrinsic or processing-based mechanisms could of course still have contributed).

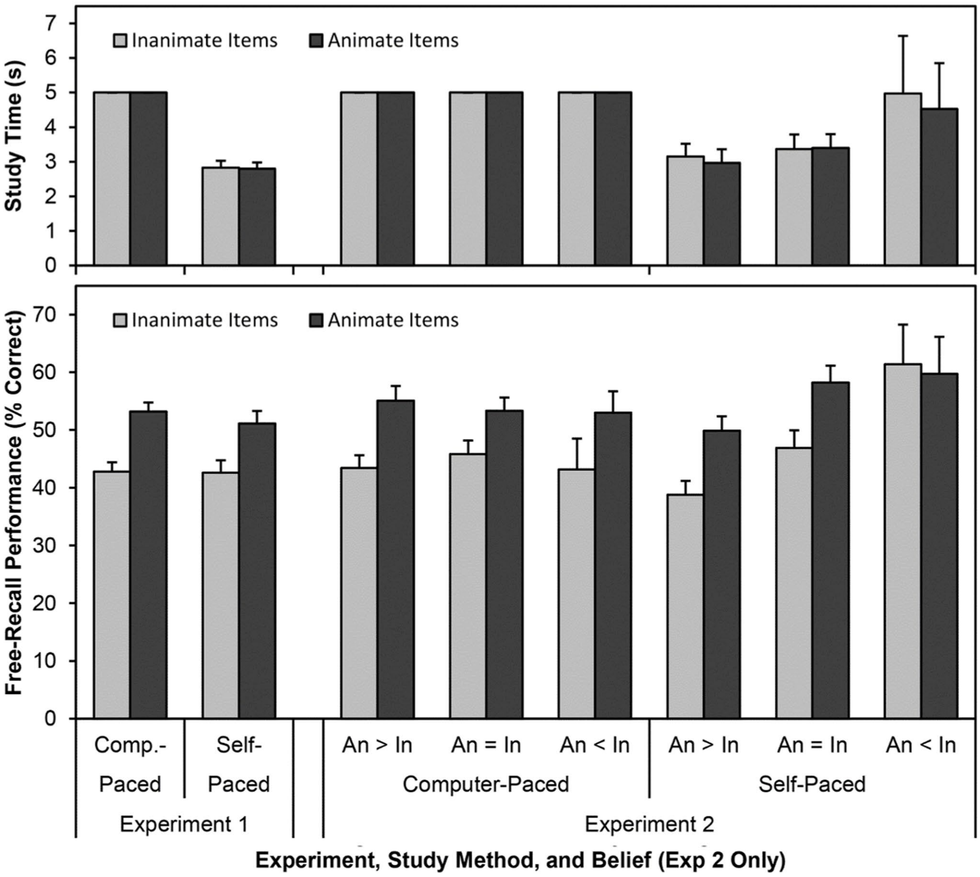


FIGURE 1
The mean study time (in seconds, top panel) and free-recall performance (percent recalled, bottom panel) for animate and inanimate words in Experiments 1 and 2, split by study method (computer-paced and self-paced). Results are collapsed on trial and list source. The results for Experiment 2 are split based on participants' self-reported beliefs about the effects of animacy on memory: animate items are more memorable than inanimate items ("An > In"), animate items are equally as memorable as inanimate items ("An=In"), and animate items are less memorable than inanimate items ("An < In"). Error bars are one standard error of the mean.

We found this outcome to be somewhat surprising, given some prior studies have found that task instructions could lead participants to alter the occurrence of the animacy advantage, *even under computer-paced study conditions*. For example, when researchers told participants to expect either an animate or inanimate advantage to occur, participants shifted their encoding to compensate for that expected outcome (DeYoung and Serra, 2021). When researchers told participants to purposely focus on encoding either animate or inanimate items, participants produced an animate or inanimate advantage in free-recall performance, respectively (Shull et al., n.d.).

It is possible that the participants in Experiment 1 did not devote study time differently to animate and inanimate items because their metacognitive beliefs about this effect were not activated prior to (or during) their study of the materials (*cf.* Dunlosky and Tauber, 2014; Tauber et al., 2019). In the present Experiment 2, we attempted to activate their pre-existing metacognitive beliefs about the effect prior to interacting with the materials using a simple metacognitive-beliefs question (a more subtle and perhaps more naturalistic approach than telling participants to expect or to purposely produce a given outcome). Although activating these beliefs prior to encoding could alter the occurrence of the animacy advantage under computer-paced conditions (*cf.* DeYoung and Serra, 2021; Shull et al., n.d.), there is an even greater opportunity for these beliefs to affect study time and the subsequent occurrence of the animacy advantage in the self-paced conditions. Knowing participants' metacognitive beliefs about the effects of animacy on memory can allow for a more nuanced consideration of the occurrence of the animacy advantage under either study method.

## 3.1. Materials and method

### 3.1.1. Participants

The participants were 210 undergraduate college students from the psychology participant pool at Texas Tech University. They participated for class credit. None had participated in Experiment 1.

Using the same power analysis as in Experiment 1, we again used 35 participants per group, and the same considerations of power would apply. In this experiment, however, we also considered whether participants' beliefs might interact with animacy. Assuming an even distribution of beliefs in the sample, we would be sufficiently powered to detect an interaction of animacy and beliefs with a *Cohen's* $d = 0.23$ or higher.

### 3.1.2. Materials

The study materials were the same as in Experiment 1.

### 3.1.3. Design

The primary design was the same as in Experiment 1: a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) × 2 (study method: computer-paced vs. self-paced) mixed design. The addition of the metacognitive beliefs question, however, allowed us to also examine participants' beliefs about the effect of animacy on free-recall performance as a group variable.

### 3.1.4. Procedure

The procedure for Experiment 2 was the same as for Experiment 1 except for the addition of a metacognitive beliefs question at the start of the task. Specifically, after reading the same set of instructions as in Experiment 1, participants in Experiment 2 read the question, "In this experiment, you will be studying a list of twenty-four words for a memory test. Half of the words will represent living (animate) concepts and half will represent non-living (inanimate) concepts. In this experiment, which statement below do you believe will be MOST ACCURATE?." They responded by picking one of the following options: "I think my memory will be better for living things than non-living things.," "I think my memory will be equal for living things than non-living things.," or "I think my memory will be better for non-living things than living things…" After answering this question, participants began the first study trial, and the rest of the procedure proceeded as in Experiment 1. Unlike in Experiment 1, participants in Experiment 2 were therefore made explicitly aware that half of the items would be animate, and half would be inanimate.

## 3.2. Results

The data for Experiment 2 is available at https://osf.io/6kndh/.

### 3.2.1. Beliefs question

We calculated the number and percentage of participants in the two study method conditions that endorsed each belief prior to beginning the task. In the computer-paced conditions, 41 participants (39.0%) endorsed the belief that animate items would be more memorable than inanimate items, 53 participants (50.5%) endorsed the belief that animate items would be equally as memorable as inanimate items, and 11 participants (10.5%) endorsed the belief that inanimate items would be more memorable than animate items. In the self-paced conditions, 50 participants (47.6%) endorsed the belief that animate items would be more memorable than inanimate items, 45 participants (42.9%) endorsed the belief that animate items would be equally as memorable as inanimate items, and 10 participants (9.5%) endorsed the belief that inanimate items would be more memorable than animate items. The proportions did not differ by study method, $X^2$ (2, $N = 210$) = 1.591, $p = 0.451$.

### 3.2.2. Self-paced study time

We analyzed study time (Table 3; Figure 1) with a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) mixed ANOVA. Self-paced study time did not differ by animacy, $F(1,102) = 1.119$, $MSE = 1.828$, $p = 0.293$, $\eta_p^2 = 0.01$, or by list, $F(2,102) = 0.123$, $MSE = 52.299$, $p = 0.885$, $\eta_p^2 < 0.01$. Polynomial contrasts indicated that self-paced study time decreased both linearly, $F(1,102) = 51.101$, $MSE = 38.491$, $p < 0.001$, $\eta_p^2 = 0.33$, and quadratically, $F(1,102) = 13.000$, $MSE = 12.908$, $p < 0.001$, $\eta_p^2 = 0.11$, across the study trials. None of the interactions were significant.

Compared to a fixed study time of five seconds per item for the computer-paced conditions, participants in the self-paced conditions spent an equivalent amount of time studying both animate and inanimate items on the first study trial (both $ps > 0.2$) but spent significantly less time studying both animate and inanimate items on the second and third study trials (all $ps < 0.001$). These values would remain significant after a Bonferroni correction ($\alpha = 0.008$).

We also repeated the ANOVA above, adding in participants' beliefs about the effect of animacy on memory (animate > inanimate vs. animate = inanimate vs. animate < inanimate) as a group variable (Figure 1). There was no difference in study time by belief, $F(2,96) = 0.805$, $MSE = 51.615$, $p = 0.450$, $\eta_p^2 = 0.02$, nor did belief interact with animacy, $F(2,96) = 1.476$, $MSE = 1.790$, $p = 0.234$, $\eta_p^2 = 0.03$. No other effects or interactions were significant either. That said, the main effect of animacy on study time approached significance after accounting for beliefs, $F(1,96) = 3.146$, $MSE = 1.790$, $p = 0.079$, $\eta_p^2 = 0.03$, with average study time being slightly higher for inanimate than animate items.

### 3.2.3. Free-recall performance

We analyzed recall (Table 3; Figure 1) with a 2 (animacy: animate vs. inanimate words) × 3 (trial: Trial 1 vs. Trial 2 vs. Trial 3) × 3 (word list source: Nairne et al., 2013 vs. Popp and Serra, 2016 vs. Popp and Serra, 2018) × 2 (study method: computer-paced vs. self-paced) mixed ANOVA. Although participants in the self-paced conditions spent less time studying the items on later trials than did those in the computer-paced conditions, overall levels of recall did not differ based on study method, $F(1,204) = 0.002$, $MSE = 1655.462$, $p = 0.963$, $\eta_p^2 < 0.01$. Recall also did not differ by list, $F(2,204) = 0.164$, $MSE = 1655.462$, $p = 0.849$, $\eta_p^2 < 0.01$. Participants recalled more animate than inanimate items, $F(1,204) = 114.145$, $MSE = 258.722$, $p < 0.001$, $\eta_p^2 = 0.36$. Animacy interacted with list, $F(2,204) = 6.358$, $MSE = 258.722$, $p = 0.002$, $\eta_p^2 = 0.06$. Follow-up comparisons indicated that participants recalled more animate than inanimate items for all three lists (all $ps < 0.001$); the effect size was again largest for the Nairne et al. (2013) list (*Cohen's* $d = 0.98$), slightly smaller for the Popp and Serra (2018) list (*Cohen's* $d = 0.84$), and smallest for the Popp and Serra (2016) list (*Cohen's* $d = 0.41$). Polynomial contrasts indicated that recall increased both linearly, $F(1,204) = 779.483$, $MSE = 126.073$, $p < 0.001$, $\eta_p^2 = 0.79$, and quadratically, $F(1,204) = 19.693$, $MSE = 148.718$, $p < 0.001$, $\eta_p^2 = 0.09$, across the study trials. Unexpectedly, the triple interaction between animacy, trial, and list was significant, $F(4,408) = 2.722$, $MSE = 111.149$, $p = 0.029$, $\eta_p^2 = 0.03$. No other interactions were significant, although the quadruple interaction between animacy, trial, list, and study method approached significance, $F(4,408) = 2.357$, $MSE = 111.149$, $p = 0.053$, $\eta_p^2 = 0.02$.

We also repeated the ANOVA above, adding in participants' beliefs about the effect of animacy on memory (animate > inanimate vs. animate = inanimate vs. animate < inanimate) as a group variable (Figure 1). A difference in the level of recall by beliefs approached significance, $F(2,192) = 2.486$, $MSE = 1606.375$, $p = 0.086$, $\eta_p^2 = 0.03$, as

TABLE 3 Mean study time and mean free-recall performance in experiment 2.

| List and trial | Computer-paced conditions | | | | Self-paced conditions | | | |
|---|---|---|---|---|---|---|---|---|
| | Inanimate | | Animate | | Inanimate | | Animate | |
| | M | SD | M | SD | M | SD | M | SD |
| Study time (in seconds) | | | | | | | | |
| Nairne et al. (2013)<br>List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 6.42 | 7.11 | 6.23 | 6.24 |
| Trial 2 | 5 | – | 5 | – | 2.85 | 3.36 | 2.61 | 3.35 |
| Trial 3 | 5 | – | 5 | – | 1.81 | 1.91 | 1.47 | 1.38 |
| Popp and Serra (2016)<br>List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 5.27 | 5.20 | 4.87 | 4.50 |
| Trial 2 | 5 | – | 5 | – | 2.82 | 2.45 | 2.88 | 3.88 |
| Trial 3 | 5 | – | 5 | – | 1.94 | 1.76 | 1.77 | 1.53 |
| Popp and Serra (2018)<br>List | | | | | | | | |
| Trial 1 | 5 | – | 5 | – | 6.21 | 8.33 | 6.34 | 7.52 |
| Trial 2 | 5 | – | 5 | – | 2.30 | 1.94 | 2.34 | 2.37 |
| Trial 3 | 5 | – | 5 | – | 1.15 | 0.68 | 1.24 | 0.79 |
| Free-recall performance (% correct) | | | | | | | | |
| Nairne et al. (2013)<br>List | | | | | | | | |
| Trial 1 | 26.19 | 17.75 | 45.95 | 18.34 | 29.29 | 24.20 | 42.62 | 26.18 |
| Trial 2 | 46.19 | 18.89 | 60.95 | 18.05 | 42.38 | 26.69 | 56.67 | 23.12 |
| Trial 3 | 62.38 | 21.90 | 70.71 | 19.11 | 55.24 | 25.57 | 64.76 | 21.40 |
| Popp and Serra (2016)<br>List | | | | | | | | |
| Trial 1 | 32.38 | 12.58 | 33.57 | 14.64 | 29.05 | 20.05 | 36.90 | 21.03 |
| Trial 2 | 49.52 | 20.00 | 54.29 | 21.80 | 47.62 | 19.23 | 53.33 | 20.83 |
| Trial 3 | 58.81 | 23.04 | 65.95 | 22.08 | 59.76 | 16.36 | 65.95 | 21.52 |
| Popp and Serra (2018)<br>List | | | | | | | | |
| Trial 1 | 22.38 | 13.37 | 35.00 | 12.59 | 31.90 | 21.81 | 38.81 | 25.08 |
| Trial 2 | 47.38 | 18.94 | 53.10 | 17.75 | 46.90 | 21.21 | 59.76 | 18.13 |
| Trial 3 | 56.19 | 21.80 | 66.43 | 20.16 | 57.62 | 22.81 | 70.71 | 18.46 |

Values are the mean study time (in seconds) of words of each type on each trial for participants in the self-paced conditions (those in the computer-paced conditions studied each item for 5 s each) and the mean percentage of words of each type that participants correctly recalled on each trial for all conditions.

overall recall was somewhat lower for those participants who endorsed the belief that animate items are more memorable than inanimate items. Follow up analyses indicated that the trend stemmed from the recall of inanimate items being lower for those who believed that animate items are more memorable than inanimate items than for those who believe that inanimate items are more memorable than animate items ($p = 0.031$) and those who endorsed no difference ($p = 0.092$). Put differently, participants who believed that animate items are more memorable than inanimate items had lower recall of *inanimate* items compared to those who believed otherwise. The interaction between animacy and beliefs approached significance, $F(2,192) = 2.866$, $MSE = 256.867$, $p = 0.059$, $\eta_p^2 = 0.03$, as did the interaction between animacy, study method, and beliefs, $F(2,192) = 2.523$, $MSE = 256.867$, $p = 0.083$, $\eta_p^2 = 0.03$. As well, the interaction between animacy, trial, study method, list and beliefs was significant, $F(8,384) = 2.719$, $MSE = 110.548$, $p = 0.006$, $\eta_p^2 = 0.05$. No other interactions with beliefs approached significance, and the inclusion of beliefs did not greatly alter other outcomes from the prior ANOVA.

Given there seemed to be a difference in whether the effect of animacy was altered by beliefs based on how participants studied the items, we performed two separate 2 (animacy) × 3 (trial) × 3 (word list source) × 3 (beliefs) mixed ANOVAs, split by study method (Figure 1). In this case, beliefs affected participants' overall level of recall under self-paced conditions, $F(2,96) = 3.790$, $MSE = 1953.195$, $p = 0.026$, $\eta_p^2 = 0.07$, but not under computer-paced conditions, $F(2,96) = 0.089$, $MSE = 1259.554$, $p = 0.915$, $\eta_p^2 < 0.01$. Whereas there were no significant differences in level of recall by belief within the computer-paced conditions (all $ps > 0.76$), the pattern within the self-paced conditions mirrored the pattern obtained in the larger ANOVA: the effect stemmed largely from the recall of inanimate items being higher for those who believed that inanimate items are more memorable than animate items than for those who believed that animate items are more memorable than inanimate items ($p = 0.002$), although recall was also higher for those who believed that inanimate items are more memorable than animate items compared to those who endorsed no difference ($p = 0.080$). As well, inanimate recall was somewhat higher for those who endorsed no difference compared to those who believed that animate items are more memorable than inanimate items ($p = 0.099$). Within the computer-paced conditions, animacy interacted with list, $F(2,96) = 3.568$, $MSE = 274.368$, $p = 0.032$, $\eta_p^2 = 0.07$, but not with beliefs, $F(2,96) = 0.760$, $MSE = 274.368$, $p = 0.470$, $\eta_p^2 = 0.02$. In contrast, within the self-paced conditions, animacy interacted with beliefs, $F(2,96) = 4.488$, $MSE = 239.366$, $p = 0.014$, $\eta_p^2 = 0.09$, but not with list, $F(2,96) = 1.184$, $MSE = 239.366$, $p = 0.310$, $\eta_p^2 = 0.02$. More specifically, in the computer-paced conditions, the effect size for animacy was largest for the Nairne et al. (2013) list (*Cohen's* $d = 1.06$), somewhat smaller for the Popp and Serra (2018) list (*Cohen's* $d = 0.77$), and smallest for the Popp and Serra (2016) list (*Cohen's* $d = 0.31$). In the self-paced conditions, the effect size for animacy was largest for participants who endorsed no difference (*Cohen's* $d = 0.93$), slightly smaller for those who believed that animate items are more memorable than inanimate items (*Cohen's* $d = 0.84$), and smallest (and trending in the other direction) for participants who believed that inanimate items are more memorable than animate items (*Cohen's* $d = -0.14$).

## 3.3. Discussion

Overall, the results of Experiment 2 were consistent with those of Experiment 1. We obtained the animacy advantage in free-recall performance, again regardless of study method (computer-paced versus self-paced study). Although participants in the self-paced conditions devoted less total study time to items compared to those in the computer-paced conditions, overall recall was again the same regardless of study method, and participants in the self-paced conditions again did not devote time differently to animate versus inanimate items.

In terms of participants' metacognitive beliefs, it does not seem that having participants self-report their metacognitive beliefs prior to beginning the task altered the occurrence of the animacy advantage in the computer-paced conditions, replicating outcomes from similar situations in some of the experiments reported by DeYoung and Serra (2021). In contrast, although participants in the self-paced conditions demonstrated about the same level of recall of animate items regardless of their beliefs, their recall of the inanimate items increased from those who endorsed the belief that animate items would be more memorable than inanimate items to those who endorsed the belief that animate items would be equally as memorable as inanimate items, and again to those who endorsed the belief that inanimate items would be more memorable than animate items. This pattern occurred without major differences in study time by group or animacy, which contradicts our prediction that participants might allocate their study time to the items differently based on their metacognitive beliefs. Nevertheless, these outcomes indirectly support extrinsic, processing-based accounts of the animacy advantage in free-recall performance, as participants with different beliefs about the effect of animacy on memory presumably processed the inanimate items differently in order to produce the different levels of recall of these items. As in several prior studies (e.g., Bonin et al., 2015; Leding, 2018; DeYoung and Serra, 2021; Shull et al., n.d.), the effects of this processing were most noticeable on inanimate items, again suggesting that animate items normally trigger greater processing than inanimate items, regardless of people's beliefs, but that people can increase their processing of inanimate items to remember more of them.

# 4. General discussion

## 4.1. Effects of study method

In both Experiments 1 and 2, we examined the animacy advantage in free-recall performance under computer-paced versus self-paced study conditions and using three different sets of animate and inanimate stimuli. We also considered whether participants' metacognitive beliefs (i.e., expectations) about the task—which we measured at the onset of the procedure in Experiment 2—altered the effects of animacy on study time or recall.

In both experiments, we obtained the animacy advantage in free-recall performance, regardless of whether participants studied the materials under computer-paced or self-paced conditions. Even though participants in the self-paced conditions spent less total time studying the items than did participants in the computer-paced conditions, the overall levels of recall and the occurrence of the animacy advantage were equivalent for the two study methods. Importantly, participants tended to devote equivalent study time to animate and inanimate items in the self-paced conditions, so the occurrence of the animacy advantage in those conditions cannot be attributed to a difference in study-time allocation (Surprisingly, total study time was lower for the self-paced conditions than for the computer-paced conditions, yet the overall level of recall was the same). The results indirectly support at least two prior suppositions about the animacy advantage: (1) under most conditions, participants do not seem to purposely produce the animacy advantage by devoting greater processing to animate over inanimate items (*cf.* DeYoung and Serra, 2021; Serra, 2021), and (2) animate items seem to trigger greater richness of encoding than do inanimate items (*cf.* Meinhardt et al., 2020; Rawlinson and Kelley, 2021; Bonin et al., 2022), likely outside of participants' awareness or control (DeYoung and Serra, 2021). Regarding this latter point, however, the present studies do not elucidate the form of this additional (or different) form of encoding or processing. Future research should continue to work to identify factors related to animacy that alter processing, or greater forms of processing such as spreading activation, that might contribute to the effect (e.g., Meinhardt et al., 2020; Rawlinson and Kelley, 2021).

## 4.2. Metacognitive beliefs

Much as in DeYoung and Serra (2021), in the present Experiment 2, participants' pre-existing beliefs about the animacy advantage did not relate to the occurrence of the effect when participants encoded the items under computer-paced study conditions, as well as for most participants in the self-paced conditions. That said, participants in the self-paced study conditions who believed that inanimate items would be more memorable than animate items studied all items for longer than did other participants, devoted equal study time to animate and inanimate items, and ultimately showed *equivalent* recall for animate and inanimate items. These results are in line with those of other studies that implicitly or explicitly encouraged participants to devote extra—and equivalent—processing effort to animate and inanimate items, resulting in a reduced or eliminated animacy advantage. For example, Bonin et al. (2015, Study 4) instructed their participants to produce interactive mental imagery during study, which reduced the size of the animacy advantage (primarily by increasing the recall of inanimate items compared to no interactive imagery instructions). Shull et al. (n.d.) crossed high and low point values (points earned for correctly recalling each item) with animate and inanimate items; an animacy advantage still occurred for low-value items, but the recall of high-value items was higher and showed no difference by animacy. Such outcomes suggest that—by default—animate items might trigger more processing than inanimate items, but conditions that encourage participants to devote greater or equivalent processing to inanimate items can reduce, eliminate, or even reverse (Shull et al., n.d.) the occurrence of the animacy advantage in free-recall. It is not immediately clear why participants who believed inanimate items were more memorable than animate items obtained equivalent recall for animate and inanimate items after studying them for the same amount of time, whereas most other participants studied animate and inanimate items for the same amount of time but still demonstrated an animacy advantage; presumably, this subset of participants processed the inanimate items in a different way than did other participants which led to enhanced recall for the inanimate items.

In the present experiments, we utilized participants' pre-existing beliefs about the effects of animacy on memory rather than trying to manipulate their beliefs. Although considering pre-existing beliefs prevents us from making causal conclusions and produces groups of unequal size, we know from prior research that trying to manipulate these beliefs is ineffective and produces unexpected effects. More specifically, DeYoung and Serra (2021, Experiment 2) attempted to manipulate participants' beliefs about the effects of animacy on free-recall performance by telling participants at the onset of the task to either expect to recall more animate than inanimate items, to recall more inanimate than animate items, to recall an approximately even number of animate and inanimate items, or they provided no expectation. In addition, their participants made metacognitive memory judgments for every item they studied. The provided beliefs did not affect participants' judgments of their memory (i.e., all groups judged animate items as more memorable than inanimate items regardless of the beliefs provided), which suggests that the provided information likely had little or no effect on their beliefs about animacy. Unexpectedly, however, the beliefs altered the effect of animacy on free-recall performance. The typical animacy advantage occurred for the group not given any expectation and for the group told to expect to recall more *inanimate* than animate items. Recall, however, did *not* differ by animacy for the group told to expect to recall more animate than inanimate items or the group told to expect no difference. Even though encoding was computer-paced in that experiment, participants were apparently able to alter their encoding effort or strategy to compensate for the provided outcomes to some extent (*cf.* Shull et al., n.d.).

## 4.3. List effects

The three lists we used in the present experiments all consistently produced an animacy advantage in free-recall performance, even when that effect was moderated by other factors. The size of that advantage, however, consistently differed by list (Tables 2, 3). Combining the experiments ($n = 420$), the Nairne et al. (2013) list produced the largest animacy advantage (*Cohen's* $d = 1.03$), the Popp and Serra (2018) list produced a somewhat smaller animacy advantage (*Cohen's* $d = 0.68$), and the Popp and Serra (2016) list produced the smallest effect (*Cohen's* $d = 0.52$). As previously shown in the context of paired associates learning (Serra and DeYoung, 2023), the selection of animate and inanimate stimuli can moderate the effects of animacy on memory. We do believe that animacy somehow aids the recall of single words in the free-recall paradigm, but researchers must understand that other factors besides animacy might also be contributing to the results obtained in any experiment. More positively, we hope that future research can more deeply examine how specific factors moderate the effects of animacy on memory, not just as potential confounding factors, but perhaps as hints to identify the mechanism(s) responsible for this effect. For example, as previously suggested by Popp and Serra (2018), the rather large animacy advantage produced by the Nairne et al. (2013) list might be partially attributable to the fact that those animate stimuli are more mentally arousing than are those inanimate stimuli, even though those Popp and Serra (2018) found an animacy advantage for items matched on arousal. As well, consider that the Popp and Serra (2016) list produced the smallest effect size in the present experiments, even though those animate and inanimate stimuli were matched on fewer factors than the Nairne et al. (2013) and Popp and Serra (2018) lists. Rather than producing a set of animate items that were favored for recall by other factors in addition to animacy, it is possible that the Popp and Serra (2016) list contains a confounding factor(s) that favors the recall of the *inanimate items,* reducing the obtained size of the animate advantage for this list. There are also some apparent differences in the attributes of the words across the three lists that could have contributed to the differing size of the animacy advantage across the lists (Table 1). For example, estimated age of acquisition is noticeably lower for words in the Nairne et al. (2013) list than in the other two lists. Although the animate and inanimate words were balanced on factors *within* each list, it is possible that animacy interacts with some of these factors in yet-unidentified ways. We recommend that researchers continue to examine the contribution of various factors besides animacy to recall using more continuous analyses such as regression or modeling (*cf.* Nairne et al., 2013; see also Gelin et al., 2017; VanArsdall and Blunt, 2022), and to consider whether those factors are independent from animacy or not.

## 4.4. Future directions

Going forward, we recommend that researchers consider two major classes of explanation for the animacy advantage: one class that focuses on *intrinsic* or *item-level* differences in memory-relevant factors that might exist between animate and inanimate words and could contribute to the effect, and one class that focuses on *extrinsic* or *processing* differences between animate and inanimate words (Rawlinson and Kelley, 2021, referred to such accounts as *controlled processing*) that could contribute to the effect. For example, many of the accounts for the animacy advantage which have already been discredited would fall into the intrinsic category: the effect does not seem to occur because animate and inanimate words differ on threat (Leding, 2019, 2020), arousal (Meinhardt et al., 2018; Popp and Serra, 2018; Leding, 2019), or categorizability (Gelin et al., 2017; VanArsdall et al., 2017; Serra, 2021). Some of the currently more viable accounts would fall into the extrinsic category: the effect might occur because animate items activate more related information (Meinhardt et al., 2020; Bonin et al., 2022) or have more semantic features (Rawlinson and Kelley, 2021). Admittedly, some factors, such as attentional capture, do not fall neatly into either an intrinsic or extrinsic category. As well, intrinsic and extrinsic mechanisms need not be exclusive and could even work together to produce the effect (*cf.* Meinhardt et al., 2020). Going forward, however, we think there can be value in separating tentative accounts of the animacy advantage into these two categories when possible. As a growing body of data discredits intrinsic or item-level accounts of the effect but supports extrinsic or processing-based accounts, it might be more efficient for researchers to focus on testing hypotheses for the effect that favor extrinsic differences between animate and inanimate words rather than intrinsic differences.

As well, researchers could examine whether the animacy advantage in memory differs across different levels and even "types" of animacy, rather than treating animacy as a living-vs.-nonliving binary as we did in the present report (and most other researchers have done as well). For example, VanArsdall and Blunt (2022) identified several subfactors related to people's concepts of animacy, such as thought, movement, reproduction, goal setting, and similarity to humans. These factors might relate to the memorability of a given concept to different degrees, or even interact.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/6kndh/.

## Ethics statement

The studies involving human participants were reviewed and approved by Texas Tech University IRB and HRPP. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

MS and CD devised the study, collected and analyzed the data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

## 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.

## Publisher's note

## References

Alberini, C. M. (1999). Genes to remember. *J. Exp. Biol.* 202, 2887–2891. doi: 10.1242/jeb.202.21.2887

Altman, M. N., Khislavsky, A. L., Coverdale, M. E., and Gilger, J. W. (2016). Adaptive attention: how preference for animacy impacts change detection. *Evol. Hum. Behav.* 37, 303–314. doi: 10.1016/j.evolhumbehav.2016.01.006

Blunt, J. R., and VanArsdall, J. E. (2021). Animacy and animate imagery improve retention in the method of loci among novice users. *Mem. Cogn.* 49, 1360–1369. doi: 10.3758/s13421-021-01175-0

Bonin, P., Gelin, M., and Bugaiska, A. (2014). Animates are better remembered than inanimates: further evidence from word and picture stimuli. *Mem. Cogn.* 42, 370–382. doi: 10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Laroche, B., Méot, A., and Bugaiska, A. (2015). The "how" of animacy effects in episodic memory. *Exp. Psychol.* 62, 371–384. doi: 10.1027/1618-3169/a000308

Bonin, P., Thiebaut, G., Bugaiska, A., and Méot, A. (2022). Mixed evidence for a richness-of-encoding account of animacy effects in memory from the generation-of-ideas paradigm. *Curr. Psychol.* 41, 1653–1662. doi: 10.1007/s12144-021-02666-8

Bugaiska, A., Grégoire, L., Camblats, A. M., Gelin, M., Méot, A., and Bonin, P. (2019). Animacy and attentional processes: evidence from the Stroop task. *Q. J. Exp. Psychol.* 72, 882–889. doi: 10.1177/1747021818771514

Crawley, J. N., Belknap, J. K., Collins, A., Crabbe, J. C., Frankel, W., Henderson, N., et al. (1997). Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology* 132, 107–124. doi: 10.1007/s002130050327

DeYoung, C. M., and Serra, M. J. (2021). Judgments of learning reflect the animacy advantage for memory, but not beliefs about the effect. *Metacogn. Learn.* 16, 711–747. doi: 10.1007/s11409-021-09264-w

Dunlosky, J., and Tauber, S. K. (2014). "Understanding people's metacognitive judgments: an isomechanism framework and its implications for applied and theoretical research" in *The SAGE handbook of applied memory*. eds. T. J. Perfect and D. S. Lindsay (Los Angeles, CA: SAGE), 444–464.

Dunlosky, J., and Thiede, K. W. (1998). What makes people study more? An evaluation of factors that affect self-paced study. *Acta Psychol.* 98, 37–56. doi: 10.1016/s0001-6918(97)00051-6

Félix, S. B., Pandeirada, J. N., and Nairne, J. S. (2019). Adaptive memory: longevity and learning intentionality of the animacy effect. *J. Cogn. Psychol.* 31, 251–260. doi: 10.1080/20445911.2019.1586716

Gelin, M., Bugaiska, A., Méot, A., and Bonin, P. (2017). Are animacy effects in episodic memory independent of encoding instructions? *Memory* 25, 2–18. doi: 10.1080/09658211.2015.1117643

Gelin, M., Bugaiska, A., Méot, A., Vinter, A., and Bonin, P. (2019). Animacy effects in episodic memory: do imagery processes really play a role? *Memory* 27, 209–223. doi: 10.1080/09658211.2018.1498108

Kazanas, S. A., Altarriba, J., and O'Brien, E. G. (2020). Paired-associate learning, animacy, and imageability effects in the survival advantage. *Mem. Cogn.* 48, 244–255. doi: 10.3758/s13421-019-01007-2

Leding, J. K. (2018). The animacy advantage in memory: manipulations of levels of processing and survival processing. *Am. J. Psychol.* 131, 273–281. doi: 10.5406/amerjpsyc.131.3.0273

Leding, J. K. (2019). Adaptive memory: Animacy, threat, and attention in free recall. *Mem. Cogn.* 47, 383–394. doi: 10.3758/s13421-018-0873-x

Leding, J. K. (2020). Animacy and threat in recognition memory. *Mem. Cogn.* 48, 788–799. doi: 10.3758/s13421-020-01017-5

Leding, J. K. (2022). Thinking of death and remembering living things: mortality salience and the animacy effect. *Memory* 30, 1258–1266. doi: 10.1080/09658211.2022.2102195

Li, P., Jia, X., Li, X., and Li, W. (2016). The effect of animacy on metamemory. *Mem. Cogn.* 44, 696–705. doi: 10.3758/s13421-016-0598-7

LiveCode Ltd. (2019). LiveCode [computer software]. Available at: https://livecode.com/ (Accessed April 10, 2023).

Magreehan, D. A., Serra, M. J., Schwartz, N. H., and Narciss, S. (2016). Further boundary conditions for the effects of perceptual disfluency on judgments of learning. *Metacogn. Learn.* 11, 35–56. doi: 10.1007/s11409-015-9147-1

Meinhardt, M. J., Bell, R., Buchner, A., and Röer, J. P. (2018). Adaptive memory: is the animacy effect on memory due to emotional arousal? *Psychon. Bull. Rev.* 25, 1399–1404. doi: 10.3758/s13423-018-1485-y

Meinhardt, M. J., Bell, R., Buchner, A., and Röer, J. P. (2020). Adaptive memory: is the animacy effect on memory due to richness of encoding? *J. Exp. Psychol. Learn. Mem. Cogn.* 46, 416–426. doi: 10.1037/xlm0000733

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *J. Exp. Psychol. Gen.* 131, 349–363. doi: 10.1037/0096-3445.131.3.349

Murphy, D. H., Friedman, M. C., and Castel, A. D. (2022). Metacognitive control, serial position effects, and effective transfer to self-paced study. *Mem. Cogn.* 50, 144–159. doi: 10.3758/s13421-021-01204-y

Murphy, D. H., Hoover, K. M., and Castel, A. D. (2023). Strategic metacognition: self-paced study time and responsible remembering. *Mem. Cogn.* 51, 234–251. doi: 10.3758/s13421-022-01307-0

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017). Remembering the living: episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory: the mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

Nelson, T. O., and Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect.". *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 676–686. doi: 10.1037/0278-7393.14.4.676

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

Popp, E. Y., and Serra, M. J. (2016). Adaptive memory: Animacy enhances free-recall but impairs cued-recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 186–201. doi: 10.1037/xlm0000174

Popp, E. Y., and Serra, M. J. (2018). The animacy advantage for free-recall performance is not attributable to greater mental arousal. *Memory* 26, 89–95. doi: 10.1080/09658211.2017.1326507

Rawlinson, H. C., and Kelley, C. M. (2021). In search of the proximal cause of the animacy effect on memory: Attentional resource allocation and semantic representations. *Mem. Cogn.* 49, 1137–1152. doi: 10.3758/s13421-021-01154-5

Scholl, B. J., and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309. doi: 10.1016/S1364-6613(00)01506-0

Serra, M. J. (2021). Animate and inanimate words demonstrate equivalent retrieval dynamics, despite the occurrence of the animacy advantage. *Front. Psychol.* 12:2027. doi: 10.3389/fpsyg.2021.661451

Serra, M. J., and DeYoung, C. M. (2023). Within-pair factors might explain the inconsistent effects of animacy on paired-associates recall. *Psychon. Bull. Rev.* 30, 688–699. doi: 10.3758/s13423-022-02184-z

Serra, M. J., and Dunlosky, J. (2010). Meta comprehension judgements reflect the belief that diagrams improve learning from text. *Memory* 18, 698–711. doi: 10.1080/09658211.2010.506441

Shull, J. N., Villalobos, N. C., Kelley, T. D., and Serra, M. J. (n.d.). (under review). Joint contributions of controlled processing and attentional prioritization to the animacy advantage in free-recall performance.

Tauber, S. K., Witherby, A. E., and Dunlosky, J. (2019). Beliefs about memory decline in aging do not impact judgments of learning (JOLs): a challenge for belief-based explanations of JOLs. *Mem. Cogn.* 47, 1102–1119. doi: 10.3758/s13421-019-00919-3

Thiede, K. W., and Dunlosky, J. (1999). Toward a general model of self-regulated study: an analysis of selection of items for study and self-paced study time. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1024–1037. doi: 10.1037/0278-7393.25.4.1024

VanArsdall, J. E., and Blunt, J. R. (2022). Analyzing the structure of animacy: exploring relationships among six new animacy and 15 existing normative dimensions for 1,200 concrete nouns. *Mem. Cogn.* 50, 997–1012. doi: 10.3758/s13421-021-01266-y

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Blunt, J. R. (2013). Adaptive memory. *Exp. Psychol.* 60, 172–178. doi: 10.1027/1618-3169/a000186

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Cogdill, M. (2015). Adaptive memory: Animacy effects persist in paired-associate learning. *Memory* 23, 657–663. doi: 10.1080/09658211.2014.916304

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Cogdill, M. (2017). A categorical recall strategy does not explain animacy effects in episodic memory. *Q. J. Exp. Psychol.* 70, 761–771. doi: 10.1080/17470218.2016.1159707

Vihman, V. A., and Nelson, D. (2019). Effects of animacy in grammar and cognition: introduction to special issue. *Open Linguistics* 5, 260–267. doi: 10.1515/opli-2019-0015

Wehner, J. M., and Silva, A. (1996). Importance of strain differences in evaluations of learning and memory processes in null mutants. *Ment. Retard. Dev. Disabil. Res. Rev.* 2, 243–248. doi: 10.1002/(SICI)1098-2779(1996)2:4<243::AID-MRDD8>3.0.CO;2-R

Westfall, J. (2016). PANGEA: Power ANalysis for GEneral Anova designs (v0.2) [computer software]. Available at: https://jakewestfall.shinyapps.io/pangea/ (Accessed April 10, 2023).

Zanini, C., Rugani, R., Giomo, D., Peressotti, F., and Franzon, F. (2020). Effects of animacy on the processing of morphological number: a cognitive inheritance? *Word Structure* 13, 22–44. doi: 10.3366/word.2020.0158

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 18 May 2023
DOI 10.3389/fpsyg.2023.1146200

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University, United States

REVIEWED BY
Christopher R. Madan,
University of Nottingham, United Kingdom
Toshiki Saito,
Waseda University, Japan

*CORRESPONDENCE
D. Stephen Lindsay
✉ slindsay@uvic.ca

Received 17 January 2023
ACCEPTED 13 April 2023
PUBLISHED 18 May 2023

CITATION
Mah EY, Grannon KEL, Campbell A, Tamburri N, Jamieson RK and Lindsay DS (2023) A direct replication and extension of Popp and Serra (2016, experiment 1): better free recall and worse cued recall of animal names than object names, accounting for semantic similarity. *Front. Psychol.* 14:1146200.
doi: 10.3389/fpsyg.2023.1146200

# A direct replication and extension of Popp and Serra (2016, experiment 1): better free recall and worse cued recall of animal names than object names, accounting for semantic similarity

Eric Y. Mah[1], Kelly E. L. Grannon[1], Alison Campbell[1],
Nicholas Tamburri[1], Randall K. Jamieson[2] and
D. Stephen Lindsay[1]*

[1]Department of Psychology, University of Victoria, Victoria, BC, Canada, [2]Department of Psychology, University of Manitoba, Winnipeg, MB, Canada

**Introduction:** Free recall tends to be better for names of animate concepts such as animals than for names of inanimate objects. In Popp and Serra's 2016 article, the authors replicated this "animacy effect" in free recall but when participants studied words in pairs (animate-animate pairs intermixed with inanimate-inanimate pairs) and were tested with cued recall, performance was better for inanimate-inanimate pairs than for animate-animate pairs ("reverse animacy"). We tested the replicability of this surprising effect and one possible explanation for the effect (semantic similarity).

**Methods:** Our Experiment 1 was a preregistered direct replication (*N* = 101) of Popp and Serra's Experiment 1 (mixed-lists condition). In a second preregistered experiment conducted in four different samples (undergraduate *N* = 153, undergraduate *N* = 143, online Prolific *N* = 101, online Prolific/English-as-a-first-language *N* = 150), we manipulated the within-category semantic similarity of animal and object wordlists.

**Results:** AIn Experiment 1, just as in Popp and Serra, we observed an animacy effect for free recall and a reverse animacy effect for cued recall. Unlike Popp and Serra, we found that controlling for interference effects rendered the reverse animacy effect non-significant. We took this as evidence that characteristics of the stimulus sets (e.g., category structure, within-category similarity) may play a role in animacy and reverse animacy effects. In Experiment 2, in three out of our four samples, we observed reverse animacy effects when within-category similarity was higher for animals and when within-category similarity was equated for animals and objects.

**Discussion:** Our results suggest that the reverse animacy effect observed in Popp and Serra's 2016 article is a robust and replicable effect, but that semantic similarity alone cannot explain the effect.

KEYWORDS

**adaptive memory, animacy, cued recall, free recall, direct replication**

# Introduction

Replicability is a hallmark of science. Direct replications are particularly valuable for assessing the report of new phenomena that are in some way surprising. Here we report a direct replication of an experiment reported by Popp and Serra (2016) having to do with memory for words that name animals (e.g., cat and dog) versus words that name inanimate objects (e.g., chair and hammer), followed by an experiment investigating one possible explanation for the effects observed by Popp and Serra. We begin by describing the context of Popp and Serra's research, then summarize their findings and explain why we believe it is important to assess their replicability. Then we report our findings and discuss their implications.

## Adaptive memory

According to evolutionary psychology, our perceptual and memory systems are adapted to notice and retain information with high survival relevance (Nairne et al., 2007; for a meta-analysis, see Scofield et al., 2017). Current accounts suggest that survival-relevant information recruits additional processing (e.g., elaboration, deep processing, simulation) that is less extensively used with other non-survival information (Kazanas et al., 2020). This argument of evolutionary psychology led to the development of the *survival processing* experimental paradigm (Nairne et al., 2007). In this paradigm, participants study stimuli in a survival context (or in a non-survival control context) and are subsequently tested on memory for the stimuli. As an example, some participants may be told to imagine that they are in a dangerous grassland and to think about study-list items in terms of their relevance to surviving in such an environment (Kazanas et al., 2020) or instructed to rate stimuli on their survival relevance (Schwartz and Brothers, 2014), whereas subjects in a control condition complete a standard study-test phase or imagine a scenario with low survival relevance (e.g., moving to a new home; Nairne et al., 2007). Recall performance is generally observed to be superior when items are encoded in a survival orientation.

## Animacy and free recall

In studies of survival processing and memory, subjects are instructed to process study-list items in survival-relevant ways or in some other comparably "deep" way that is not related to survival. In a related but distinct line of work (again initiated by James S. Nairne), researchers have compared memory for names or images of animate things (e.g., animals/humans) versus names or images of inanimate objects (e.g., tools/toys). For example, Nairne et al. (2007) sought to examine the effect of animacy on memory by asking participants to free recall a randomly intermixed list of words representing animate (e.g., baby, soldier, duck) and inanimate (e.g., doll, purse, hat) items and discovered that participants recalled more animate than inanimate items. In related work, VanArsdall et al. (2013) demonstrated a corresponding animacy advantage in free recall when nonwords were associated with animate versus inanimate features. That is, when a made-up nonword was paired with a living property (e.g., "FRAV dislikes tomatoes"), free recall for that nonword was better than when it was paired with a non-living property (e.g., "FRAV runs on gasoline"). This "animacy effect" has been well-documented across a variety of designs and tasks, including recall and recognition memory for both word and picture stimuli (Bonin et al., 2014; Scofield et al., 2017).

## Animacy and cued recall

In contrast to the consistent finding of an animacy advantage in free recall and recognition, studies of the animacy effect in paired-associate cued recall have yielded mixed results. Initial findings suggested that participants were better at learning Swahili-English pseudo-vocabulary translations when a Swahili stimulus word was randomly paired with an English animal name compared to when it was matched with an English object name (VanArsdall et al., 2015). Popp and Serra (2016) noted that vocabulary translation tasks such as that used by VanArsdall et al. (2015) differ from other forms of paired-associate learning and that variables that affect one operationalization of paired-associate learning do not necessarily affect other operationalizations. To address this, Popp and Serra' Experiment 1 examined participants free and cued recall performance for pairs of English words using 84 animal names and 84 inanimate object names matched on a number of relevant features including number of letters, imagability, concreteness, and frequency. When subjects studied and attempted free recall of individual words, performance was better for animate than inanimate words (i.e., an animacy effect in free recall). However, when they studied animate-animate and inanimate-inanimate pairs, subsequent cued recall of targets was better for the inanimate-inanimate pairs (i.e., a reversed animacy effect in cued recall). Popp and Serra reported converging evidence for this reversed animacy effect in two additional experiments (although, in Experiment 2 they obtained ambiguous results in a condition in which one member of the pair was a Swahili word and the other an animate or inanimate English word).

Kazanas et al. (2020) also reported converging evidence for a reverse animacy effect on cued recall using English-Spanish translation pairs. They reported an experiment in which they orthogonally varied orientation (survival vs. controls) and animacy (animate vs. inanimate). English-speaking monolinguals studied recordings of spoken English-Spanish translation pairs (e.g., cat: gato) with varying instructions as to how to think of the words during study. Later, participants were tested on sentence completion, matching, or picture naming. On all of these associative-learning tasks, performance was better for inanimate than animate pairs. The findings of Kazanas et al. contrast with those of VanArsdall et al. (2015)–who, as noted above, observed an animacy advantage with cued recall of English-Swahili translation pairs–but converge with the Popp and Serra (2016) findings.

In contrast with these findings in support of the reverse animacy effect, recent experiments from DeYoung and Serra (2021, Experiments 4 & 5) with similar procedures and same/different wordsets did not yield a reverse animacy effect. In one experiment, the authors failed to replicate the reverse animacy effect with the Popp and Serra (2016) wordset, and in both experiments the authors observed an animacy advantage for cued recall with a new wordset. These results raise several possibilities, the most salient being that (a) the reverse animacy effect is not replicable, and/or (b) wordset-specific characteristics *other than* animacy explain or moderate the effect. We conducted two experiments to address these possibilities. Experiment 1 is a direct replication of Popp and Serra (2016), and Experiment 2 is an extension of their experiment that examines an alternate explanation for the reverse animacy effect in cued-recall.

# Experiment 1

To the best of our knowledge, no direct replication of the Popp and Serra (2016) experiment has been reported. DeYoung and Serra (2021; Experiment 5) used the Popp and Serra (2016) wordset in a cued recall-only experiment and failed to observe an animacy advantage or disadvantage. Kazanas et al. (2020) observed a reverse animacy effect, but their study was far from a *direct* replication of the original Popp and Serra design. As Popp and Serra noted, language-translation tasks differ from other paired-associate learning tasks. Also, Kazanas et al. used only 12 words from each category.

The animacy advantage in free recall has been reported by multiple labs using different sets of materials and a variety of procedures, and while the animacy literature thus far has allowed researchers to hypothesize mechanisms (e.g., mental imagery, attention, semantic features; Bonin et al., 2015, and Xiao et al., 2016), the matter of underlying mechanisms is far from settled. Because most previous animacy-memory studies have shown a memory advantage for animate stimuli, and because proposed underlying theories (i.e., adaptive memory) predict a general animacy advantage, the observation of an animacy *disadvantage* in cued recall warrants verification. If a reverse animacy effect for cued recall proves to be robust, that may help advance our understanding of the mechanisms underlying the relationship(s) between animacy and memory. Thus, the principal aim of this experiment was to replicate the key finding from Popp and Serra's Experiment 1 (i.e., an animacy effect in free recall paired with a reversed animacy effect in cued recall). We also collected self-report measures regarding participants' perceptions of the experiment.

## Method

The plans for this experiment were preregistered on the Open Science Framework[1] with all study materials, the program used to collect the data, details of the procedure, and the specifications for the planned statistical analyses; see https://osf.io/hcp4m. The data and scripts used to process/analyze the data can be accessed at https://osf.io/pbec9/. We also include a report from the Transparency Checklist (Aczel et al., 2020) that may be useful in assessing our preregistration.

## Design

The experiment conformed to a 2 (animacy: animals, objects) × 2 (memory type: free recall, cued recall) within-subjects design, with the main dependent variable being recall accuracy. Popp and Serra tested half of their Experiment 1 subjects with study lists that intermixed names of animals and objects and half with blocked lists (i.e., all animals for one study-test cycle, all objects for another). In our experiment, we used only mixed lists because we were not particularly interested in list type (which had no statistically significant effects in Popp and Serra).

### Sample size planning

When planning a direct replication for which the primary outcome will be based on null hypothesis significance testing (NHST), it is desirable to test a sufficiently large sample of subjects to attain high statistical power to detect the hypothesized effect, if it is real. Sample-size planning is complicated by the fact that publication bias favors the publication of large effects. Consequently, sample-size plans based on published literature may have low power to detect the average effect of the manipulation in question (Anderson et al., 2017). Additionally, for under-studied effects (such as the reverse animacy effect), there are few effect size estimates upon which to rely. We used Simonsohn's (2015) "small telescopes" approach, which suggests setting a sample size of 2.5 times that of the to-be-replicated study. According to Simonsohn, this gives the replication about 80% power to reject the null hypothesis of a detectable effect (i.e., an effect that the original study had 33% power to detect). In other words, the "small telescopes" heuristic allows one to power replication studies for effects that would have been minimally detectable in the original study (rather than observed effects, which may be inflated). Following this heuristic, we set our minimum sample at 2.5 times the number of subjects in Popp and Serra (2016), Experiment 1 mixed-list conditions ($N = 36$), at 90 participants. We also preregistered that if more than 90 subjects met the inclusion criteria at the end of the available data-collection period they would be included in our analyses.

## Sample

Participants ($N = 104$) were recruited via our university's psychology research participation pool. All members of the participation pool were eligible to participate. Age in our sample ranged from 18 to 36 years old ($M = 21.2$, $SD = 3.8$). We did not collect information on gender or ethnicity, but participants drawn from this pool tend to self-identify as female (72%) and European/Caucasian (71%). We asked participants whether they learned English as a first language, second language, or simultaneously with another language: 82% of participants reported English as a first language, 7% reported English as a second language, and 11% reported bilingual English and another language. Participants were compensated with optional extra credit in a psychology course. Data were collected from 104

---

[1] We created a preregistration prior to data analysis, but due to technical issues this registration was not archived and we had to upload a copy of the registration after we had conducted our analyses. The new copy of the preregistration was identical to the pre-analysis version.

participants, but three participants were excluded from analysis based on preregistered exclusion criteria and so the final sample size was 101.[2]

## Materials

The computer program used in this study was generously provided by Michael Serra of Popp and Serra (2016). We made a few minor modifications to the instructions and to the informed consent statement and added some self-report items after the main task. The LiveCode (https://livecode.com/) program for the experiment is accessible at https://osf.io/jpd5k. The Popp and Serra program included 84 each of animal and object words, with lists matched on mean number of letters, mental imagery, concreteness, and word frequency. Animal words included mammals, insects, reptiles, amphibians, birds, and fish. Object words included household objects, tools, instruments, clothing, sports-related objects, appliances, and some miscellaneous objects (e.g., anchor, cannon). Readers can refer to the original Popp and Serra article for a detailed description of words used in this experiment and how they were selected.

## Procedure

We posted the study on our university's online psychology research participation system that students use to sign up and participate in psychology experiments for optional bonus points in certain courses. Students who signed up to participate were given instructions and a Dropbox link to download the experiment program. After completing the experiment remotely on their own computers participants emailed us their data files. The remote/online nature of our replication differed from the original study, and was a result of COVID-19 closures.

The program began by presenting a description of the study and inviting informed consent. Participants were given an option to withdraw from the study on the consent form, in which case the program collected no data. Participants could also withdraw at any point by simply opting not to send us their data. Due to the remote nature of the data collection, we do not know how many participants opted to withdraw. Of our final sample, 53 participants were randomly assigned to perform two free recall study/test blocks followed by two cued-recall study/test blocks; the other 48 did the two types of study/test blocks in the opposite order.

In each study/test block, the program first presented study instructions. For free recall study/test blocks, participants were told that a series of nouns would be presented one at a time and that they would later be asked to recall those words in any order. For cued recall study/test blocks, they were told that a series of pairs of nouns would be presented and that they would then be shown the first word in each pair and invited to recall its partner. The program then presented study items (words or pairs) one at a time for 5 s each, preceded by a 1-s fixation cross. For each participant and for each study list, the program randomly selected 15 words (or word pairs) from the stimulus set without replacement, under the constraint that across the two study lists of a given type there be a total of 15 animal names (or pairs) and 15 object names (or pairs). For instance, if a participant's first free recall list contained 5 animals and 10 objects, the second free recall list would contain 10 animals and 5 objects[3]. Then the test instructions were presented. For free recall, participants were instructed to type in all the words they could remember from the previously studied list in any order, pressing enter after each word they recalled. For cued recall, participants were informed that the first word of each studied pair would be presented on the computer screen one-at-a-time and that they were to type in the target word of the pair, after which they pressed enter to proceed to the next cue word. For both types of tests, participants could see all of their responses throughout the test. Participants were told they could guess or leave free and cued responses blank, and were given unlimited time to complete each test phase. The program recorded reaction times as a matter of completeness (i.e., time from test start or last word submission to current word submission).

After pressing enter on the last cued recall item on the cued recall test, participants were automatically shown the study instructions for the next study/test cycle. When participants decided that they had completed a free recall test, they clicked a "Finished" button to advance to the next study instructions. This process repeated until the participant had completed all four study/test blocks. Other than minor changes to the instructions and an informed consent statement, the foregoing parts of the procedure were identical to the procedure used by Popp and Serra.

Following the experiment, participants were invited to enter their age in years and to indicate whether or not English was their first language[4]. We then asked questions that assessed participants' (a) awareness of the animal/object categories, (b) study and test strategies, and (c) perceived relative difficulty of recalling animals versus objects for each type of test (Table 1). Participants were also asked to indicate whether they experienced distractions during the experiment, with the options "No distractions," "Minor/brief/few distractions," "A major distraction that affected my ability to pay attention to the experiment." After answering all of the questions, participants were debriefed and thanked for their participation.

---

[2] Two of these participants did not correctly recall at least one word from both free recall lists ($N=2$), and one participant reported experiencing a "major distraction that affected your ability to pay attention to the experiment." There were no missing/incomplete data in our final sample.

[3] In our preregistered plan, we initially intended that the ratio of animal:object words in each list be restricted to 7:8 and 8: 7, and mistakenly assumed that the program (and original experiment) worked this way. However, our replication (and the original experiment) included variable ratios as extreme as 4: 11. Along with testing additional hypotheses, we addressed this in three of the samples collected for Experiment 2.

[4] We did not plan or preregister any analyses of these demographics variables, and included them only to provide a picture of our sample and to allow other researchers to examine these variables.

TABLE 1 Self-report strategy questionnaire.

| | |
|---|---|
| 1. | "Please describe the words you studied. That is, what characteristics, properties, or attributes did the words have?" |
| 2. | "Some of the words were names of animals and others were names of inanimate objects. Did you notice that fact when you were studying the words?" [Definitely not/Maybe not/Do not know/Maybe yes/Definitely yes] |
| 3. | "When words were presented one at a time for study, what if anything did you do to try to remember them?" |
| 4. | "When words were presented one at a time for study, did you use the same strategy for animals and objects, or different strategies for animals versus objects? If you used different strategies, please describe them below." |
| 5. | "When tested on free recall of words that you had studied one at a time, did you find it easier to recall one category of words than the other? Drag the slider below to indicate the relative ease of remembering words from each category" [Objects much easier vs. Animals much easier] |
| 6. | "When words were presented as pairs for study, what if anything did you do to try to remember them?" |
| 7. | "When words were presented as pairs for study, did you use the same strategy for animal pairs and object pairs or different strategies for animal pairs versus object pairs? If you used different strategies, please describe them below." |
| 8. | "When tested on recall of words that you had studied in pairs, did you find it easier to recall one category of word pair than the other? Drag the slider below to indicate the relative ease of remembering words from each category." [Objects much easier vs. Animals much easier] |

# Results

## Analytic strategy

Our primary analyses were NHST ANOVAs and follow-up *t* tests, mirroring the analyses of the original experiment. However, we supplemented those frequentist analyses with Bayesian analogs (see Supplementary online material Section 1A for more details). We applied these analyses to the following 6 questions. Did we replicate the animacy and reverse animacy effects observed in Popp and Serra (2016), using their scoring method? Did we replicate the animacy and reverse animacy effects observed in Popp and Serra (2016), using a more involved manual scoring method? Does paired-associates interference (e.g., differential benefits of guessing for animals vs. objects) account for the effects we observed? Were most participants aware of animal and object categories, and if so, how specific or general were the categories they perceived? Did category awareness relate to observed animacy effects? What memory strategies did participants use for free and cued recall, and did participants use different strategies for animals and objects (and if so, did this relate to observed animacy effects)? Were participants sensitive to differences in recall difficulty (e.g., animals easier in free recall, harder in cued recall) at a metacognitive level? Did participant self-reports of relative recall difficulty map onto actual performance (i.e., were participants calibrated)?

## Primary analysis

We assessed recall accuracy with the scoring method used by Popp and Serra (2016). For free recall, responses that exactly matched a study-list word were automatically counted as correct. All other responses were judged by two independent scorers (blind to type of test) and responses that both scorers judged to be acceptable matches to a study-list word (e.g., "harpsicord" for "harpsichord") were counted as correct (103 cases, or 3.3% of all free recall trials). When there was disagreement between the two scorers about the match of the response word and a word from the study list (10 cases), a third independent scorer who was blind to the test condition resolved it. Free recall performance was operationalized as the proportion of animal and object words that participants correctly recalled (out of 15 each of studied animal and object words) across the two lists. For cued recall, response words were computer scored: If the first three letters of the response matched the first three letters of the studied target word, the response was counted as correct. Cued recall performance was operationalized as the proportion of animal–animal and object-object pairs that participants correctly completed (out of 15 each of studied animal–animal and object-object pairs) across the two lists of pairs. Figure 1 shows the means, distributions, and data points for each condition.

Proportion of targets accurately recalled was analyzed with a 2 (animacy: animals, objects) × 2 (memory type: free recall, cued recall) within-subjects ANOVA. We also evaluated main and interaction effects via Bayes Factor (BF)[5] analysis. Consistent with the results of Popp and Serra (2016), mean proportion of targets accurately recalled was higher for free recall ($M = 0.61$, $SD = 0.19$) than cued recall ($M = 0.51$, $SD = 0.28$), $F(1, 100) = 21.04$, $p < 0.001$, $\eta_p^2 = 0.17$. The corresponding BF for this analysis was >100; "extreme" evidence in favor of a memory type main effect (Wagenmakers et al., 2011). Also replicating Popp and Serra, the main effect of animacy was not significant, $F(1, 100) = 1.68$, $p = 0.20$, $\eta_p^2 = 0.02$ (BF = 0.6; anecdotal evidence against an effect), and the interaction between animacy and memory type was significant, $F(1, 100) = 41.70$, $p < 0.001$, $\eta_p^2 = 0.29$ ($BF > 100$). Paired-samples *t* tests[6] (Bonferroni corrected) were used to examine the interaction. This revealed that mean free-recall proportion correct for animals ($M = 0.64$, $SD = 0.17$) was better than that for objects ($M = 0.58$, $SD = 0.21$), $t(100) = 3.24$, $p = 0.002$, $d_z$[7] $= 0.32$ [0.12, 0.52]. By contrast, mean cued-recall proportion correct was

---

[5] These Bayes Factors captured ratios of prior-weighted marginal likelihoods for models with and without effects of interest. In other words, the ratio of how well one model explains the data relative to another.

[6] To account for apparent non-normality in proportion correct, we also conducted nonparametric Wilcoxon signed-rank tests. In all cases, the parametric and nonparametric analyses agreed (i.e., both were either significant or non-significant).

[7] $d_z$ = Cohen's *d* for within-subjects comparisons, with the 95% Confidence Interval on the effect size estimate in brackets.

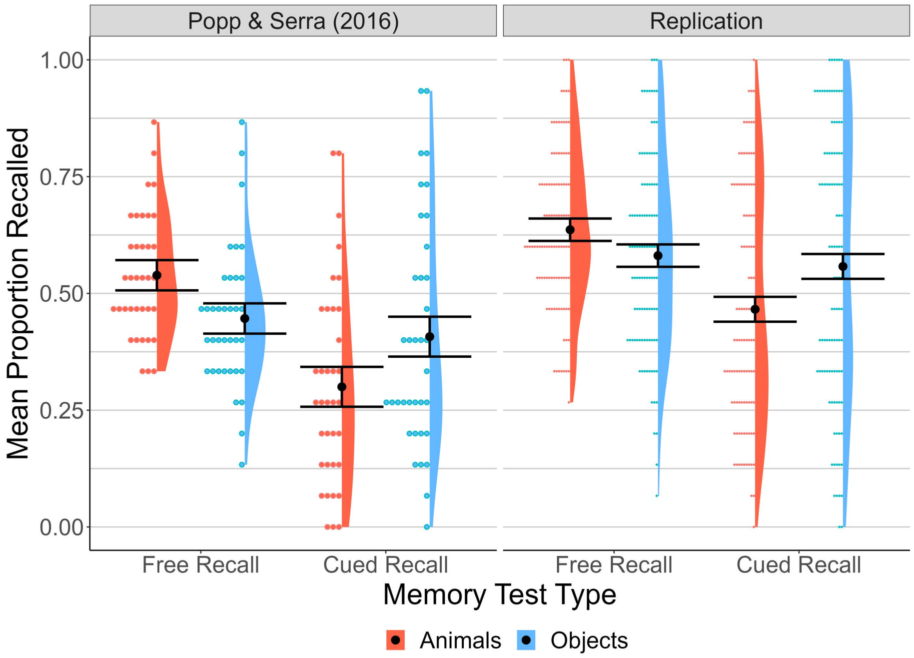


FIGURE 1
Proportion of targets correctly recalled by Memory Test Type and Animacy. Means, distributions, and individual participant data points for proportion of targets correctly recalled across memory test type and animacy conditions for the original study data (Popp and Serra, 2016, Exp 1, mixed-list condition) and the current replication. Error bars represent 95% within-subject confidence intervals based on the animals versus objects comparison for each memory test type (calculated as per Loftus and Masson, 1994).

better for object pairs ($M = 0.56$, $SD = 0.29$) than for animal pairs ($M = 0.47$, $SD = 0.25$), $t(100) = 4.84$, $p < 0.001$, $d_z = 0.48$ [0.27, 0.69].

## Alternative analyses

We explored an alternate scoring procedure in which cued recall responses (as well as free recall responses) were scored manually. As detailed in Supplementary online material (Section 1B), results with this measure mirrored those reported above. We also report in Supplementary online material (Section 1C) an unplanned 2 (Animacy) × 2 (Memory Test Type) × 2 (Test Order) × 2 (Study/Test Block) ANOVA to examine all test-related effects. We coded and analyzed qualitative self-reported recall strategy data, but found no clear relationships between strategy use and animacy/reverse animacy effects (Supplementary online material 1F.4). We also analyzed self-reported recall difficulty, and found that participants showed some (but imperfect) metacognitive awareness of animacy/reverse animacy effects (e.g., objects rated as more difficult than animals, moreso for cued recall; mixed evidence for a relationship between reported recall difficulty and actual performance; Supplementary online material 1F.5). Finally, we conducted several exploratory analyses of reaction time (e.g., reaction times for correct answers vs. commission errors vs. omission errors)—these can be found in Supplementary online material 1F.4 (Section 1D). Below, we briefly discuss two additional alternative analyses relevant to potential explanations for the reverse animacy effect–paired-associates interference and category awareness data.

## Paired-associates interference

Popp and Serra (2016) noted that subjects might report words that they recalled as guesses. Even a word presented on the study list might be generated as a guess (e.g., if it was not encoded at study). Moreover, Popp and Serra speculated that their animate names might be a "narrower," more closely associated set than their inanimate names. That would promote guessing of studied animate names, relative to guessing of studied inanimate names. On the free recall test, such guesses would inflate performance of animate names despite being "lucky intrusions." On the paired-associates recall test, in contrast, guessing would be less helpful, and might even interfere with correct report of animate names. For one thing, most studied words generated as paired-associate guesses would have been studied with a different cue word. For another, having a guessed word come to mind might interfere with recall of the target. Popp and Serra referred to this as paired-associate interference, and explored the possibility of interference via an exploratory analysis in which any incorrect animal/object word recalled in place of a correct animal/object target was counted as correct, irrespective of whether the incorrect word was studied or not. For example, if a participant studied the cue-target pair "glasses – motorcycle" but then responded to the test cue "glasses" with the previously non-studied "car," this was counted as correct. Popp and Serra found that the cued recall reverse animacy effect was still present in their overall sample when interference responses (i.e., same-category commission errors) were counted as correct.

We applied a similar analysis to our own data, and found that when treating same-category cued recall commission errors as correct, the reverse animacy effect was no longer significant (with Bayesian evidence in support of *no* effect) in both our dataset and in the restricted mixed-lists condition of Popp & Serra. Results were similar when only treating *studied* same-category commission errors as correct (see Supplementary online material Section 1E for detailed results for these analyses). Although these results seem to suggest an interference account for the reverse animacy effect, even with the liberal scoring criterion mean cued-recall proportion correct was not better for animals than objects. That is consistent with Popp and Serra's argument that interference and category size effects do not fully explain differential animacy effects in free versus cued recall.

### Category awareness

To further explore potential mechanisms underlying the animacy/reverse animacy effects, we coded and analyzed qualitative data from the post-study survey (Table 1). The final coding scheme can be found in Supplementary online material (Section 1F.1). Although most participants reported awareness of animal and object categories (85%, Supplementary online material Section 1F.2), casual inspection of participants' open-ended descriptions of the materials suggested a difference in the *granularity* or *specificity* of category awareness (e.g., whether participants tended to mention animals as a general or superordinate category, but used more specific subordinate subcategories for objects). We therefore coded for specificity, separately for animals and objects (see Supplementary online material Section 1F.1 for details on the coding scheme). The most striking findings here were: (a) 76% of participants reported *only* a superordinate category for animals while only 42% of participants reported *only* a superordinate category for objects (comparing these proportions, $\chi^2(1) = 23.64$, $p < 0.001$), and (b) No participants reported *only* a subordinate category for animals, compared to 21% of participants who did for objects (comparing these proportions, $\chi^2(1) = 21.26$, $p < 0.001$). Proportions of responses in the other categories can be found in Supplementary online material (Section 1F.2).

Although the magnitude of the reverse animacy effect in participants who indicated an awareness of a general object category versus participants who did not did not significantly differ (see Supplementary online material Section 1F.3), these data at least suggest that animal and object categories were perceived differently.

## Discussion

On average, our participants obtained somewhat better scores than the Popp and Serra (2016) subjects, but our primary analyses yielded results that closely paralleled theirs. Free recall was better for animal names than object names whereas cued recall was better for object-object pairs than for animal–animal pairs. These results suggest that, as proposed by Popp and Serra (2016) and Kazanas et al. (2020), the relationship between animacy and memory performance is moderated by some unknown factor(s) related to the type of memory task.

On a cued recall test, nontarget words may sometimes come to mind in response to recall cues (based on semantic and/or implicit memory). Popp and Serra (2016) raised the possibility that nontarget animal names might come to mind in response to animal cue words more often than nontarget object names come to mind in response to object cue words. That might occur if the animal category was more salient and/or narrower than the object category. If so, then nontarget animal names might interfere with retrieval of target animal names more often than nontarget object names interfere with retrieval of target object names. To assess the role of such interference, Popp and Serra re-analyzed their data with a liberal scoring criterion in which any within-category response was treated as correct. They reported an analysis of cued recall accuracy as indexed by this liberal criterion among subjects for whom category was blocked (a condition we did not include in our experiment) as well as subjects for whom categories were mixed. In that analysis, the reverse animacy effect was significant, suggesting that interference alone could not account for the reverse animacy effect on cued recall. But in our larger sample, the reverse animacy effect was not significant with this liberal scoring criterion, and Bayesian analysis provided modest support for the null hypothesis. Also, although Popp and Serra reported no significant List (themed vs. mixed) × Animacy (animal vs. object) interaction on cued-recall proportion correct with liberal scoring, the reverse animacy effect was directionally larger in their themed list than in their mixed list condition. We found that an analysis restricted to the latter yielded the same outcome as our data: Evidence for the absence of an animacy effect on cued recall accuracy with liberal scoring in the mixed-lists condition.

Our post-experiment questions regarding subjects' perceptions yielded some interesting findings. For one thing, when asked to describe the words our participants much more often mentioned a general animal category than a general object category. Also, few of our participants mentioned only subordinate categories of animals (e.g., "mammals") whereas more of them mentioned only subordinate object categories (e.g., "tools") without reference to a superordinate "object" category. These results provide further support for the idea that animals (generally or the ones used in the current stimulus set) represent a more cohesive and singular category than objects. As mentioned previously, it is possible that a tighter category structure for animals benefited free recall but hampered cued recall, while reduced awareness of a general objects category may have hampered free recall but benefited cued recall (due to reduced interference). We investigated this possibility–a category-structure-based explanation for the reverse animacy effect–in Experiment 2.

## Experiment 2

The liberal scoring data from Popp and Serra (2016) and our study suggest that the "reverse animacy effect" might better be described as a "reduced animacy due to interference effect." Additionally, our exploratory analyses of perceived category specificity suggest that the animal and object categories in the current stimulus set have different semantic structures. These findings also raise the possibility that both the animacy effect in free recall and the reverse animacy effect in cued recall may have more to do with differences in the semantic structure of the two categories (either in general or in the particular items selected by Popp and Serra and by Kazanas et al., 2020) than with the evolutionary significance of animal names. We do not think these researchers deliberately crafted lists with different similarity structures, but neither did they (or we) explicitly measure or control for category size/structure. Experiments directly controlling for or manipulating category semantic structure are necessary to adjudicate between

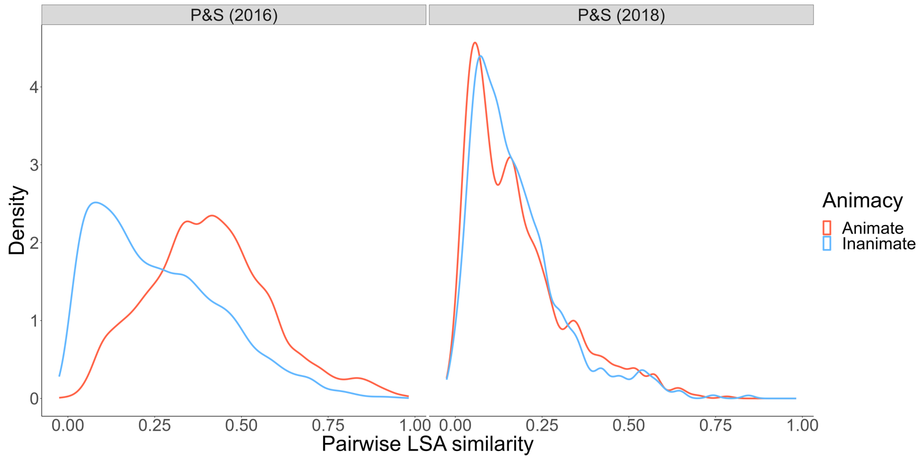


FIGURE 2
Pairwise semantic similarity from Popp and Serra (2016, 2018) wordsets. Pairwise semantic similarity indexed via LSA cosine similarity, calculated using the LSAfun R package and the English LSA 100k space with 300 dimensions (Günther et al., 2015).

semantic-structure-based explanations for animacy effects (e.g., the *overlapping semantic features hypothesis*; McRae et al., 1997; Xiao et al., 2016) and survival-processing-based accounts (e.g., the *animate monitoring hypothesis*; New et al., 2007).

One way to measure the semantic structure of wordsets is Latent Semantic Analysis (LSA; Landauer and Dumais, 1997). LSA is a method for computing distributional semantics that involves computing the co-occurrence of words across large sets of documents. Words that co-occur often in the same contexts across multiple documents are assumed to be highly similar (Landauer et al., 1998). Validating this assumption are experiments showing correlations between LSA-based measures of similarity and behavioral results (e.g., priming, sentence completion times; Günther et al., 2016).

Using word representations derived with Latent Semantic Analysis as applied to a large document corpus (Günther et al., 2015), we examined wordsets that have been used to test the reverse animacy effect (Popp and Serra, 2016; DeYoung and Serra, 2021). Specifically, we simulated 10,000 animal- and object-word pairs from each wordset and computed pairwise semantic similarities. The distributions of word-to-word similarity are shown in Figure 2.

As shown, the within-category similarities between animate words in the Popp and Serra (2016) animate versus inanimate wordsets (where a reverse animacy effect was observed) is striking. Animate cues and targets were more similar to one another than inanimate cues and targets. Conversely, in the Popp and Serra (2018) wordset (where a reverse animacy effect was *not* observed; DeYoung and Serra, 2021), within-pair similarity was generally lower and, more critically, roughly equal for animates and inanimates. These results, along with our interference analyses and qualitative data on people's category awareness data from Experiment 1, point increasingly toward a semantic structure-based explanation for the reverse animacy effect. Such an explanation is also consistent with the *overlapping semantic feature* hypothesis (McRae et al., 1997) that suggests animals may be more memorable than objects because animals share more overlapping features (e.g., fur, four legs, teeth) relative to objects, which tend to have wider-ranging features (e.g., features of different musical instruments such as trumpets and guitars have less featural overlap). This hypothesis is supported by studies showing benefits of greater feature and neural global pattern overlap for subsequent memory (Ilic et al., 2013; Xiao et al., 2016). Overlapping semantic features can explain animacy and reverse animacy effects by drawing on the *spreading-activation* theory of semantic memory (Collins and Loftus, 1975) that posits concepts activate other semantically-related concepts in memory proportional to the degree of relatedness. If animal words share more overlapping features, the semantic network will contain stronger links (and vice versa for object words) that benefit free recall of animals due to increased activation of related studied animal targets. However, stronger links might harm performance in cued recall because a given animal cue will activate more related non-target concepts. This explanation is consistent with the results of Popp and Serra (2016) and our Experiment 1, in which there was a clear performance cost in cued recall of animals relative to free recall of animals (whereas free recall and cued recall performance for objects was quite similar).

Of course, semantic similarity lay at the center of these theories and explanations. Thus, we set out to directly test these ideas by comparing animacy and reverse animacy effects in animal versus object free and cued recall with new wordsets that we designed to control pairwise similarity relationships. Specifically, we created one “animals-more-similar” wordset in which listwise/pairwise within-category similarity was higher, on average, for animals than objects (like the Popp and Serra, 2016 wordset), and one “equal” wordset in which within-category similarity was equal, on average, for animals and objects (like the Popp and Serra, 2018 wordset). We initially set out to make an “objects-more-similar” wordset but, perhaps tellingly, were unable to do so. We return to this point in the Discussion.

To create our wordsets, we used a random, iterative word-sampling process to choose words to meet the similarity conditions we wanted to satisfy but that also kept other salient word characteristics as similar as possible across categories (i.e.,

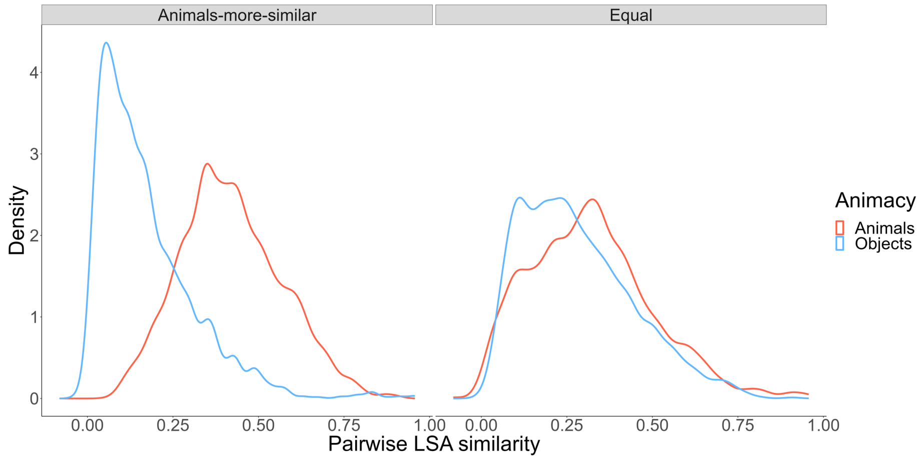


FIGURE 3
Pairwise semantic similarity for simulated pairs from our experimental conditions.

word frequency, age of acquisition, context diversity, imageability; Madan et al., 2010; Madan, 2021).[8] Ultimately, we generated two wordsets. The "animals-more-similar" wordset included 48 animal words and 48 object words, with animal-to-animal semantic similarity higher than object-to-object similarity. The "equal" wordset included 48 animal words and 48 objects words, with animal-to-animal and object-to-object semantic similarity roughly equated. The wordsets were mutually exclusive except for 17 animal words and 3 object words that appeared in both wordsets. The similarity profiles in the two wordsets are shown in Figure 3.

Using these wordsets with the Popp and Serra (2016) paradigm, we expect to replicate Popp and Serra (2016) in the animals-more-similar wordset such that participants tested on the "animals-more-similar" wordset will show an animacy effect in free recall but a reverse animacy effect in cued recall (Hypothesis 1). Second, we expect to eliminate the reverse animacy effect in the equal wordset such that we will observe a within-category similarity × animacy interaction for cued recall (Hypothesis 2). Finally, we had no specific predictions about how the "equal" wordset might affect the free recall animacy effect. But one might expect to observe that effect if there is something "special" about the animate words (above and beyond within-category similarity) that eases free recall.

## Methods

We tested these hypotheses in four different samples: two university undergraduate samples and two online samples via Prolific.co. The plans and hypotheses for the latter three samples were preregistered on the Open Science Framework with (a) all study materials, (b) the final version of the program used to collect the data, (c) details of the procedure, and (d) the specifications for the planned statistical analyses[9] (see https://osf.io/t7qfa). The data and scripts used to process/analyze the data can be accessed at https://osf.io/k4emy/.

## Design

The experiment took the form of a 2 (animacy: animals, objects) × 2 (memory type: free recall, cued recall) × 2 (within-category similarity: animals-more-similar, equal) mixed design. The first two factors were manipulated within-subjects and the third between-subjects. The primary dependent variable was, once again, recall accuracy.

### Sample size planning

We conducted power simulations using cell mean estimates obtained in Experiment 1 as well as two hypothetical effects ("small" and "large") of our similarity manipulation for Hypothesis 2. These simulations suggested that we would need $150 \leq N \leq 500$ to detect effects of interest (for the "large" and "small" effects respectively; see the preregistration for more details). For our first sample (undergraduate), we set a target $N = 150$. To maximize the efficiency of sampling for the subsequent samples, we chose to adopt a sequential Bayesian approach (Schönbrodt et al., 2017) whereby we set a minimum

---

[8] See Supplementary online material X for an exhaustive description of our word-selection process.

[9] We drafted but did not preregister a plan for the first undergraduate sample because we discovered after data collection but before analysis that the program was not assigning equal numbers of animals/objects to each study list as we had intended (see footnote 2). We include the experiment here because both Popp and Serra (2016) and our replication also had unequal category-list assignment, and our core hypotheses described above did not change after the discovery of the error. The draft version of the preregistration can be viewed here: https://docs.google.com/document/d/18oKBq7jxl-4Pz7zZ6iQBIVeflmaU83QU7oXk0UzGi1s/edit?usp=sharing. For consistency across experiments, we analyzed all data according to the preregistration for the latter three samples.

testing $N = 100$ and tested our core hypotheses via Bayes Factors at this threshold and at each additional $n = 10$ subsequent participants, up to a maximum total $N = 150$. At each sequential testing threshold, we tested three effects of interest via model-comparison Bayes Factors: the free recall animacy effect in the "animals-more-similar" condition, the cued recall reverse animacy effect in the "animals-more-similar" condition, and the cue recall reverse animacy effect in the "equal" condition. Effects were tested via comparison of a model with the effect of interest to a model without the effect of interest. If evidence in favor or against all three effects exceeded a Bayes Factor of 3, we would terminate data collection, otherwise continue to the next threshold.

## Sample

We collected four different samples for Experiment 2: two undergraduate samples, one online (Prolific.co) sample including all people who reported English fluency, and a fourth sample from online (Prolific.co) restricted to people who reported English fluency and English as a first language (post exclusion $Ns = 150$, 143, 101, and 150 respectively). Additional information about the four samples including information about exclusions based on preregistered exclusion criteria (increased from Experiment 1), condition assignments, and demographics can be found in Supplementary online material 2X.

Our objectives in collecting these four samples were to (a) test the generality of effects across different populations and (b) self-replicate our results to ensure internal as well as external convergence[10].

## Materials

The computer program was unchanged from Experiment 1, except for the following: First, we replaced the Popp and Serra wordset with our new wordsets ("animals-more-similar" and "equal"), to which participants were randomly assigned. Second, we increased the number of words/word-pairs per list to 16 to allow for equal numbers of animal/object words/pairs. Third, For samples 2–4, we ensured that each study list contained 8 animal and 8 object words/pairs. Fourth, we added a line to the cued recall instructions to discourage fast skips: "It is OK to leave items blank if you cannot recall the target word, but please make an effort to recall each target, do not just quickly skip through." Lastly, we added debriefing/exclusion questions to evaluate cheating ("Did you take notes?"), words understood (0, 25, 50, 75, 100%), and general experiences (general comments, technical difficulties). The LiveCode (https://livecode.com/) program for the experiment is accessible at https://osf.io/khsj9.

## Procedure

Aside from the changes noted above, the procedure was the same as in Experiment 1. In addition to recruiting participants via the University research pool, we posted our study online at Prolific.co, where eligible online participants who took part in the +/− 30-min experiment received $6.00 USD.

# Results

## Analytic strategy

As our sampling plan depended on sequential Bayesian tests, our primary analyses were Bayesian linear mixed-effects models, though we supplemented these analyses with NHST ANOVAs and *t*-tests to facilitate comparison of our results to the original Popp and Serra (2016) work.

Our hypothesized interactions were (a) an animacy by memory type interaction in the "animals-more-similar" condition (i.e., free recall animacy effect paired with a cued recall reverse animacy effect) and (b) a within-category similarity by animacy interaction for cued recall (i.e., reduced/eliminated reverse animacy effect in the "equal" condition). At the granularity of individual cells/conditions, our hypothesized pairwise comparisons of interest, and the basis for our sequential sampling, were (a) an animacy effect in the "animals-more-similar" condition, (b) a cued recall reverse animacy effect in the "animals-more-similar" condition, and (c) *no* cued recall reverse animacy effect in the "equal" condition. To evaluate each of these effects, we conducted Bayes Factor model comparisons of a model with that effect to a model without that effect.

For Experiment 2, accuracy measures were manually coded. That is, verbatim correct answers were counted as correct in addition to misspelled responses that were manually coded (See Supplementary online material Section 1B for details regarding this coding procedure and Supplementary online material Section 2E for manual coding statistics for each sample). Overall, the number of manually corrected responses was small (no more than 5% of total responses).

## Primary analysis

To test our replication of Popp and Serra (2016) in the animals-more-similar wordset, we conducted a 2 (animacy: animals, objects) × 2 (memory type: free recall, cued recall) ANOVAs/Bayesian linear models in the "animals-more-similar" condition. In all samples, the hypothesized interaction was observed (see Table 2 for more details). To test our second hypothesis that we would eliminate the reverse animacy effect in the "equal" wordset, we conducted a 2 (animacy: animals, objects) × 2 (within-category similarity: animals-more-similar, equal) ANOVAs/Bayesian for cued recall only. In all samples (and combining across samples), significant interactions and "extreme" Bayesian evidence (Jeffreys, 1961; $BF > 100$) supported our first hypothesis. For our second hypothesis, in three out of the four samples the hypothesized interaction was not observed, with moderate

---

[10] Although we preregistered a general experiment plan, we did not initially plan to collect four samples. Rather, we conducted the experiments one at a time, making decisions as to the details as we went along.

TABLE 2 Experiment 2 interaction hypothesis tests.

| | H1: Animacy and reverse animacy effects in the “animals-more-similar” condition | | H2: Reverse animacy effect in the “animals-more-similar” but not “equal” condition | |
|---|---|---|---|---|
| Sample | Interaction BF10 | Interaction NHST | Interaction BF10 | Interaction NHST |
| Undergraduate 1 | **$2.89 * 10^{8}$** | $F(1, 67) = 24.04$, **$p < 0.001$**, $\eta^2_p = 0.26$ | **0.28** | $F(1, 148) = 0.46$, $p = 0.50$, $\eta^2_p < 0.001$ |
| Undergraduate 2 | **$1.48 * 10^{10}$** | $F(1, 71) = 52.17$, **$p < 0.001$**, $\eta^2_p = 0.42$ | **0.29** | $F(1, 141) = 0.68$, $p = 0.41$, $\eta^2_p = 0.005$ |
| Prolific | **$9.90 * 10^{6}$** | $F(1, 54) = 12.02$, **$p = 0.001$**, $\eta^2_p = 0.18$ | **6.03** | $F(1, 99) = 3.34$, $p = 0.07$, $\eta^2_p = 0.03$ |
| Prolific (EFL) | **$5.94 * 10^{7}$** | $F(1, 78) = 14.55$, **$p < 0.001$**, $\eta^2_p = 0.16$ | 0.42 | $F(1, 148) = 0.007$, $p = 0.93$, $\eta^2_p < 0.001$ |
| Combined | **$2.11 * 10^{23}$** | $F(1, 273) = 91.32$, **$p < 0.001$**, $\eta^2_p = 0.25$ | **0.06** | $F(1, 542) = 0.35$, $p = 0.55$, $\eta^2_p < 0.001$ |

The H1 interaction refers specifically to the 2 (animacy: animals, objects) × 2 (memory type: free recall, cued recall) within-subjects interaction in the “animals-more-similar” condition, with a significant interaction providing support for our hypothesis. The H2 interaction refers specifically to the 2 (animacy: animals, objects) × 2 (within-category similarity: animals-more-similar, equal) within-between interaction for cued recall only, with a significant interaction providing support for our hypothesis. BF10 values are Bayes Factors indicating the ratio of support for the interaction model over a null model with no interaction (i.e., values above 1 provide support for an interaction, values below 1 provide support against an interaction). Bolded values indicate BFs or p-values that exceeded our preregistered critical values.

Bayesian evidence ($BF10 < 0.33$) against an interaction in two of those samples (Undergraduate 1 & 2), and anecdotal evidence ($1 > BF10 > 0.33$) against an interaction in the remaining sample (Prolific EFL). When combining all samples, the interaction was non-significant and Bayesian evidence against was strong ($BF > 10$). Table 2 lists statistics for these analyses in all samples (see Supplementary online material 2C for individual experiment figures depicting the raw data similar to Figure 1):

We observed consistent evidence for Hypothesis 1 (Animacy and reverse animacy effects in the “animals-more-similar” condition), replicating the basic Popp and Serra results pattern in our “animals-more-similar” condition. In three out of the four samples (Undergraduate 1 & 2, Prolific EFL), we found evidence *against* Hypothesis 2 (Reverse animacy effect in the “animals-more-similar” but not “equal” condition), suggesting that in these samples, semantic similarity did not modulate the reverse animacy effect. The exception was in our first Prolific sample, where the reverse animacy effect was eliminated in the “equal” semantic similarity condition, with moderate Bayesian evidence for an interaction. As these analyses were conducted using manually coded accuracy, we also conducted versions of these analyses using verbatim/automatic accuracy, with similar results (see Supplementary online material Section 2F). In Table 3, we show the Bayes Factors, *t* tests, and Cohen’s $d_z$ values for all pairwise animal-object comparisons in Figure 4, we visualize the pairwise effects in a forest plot.

Importantly, in all samples except the third, the cued recall reverse animacy effect was significant and supported by moderate-or-greater Bayesian evidence in both similarity conditions. The free recall animacy effect was less consistent (significant in only 4/8 cells), but was always in the expected direction and significant and supported by extreme Bayesian evidence when combined across samples. Similar results were obtained for verbatim/automatic accuracy (see Supplementary online material Section 2F).

## Discussion

Overall, our results do not support a “semantic similarity” explanation for the reverse animacy effect. The majority-English as a Second Language sample in which we did find support for such an explanation might suggest that first-language status interacts with similarity in some way. There is some evidence that the structure of semantic networks differ between monolinguals and bilinguals (Bilson et al., 2015), but we are hesitant to draw any conclusions about language status on the basis of only one sample. Why then did semantic similarity not influence animacy and reverse animacy effects?

One possibility is that our manipulation was not strong enough. If the reader recalls that we created one wordset in which animal cue-target pairs had (on average) higher within-pair semantic similarity than did cue-target object pairs, and one wordset in which animal and object pairs had (on average) equivalent within-pair semantic similarity. This manipulation was successful in the sense that animal and object word similarity distributions appeared as intended in all samples (see Supplementary online material Section 2D). Both our primary basis for estimating word similarity (LSA) and an additional/newer similarity measure that we computed after the fact (Global Vectors for Word Representation (GloVe)[11]; Pennington et al., 2014) were generally related to memory accuracy (i.e., higher similarity = higher cued recall accuracy, see Supplementary online material Section 2D). However, even in our “equal” similarity condition, semantic similarity was slightly higher for animal–animal than object-object word pairs. It could be that a more forceful difference (e.g., higher within-pair similarity for objects vs. animals) is required to observe statistically corroborated differences in the reverse animacy effect.

However, we think it more likely that semantic similarity is not the primary mechanism behind the reverse animacy effect. Initially, it seemed that patterns in semantic similarity distributions coincided with the presence (or not) of a reverse animacy effect, but that is less clear with our results (see Figure 5).

Crucially, with high distributional overlap, we observed a reverse animacy effect but Popp and Serra (2018) did not (in fact, they observed a *reverse* reverse animacy effect – higher recall of animal than object targets). As per the Popp and Serra (2018) similarity distributions, it could be that similarity has to

---

[11] Using the common crawl space with 840B tokens, 2.2M vocabulary, cased, and 300d vectors.

**TABLE 3** Experiment 2 pairwise animal-object comparison tests.

| | "Animals-more-similar" condition | | | | "Equal" condition | | | |
|---|---|---|---|---|---|---|---|---|
| | FR | | CR | | FR | | CR | |
| Sample | BF10 | *t*-test | BF10 | *t*-test | BF10 | *t*-test | BF10 | *t*-test |
| Undergraduate 1 | 1.12 | $t(67) = 1.78$, $p = 0.08$ / 0.10, $d_z = 0.22$ [−0.02, 0.46] | **2.61 * 10$^4$** | $t(67) = 5.18$, **$p < 0.001$**, $d_z = -0.63$ [−0.89, −0.37] | **0.01** | $t(81) = 0.83$, $p = 0.41$ / 0.49, $d_z = 0.09$ [−0.13, 0.31] | **8.31 * 10$^7$** | $t(81) = 5.14$, **$p < 0.001$**, $d_z = -0.57$ [−0.80, −0.33] |
| Undergraduate 2 | **6.34** | $t(71) = 3.37$, **$p = 0.001$ / 0.003**, $d_z = 0.40$ [0.16, 0.64] | **9.08 * 10$^5$** | $t(71) = 6.21$, **$p < 0.001$**, $d_z = -0.73$ [−0.99, −0.47] | **11.09** | $t(70) = 2.63$, **$p = 0.01$**, $d_z = 0.31$ [0.07, 0.55] | **6.48 * 10$^5$** | $t(70) = 4.60$, **$p < 0.001$**, $d_z = -0.55$ [−0.79, −0.29] |
| Prolific | **10.24** | $t(54) = 2.00$, $p = 0.05$, $d_z = 0.27$ [0.001, 0.54] | **9.93** | $t(54) = 2.91$, **$p = 0.01$**, $d_z = -0.39$ [−0.67, −0.12] | **10.19** | $t(45) = 2.46$, **$p = 0.02$**, $d_z = 0.36$ [0.06, 0.66] | **0.02** | $t(45) = 0.15$, $p = 0.88$ / 1, $d_z = -0.02$ [−0.31, 0.27] |
| Prolific (EFL) | 2.81 | $t(78) = 2.35$, **$p = 0.02$ / 0.03**, $d_z = 0.26$ [0.04, 0.49] | **4.96** | $t(78) = 2.47$, **$p = 0.02$**, $d_z = -0.25$ [−0.50, −0.05] | **23.18** | $t(70) = 2.75$, **$p = 0.007$**, $d_z = 0.36$ [0.09, 0.56] | **4.65** | $t(70) = 2.90$, **$p = 0.005$ / 0.008**, $d_z = -0.34$ [−0.58, −0.10] |
| Combined | **6,705.42** | $t(273) = 4.70$, **$p < 0.001$**, $d_z = 0.28$ [0.16, 0.40] | **4.31 * 10$^{11}$** | $t(273) = 8.10$, **$p < 0.001$**, $d_z = -0.49$ [−0.61, −0.36] | **2,238.76** | $t(269) = 4.32$, **$p < 0.001$**, $d_z = 0.26$ [0.14, 0.38] | **1.16 * 10$^{11}$** | $t(269) = 6.68$, **$p < 0.001$**, $d_z = -0.41$ [−0.53, −0.28] |

FR = Free recall, CR = Cued recall. Each cell refers to a comparison of memory performance for animals and objects (positive Cohen's $d_z$ effect sizes = better memory performance for animals than objects). BF10 values are Bayes Factors indicating the ratio of support for a model with an animal-object difference relative to a model with no difference (i.e., values above 1 provide support for a difference, values below 1 provide support against a difference). Bolded values indicate BFs or p-values that exceeded our prespecified critical values.

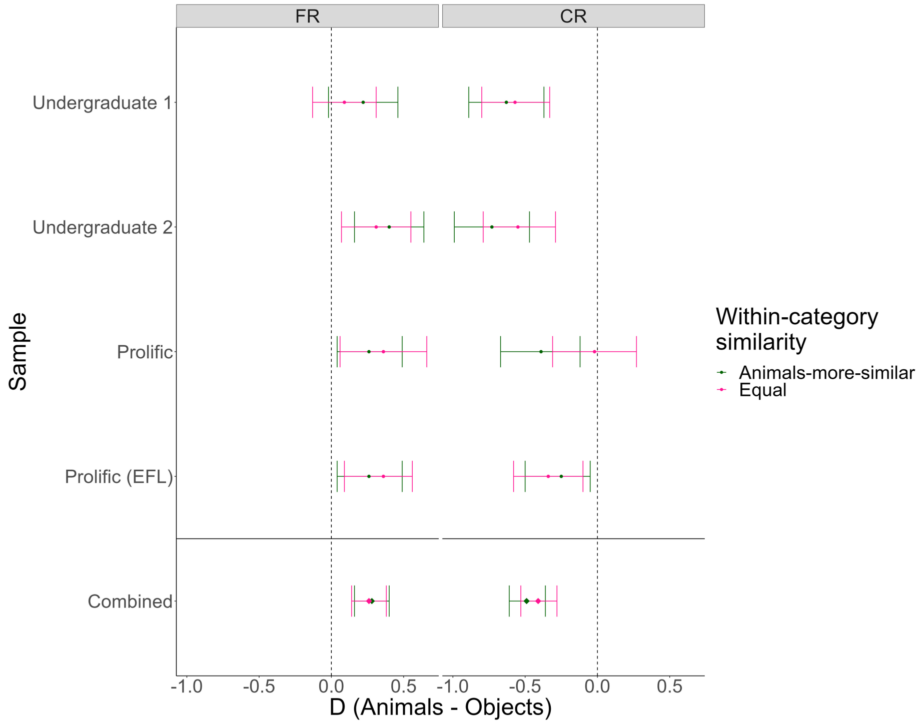


FIGURE 4
Forest plot of animacy effects by condition and test type in Experiment 2. FR=Free recall, CR=Cued recall. Error bars=95% CIs on the within-subjects Cohen's $d_z$s. Points to the right of the dotted lines represent animacy effects, points to the left of the lines represent reverse animacy effects.

be equated *and/or* low overall. In a recent independent study published during our data collection, Serra and DeYoung (2022) also explored potential within-pair factors that might account for the reverse animacy effect. They used a smaller, fixed set of animal and object word pairs that were manipulated to involve either two typical exemplars from the same category (e.g.,

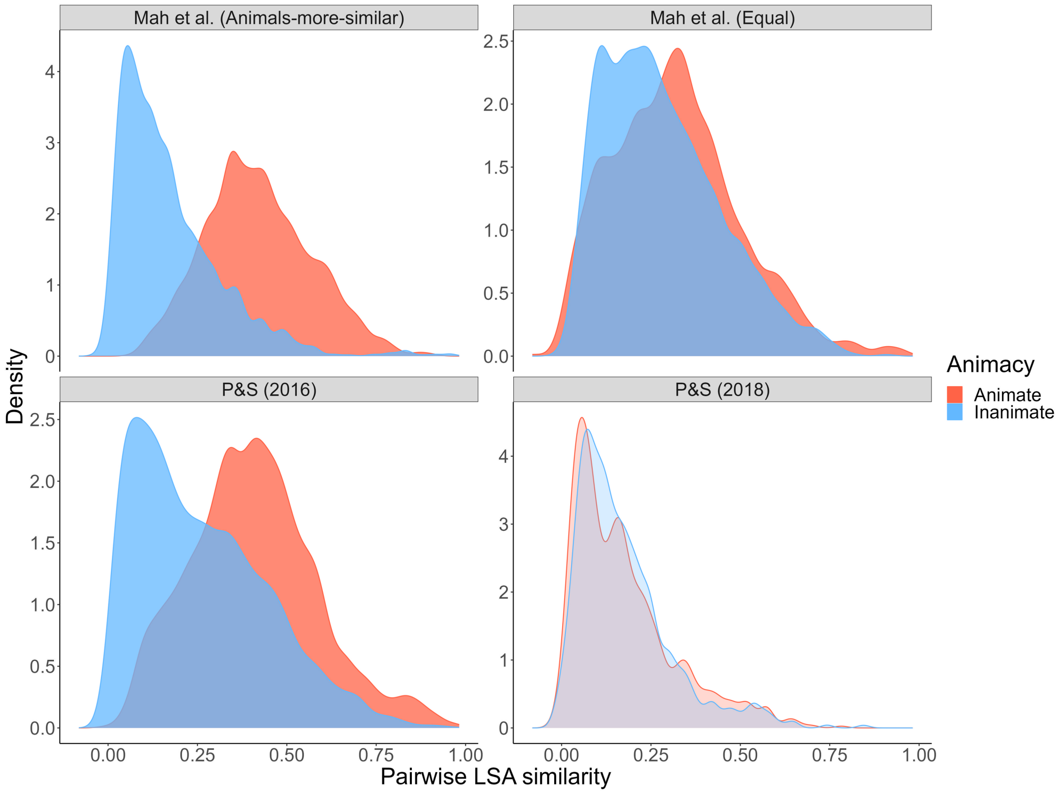


FIGURE 5
Simulated semantic similarity distributions: Current experiment and Popp and Serra (2016, 2018). Distributions represent pairwise Latent Semantic Analysis (LSA) similarity values for 10,000 randomly generated animal/object pairs from each wordset. High-opacity denotes experiments/samples in which a reverse animacy effect was observed.

SALMON–TROUT; FORK–SPOON), one typical and one atypical exemplar from the same category (e.g., SALMON–SNAPPER; FORK–STOVE), or two unrelated exemplars (e.g., SALMON–RECEIVER; FORK–NAIL). With these pairs types respectively, they observed the standard cued recall reverse animacy effect (i.e., higher recall of inanimate than animate typical-typical pairs), a cued recall animacy effect (i.e., higher recall of animate than inanimate typical-atypical pairs), and no difference by animacy (with unrelated pairs). As Figure 6 shows, a *post hoc* analysis of their wordsets similarly failed to reveal a consistent relationship (at least a clear one that we can see) between LSA semantic similarity and the presence/absence of a reverse animacy effect:

On the basis of these results, one might speculate that (a) higher animal than object similarity produces an animacy effect, (b) equal-and-moderate similarity produces a reverse animacy effect, and (c) equal-and-low similarity produces no animacy difference. However, these patterns are not consistent when one considers all the wordsets we examined.[12] In their study, Serra and DeYoung (2022) conducted additional exploratory analyses using word similarity ratings grounded in the LSA and GloVe theories and found that their constructed typicality categories did not map onto the within-pair similarity measures (and analyses of the within-pair similarity measures were generally inconclusive). As a result, they argue that paired-associate animacy and reverse animacy effects are likely not due to typicality or similarity *per se*, but some other aspect of memorability imperfectly related to these factors (and perhaps more related to typicality than similarity). Having now conducted a fulsome and high power experimental examination of the issue and nevertheless failing to reach a clear and forceful conclusion, we are inclined to agree with them.

## General discussion

In two experiments (and five samples), we investigated animacy and reverse animacy effects in free recall and cued recall. Overall, we replicated the basic pattern observed by Popp and Serra (2016) – better free recall for animals than objects, and better cued recall for objects than animals. Our results (and the results of other recent studies; e.g., Kazanas et al., 2020; Serra and DeYoung, 2022) provide further evidence for the robustness of these effects. Although the cued recall reverse animacy effect is replicable, results seem to vary as a function of materials.

In an effort to investigate underlying mechanisms and determine what specific materials-based factors might account for these inconsistencies, we considered one potential moderator – semantic

---

[12] Results were similar when analyzing GloVe similarity, see Supplementary online material Section 2D. See Supplementary online material Section 2G for descriptive statistics for the similarity metrics and the presence of reverse animacy effects.

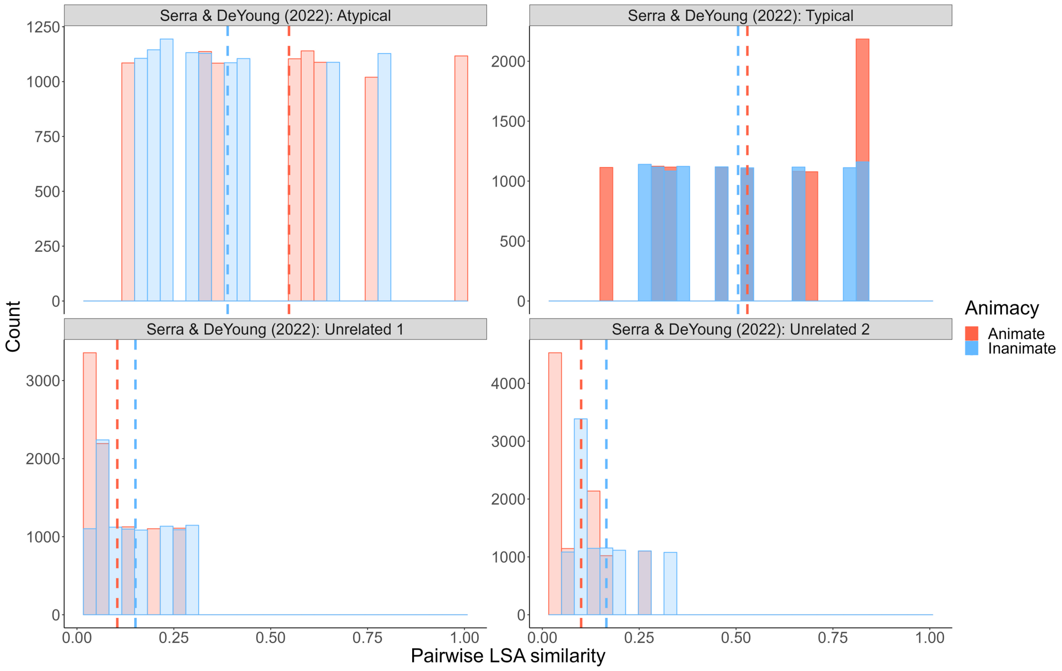


FIGURE 6
Simulated semantic similarity distributions: Serra and DeYoung (2022). Distributions represent pairwise similarity values for 10,000 randomly sampled animal/object pairs from each wordset. High-opacity denotes experiments/samples in which an animacy effect (animacy advantage) was observed.

similarity. Our key hypothesis was that equating animal and object within-category semantic similarity would eliminate or reduce the reverse animacy effect. In three out of the four samples that we collected, we either observed evidence against this hypothesis or failed to find evidence to support it. Based on these results (and examinations of other datasets), we argue that semantic similarity is not likely to be the mechanism behind the reverse animacy effect. Rather, the patterns of semantic similarity that we hypothesized to be causal (e.g., in Popp and Serra, 2016 and Popp and Serra, 2018) were perhaps confounded with the true underlying mechanism.

We are left to consider what other potential mediators or mechanisms might be at play here. As Serra and DeYoung (2022) suggest, it could be category typicality (or some other aspect affected by typicality). Another possibility is category size or specificity. In Experiment 1, and in the combined samples of Experiment 2, we coded and analyzed participants' post-experiment self-reported perceptions about the words they studied. Data from Experiment 2 suggested that our similarity manipulation did not affect participant perceptions of the *granularity* or *specificity* of the animal and object categories (e.g., participants were no more likely to report a more general or cohesive "objects" category in the "equal" similarity condition vs. the "animals-more-similar" condition, see Supplementary online material Section 2H). The vast majority of participants in Experiment 2 indicated awareness of a superordinate "animals" category (82%), while less than half of participants indicated awareness of a superordinate "objects" category (47%), with minimal differences across conditions. Thus, it is possible that participants may have been influenced more by the general categories themselves than the specific semantic relationships for any particular list (see Higham and Brooks, 1997, for data on the issue even when participants do not express awareness of category structure in memory lists).

Awareness of the categories could have primed participants' animal and object conceptual networks generally, irrespective of the similarity condition. So, although we intended for less spreading-activation-based interference (Collins and Loftus, 1975) for animals in our "equal" similarity condition (by creating a list of words less strongly linked with one another), it could be that participants became aware of the general animal category, resulting in widespread activation of animal concepts anyway. In other words, perhaps the influence of these general categories overshadowed activations of specific cues within the condition-varying semantic networks, leading to similar results in both conditions. As a final exploratory analysis in this vein, we compared the magnitude of the reverse animacy effect in two groups of Experiment 2 participants: those who indicated an awareness of superordinate "animal" and "object" categories, and those who indicated only an awareness of a superordinate "animals" category. The reverse animacy effect did not differ as a function of category awareness (see Supplementary online material Section 2Hd), but again, the vast majority of participants in our samples indicated awareness of a general animals category (i.e., we did not have enough unaware participants for a 3rd comparison group). We suggest that a "lack of awareness of a general animals category" may be more important than "awareness of a general objects category" for modulation of reverse animacy effects.

Similar effects (differential impacts of a manipulation on item versus associative memory tasks) have been observed in other paradigms. For example, Madan et al. (2012) and Bisby and Burgess (2013) found better item memory (free recall) but worse associative memory (cued recall in the former, cued recall/context memory in the latter) of emotionally arousing negative words/pairs than neutral words/pairs. Zimmerman and Kelley (2010) replicated this general pattern and also found substantial overconfidence in judgments of learning for negative relative to neutral pairs. These results suggest two possibilities – that the individually arousing words in studied pairs drew attention away from the association, or that subjects may have been overconfidence in their ability to remember emotionally arousing pairs (Madan et al., 2012). It is possible that these mechanisms might extend to animacy and reverse animacy effects.

Many studies have tested and ostensibly ruled out arousal (and proximate mechanisms such as attention) as an explanation for animacy effects (e.g., Meinhardt et al., 2018; Popp and Serra, 2018; Leding, 2019; Rawlinson and Kelley, 2021). However, metacognitive explanations for animacy and reverse animacy effects have received less attention. DeYoung and Serra (2021) did find higher judgments of learning for animate than inanimate pairs, but did not replicate the reverse animacy effect. Future studies replicating the reverse animacy effect while measuring participant judgments of learning are needed to determine the plausibility of metacognitive explanations – namely, that participants are overconfident in their ability to memorize and recall animate pairs and thus exert less effort studying them.

Like semantic similarity, specific animate and inanimate word categories vary across experiments. In our experiments and Popp and Serra (2016), the cued recall reverse animacy effect was observed (although this was not the case in DeYoung and Serra, 2021). The same wordset (consisted *only* of animals and objects) was used in these three studies. In Kazanas et al. (2020), reverse animacy effects were observed in three paired-associates tasks. Their wordset consisted of animate animals and inanimate clothing, fruits, and weapons. In Popp and Serra (2018), where a reverse animacy effect was not observed, the wordset consisted of a mix of animate humans and animals, and inanimate natural and manmade objects (DeYoung and Serra, 2021). In Serra and DeYoung (2022), the reverse animacy effect fluctuated based on the relative typicality of cues and targets. The wordset in their study consisted of a mix of animate humans described in terms of roles (e.g., doctor, quarterback) and animals versus exclusively manmade inanimate objects. Although these experiments were conducted in different samples (and in some cases using different tasks), it is suggestive that the reverse animacy effect was most consistently observed when the animate category consisted only of nonhuman animals. In line with our category size/specificity explanation, it could be that the use of more ambiguous categories (e.g., animate vs. animals) allows within-pair characteristics (like typicality; DeYoung and Serra, 2021) to modulate cued recall animacy effects. But when the general category is highly salient, it could be that category-specific effects dominate the influence of pair-level variables. We do not claim to answer these questions here, but it seems worthwhile to examine and compare animate and inanimate categories of varying levels of specificity. This idea of category specificity has been explicitly tested in only two studies. In one, VanArsdall et al. (2017) found a reverse animacy effect with constrained categories (i.e., four-footed animals, furniture), but they used a different task: English-Swahili word pairs in an English-speaking sample. In another, Gelin et al. (2017) controlled for category size and cohesion, but only tested participants on free recall. Perhaps the field needs a larger multi-lab replication effort to examine the benefit/cost of animacy in free/cued recall to settle the issue once and for all?

## Limitations

As mentioned previously, one possible limitation is that our Experiment 2 manipulation of within versus between category similarity could have been more forceful. Our "equal" similarity condition was noticeably different from the "animals-more-similar" condition in terms of semantic similarity distributions, but within-pair similarity was still slightly higher for animals relative to objects. A re-analysis of our wordset using an alternate similarity measure (GloVe; Pennington et al., 2014) revealed an even larger discrepancy between similarity distributions in favor of animals. Although our manipulation aimed to control for semantic similarity by equating average listwise and pairwise values across animacy categories, an "ideal" manipulation of category semantic similarity might have been a wordset in which within-pair similarity was higher for objects than animals. Over the course of many 1,000s of simulated wordsets of common animal and object words, we did not obtain a single wordset in which this was the case. Perhaps this speaks to the general categories themselves – the animal category might be inherently more cohesive and constrained than objects, restricting the degree to which manipulations can affect category-related outcomes. Such an explanation is consistent with the *overlapping semantic features* hypothesis (McRae et al., 1997). That is, animals may naturally share more overlapping features than is the case for objects. Popp and Serra (2018) and Serra and DeYoung (2022) were able to manipulate the reverse animacy effect using more general animate and inanimate categories. It might be that in addition to reducing the salience of the overall categories (as we have suggested above), their wordsets allowed for control of effects via a reduction of semantic feature overlap. If so, our hypothesis that there may be uninteresting but important psycholinguistic factors driving the reversed animacy effect in cued recall might be revived. However, the story is more complex than semantic similarity models like LSA and GLoVe currently capture.

## Constraints on generality

The remaining limitations relate to potential constraints on the generalizability of our findings (Simons et al., 2017). First, our observations and conclusions are specific to memory performance on the particular (quite artificial) free and cued recall tasks used in our experiments (i.e., words with certain characteristics presented singly or in random word-pairs for one at a time in random order with instructions to remember them for a subsequent test, followed by a brief retention interval and then tests of the sorts we have reported). Animacy effects in memory are often couched in real-world, evolutionarily relevant contexts (e.g., remembering predators or useful objects), so caution is advised in generalizing from these artificial tasks to more realistic remembering. Some experiments have examined the effects we report on here in more ecologically valid contexts (e.g., using survival-processing paradigms; Kazanas et al., 2020; or more complex/realistic stimuli like images; Bonin et al., 2014). Further experiments specifically examining the reverse animacy effect in more realistic contexts would help determine the generalizability of the effect, and shed light on the degree to which the effect is specific to word stimuli or word-related characteristics.

Our experiments also relied on two relatively constrained wordsets. As we have discussed, the wordsets used likely impact more than the generalizability of findings–rather, wordset characteristics likely relate to substantive mechanisms underlying the animacy and reverse animacy effects. Still, it is possible that even with another set of similar animals and objects, we might have observed different results. Although we considered and attempted to control for various other word characteristics related to memory (e.g., word frequency, age of acquisition, context diversity, imageability; Madan et al., 2010; Madan, 2021), it is impossible to perfectly match stimulus sets on all these categories (e.g., Clark, 1973). Even if it was possible, the extreme set of constraints on word selection would probably generate a list of words that resembles the natural category in some unusual and odd fashion that would render distinctiveness a going factor in understanding people's memorial performance and study strategies. Although unlikely, it is possible that a wordset differing on one or more of these characteristics (e.g., lower overall word frequency) could have led to different patterns of results.

Finally, our experiments tested English-speaking participants sourced from undergraduate and online populations, with English concrete nouns carrying various psycholinguistic characteristics. Although the animacy effect on free recall has been replicated with multilingual stimulus sets (e.g., French, German, Chinese, Portuguese; Mieth et al., 2019), to our knowledge the reverse animacy effect has only been examined in English samples, with English stimuli. Our point here is not to claim that we have a theory that predicts different patterns of animacy and reverse animacy effects as a function of varying word frequency, context diversity, participant samples, languages and so forth. Would that we did. We are merely acknowledging potential constraints on the generality of our findings (Simons et al., 2017).

## Conclusion

The results of our experiments (and other experiments from Serra and colleagues) suggest that the reverse animacy effect in cued recall is a robust, replicable effect. Moreover, it does not appear that semantic similarity explains the effect. That is not to say that the effect cannot be ascribed to non-animacy pair-level factors. In fact, recent work from Serra and DeYoung (2022) suggests that within-pair factors such as typicality can explain the reverse animacy effect. However, like Serra and DeYoung, we do not think that these results rule out "adaptive memory" explanations for animacy and reverse animacy effects. Pair-level factors such as typicality or category-level factors such as category specificity might explain these effects, but could in turn be related to evolutionary factors.

Although we have not identified the specific mechanisms underlying the reverse animacy effect, our experiments bring the field closer. We provide evidence against an initially attractive candidate–word similarity–consistent with Serra and DeYoung's suggestion of some mechanism related to typicality. Additionally, our exploratory analyses of category awareness and an informal review of wordsets in reverse animacy experiments point to a potential influence of the size, specificity, or granularity of the animate and inanimate categories used. We have also added to a growing body of openly available free and cued recall animacy data (e.g., see https://osf.io/7cx2r/) that we hope will be of use to other researchers examining these effects (e.g., allowing analyses of different word/category characteristics). Finally, we have helped to establish the replicability of the reverse animacy effect, but it remains unclear whether and how the effect generalizes to more realistic memory tasks.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/hc2af/.

## Ethics statement

The studies involving human participants were reviewed and approved by University of Victoria Human Research Ethics. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## Author contributions

EM, AC, NT, KG, and DL planned and conducted Experiment 1 and drafted the original writeup. EM, KG, RJ, and DL planned Experiment 2, and EM, KG, and DL conducted Experiment 2 (as part of KG's Honours Thesis), and EM and KG drafted the writeup for Experiment 2. EM drafted the final version of the combined manuscript, and all authors revised this version before final submission.

## Funding

This work was supported by an NSERC Discovery Grant (#RGPIN-2016-03944) awarded to DL. Funds from this grant were used to pay online (Prolific.co) participants.

## Acknowledgments

The authors thank Michael Serra for providing the original experiment program and data, and for his helpful suggestions and feedback on the replication study design and on a draft of this manuscript, and Michael Masson and Adam Krawitz for their helpful suggestions on the manuscript. The authors would also thank Lorne Campbell for providing an inspiring example of a graduate seminar undertaking a direct replication (https://osf.io/9m4ed/)–the context within which the current work began.

## 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.

## Publisher's note

## Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1146200/full#supplementary-material

## References

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., et al. (2020). A consensus-based transparency checklist. *Nat. Hum. Behav.* 4, 4–6. doi: 10.1038/s41562-019-0772-6

Anderson, S. F., Kelley, K., and Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: a method adjusting sample effect sizes for publication bias and uncertainty. *Psychol. Sci.* 28, 1547–1562. doi: 10.1177/0956797617723724

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., and Hills, T. T. (2015). Semantic facilitation in bilingual first language acquisition. *Cognition* 140, 122–134. doi: 10.1016/j.cognition.2015.03.013

Bisby, J. A., and Burgess, N. (2013). Negative affect impairs associative memory but not item memory. *Learn. Mem.* 21, 21–27. doi: 10.1101/lm.032409.113

Bonin, P., Gelin, M., and Bugaiska, A. (2014). Animates are better remembered than inanimates: further evidence from word and picture stimuli. *Mem. Cogn.* 42, 370–382. doi: 10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Laroche, B., Meot, A., and Bugaiska, A. (2015). The "how" of animacy effects in episodic memory. *Exp. Psychol.* 62, 371–384. doi: 10.1027/1618-3169/a000308

Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *J. Verbal Learn. Verbal Behav.* 12, 335–359. doi: 10.1016/S0022-5371(73)80014-3

Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428. doi: 10.1037/0033-295X.82.6.407

DeYoung, C. M., and Serra, M. J. (2021). Judgments of learning reflect the Animacy advantage for memory, but not beliefs about the effect. *Metacogn. Learn.* 16, 711–747. doi: 10.1007/s11409-021-09264-w

Gelin, M., Bugaiska, A., Méot, A., and Bonin, P. (2017). Are animacy effects in episodic memory independent of encoding instructions? *Memory* 25, 2–18. doi: 10.1080/09658211.2015.1117643

Günther, F., Dudschig, C., and Kaup, B. (2015). LSAfun - an R package for computations based on latent semantic analysis. *Behav. Res. Methods* 47, 930–944. doi: 10.3758/s13428-014-0529-0

Günther, F., Dudschig, C., and Kaup, B. (2016). Latent semantic analysis cosines as a cognitive similarity measure: evidence from priming studies. *Q. J. Exp. Psychol.* 69, 626–653. doi: 10.1080/17470218.2015.1038280

Higham, P. A., and Brooks, L. R. (1997). Learning the experimenter's design: tacit sensitivity to the structure of memory lists. *Q. J. Exp. Psychol.* 50A, 199–215.

Ilic, O., Kovic, V., and Styles, S. J. (2013). In the absence of animacy: superordinate category structure affects subordinate label verification. *PLoS One* 8:e83282. doi: 10.1371/journal.pone.0083282

Jeffreys, H. (1961). *Theory of probability* (*3rd*). Oxford, UK: Oxford University Press.

Kazanas, S. A., Altarriba, J., and O'Brien, E. G. (2020). Paired-associate learning, animacy, and imageability effects in the survival advantage. *Mem. Cogn.* 48, 244–255. doi: 10.3758/s13421-019-01007-2

Landauer, T. K., and Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* 104, 211–240. doi: 10.1037/0033-295X.104.2.211

Landauer, T. K., Foltz, P. W., and Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Process.* 25, 259–284. doi: 10.1080/01638539809545028

Leding, J. K. (2019). Adaptive memory: Animacy, threat, and attention in free recall. *Mem. Cogn.* 47, 383–394. doi: 10.3758/s13421-018-0873-x

Loftus, G. R., and Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1, 476–490. doi: 10.3758/BF03210951

Madan, C. R. (2021). Exploring word memorability: how well do different word properties explain item free-recall probability? *Psychon. Bull. Rev.* 28, 583–595. doi: 10.3758/s13423-020-01820-w

Madan, C. R., Caplan, J. B., Lau, C. S. M., and Fujiwara, E. (2012). Emotional arousal does not enhance association-memory. *J. Mem. Lang.* 66, 695–716. doi: 10.1016/j.jml.2012.04.001

Madan, C. R., Glaholt, M. G., and Caplan, J. B. (2010). The influence of item properties on association-memory. *J. Mem. Lang.* 63, 46–63. doi: 10.1016/j.jml.2010.03.001

McRae, K., de Sa, V. R., and Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *J. Exp. Psychol. Gen.* 126, 99–130. doi: 10.1037//0096-3445.126.2.99

Meinhardt, M. J., Bell, R., Buchner, A., and Röer, J. P. (2018). Adaptive memory: is the animacy effect on memory due to emotional arousal? *Psychon. Bull. Rev.* 25, 1399–1404. doi: 10.3758/s13423-018-1485-y

Mieth, L., Röer, J. P., Buchner, A., and Bell, R. (2019). Adaptive memory: Enhanced source memory for animate entities. *Memory*. 27, 1034–1042. doi: 10.1080/09658211.2019.1617882

Nairne, J. S., Thompson, S. R., and Pandeirada, J. N. S. (2007). Adaptive memory: survival processing enhances retention. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 263–273. doi: 10.1037/0278-7393.33.2.263

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

Pennington, J., Socher, R., and Manning, C. D. (2014). GloVe: global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1532–1543).

Popp, E. Y., and Serra, M. J. (2016). Animacy enhances free recall but impairs cued recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 186–201. doi: 10.1037/xlm0000174

Popp, E. Y., and Serra, M. J. (2018). The animacy advantage for free-recall performance is not attributable to greater mental arousal. *Memory* 26, 89–95. doi: 10.1080/09658211.2017.1326507

Rawlinson, H. C., and Kelley, C. M. (2021). In search of the proximal cause of the animacy effect on memory: Attentional resource allocation and semantic representations. *Mem. Cogn.* 49, 1137–1152. doi: 10.3758/s13421-021-01154-5

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., and Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: efficiently testing mean differences. *Psychol. Methods* 22, 322–339. doi: 10.1037/met0000061

Schwartz, B. L., and Brothers, B. R. (2014). "Survival processing does not improve paired- associate learning" in *What is adaptive about adaptive memory?* eds. B. L. Schwartz, M. L. Howe, M. P. Toglia and H. Otgaar (New York: Oxford University Press), 159–181.

Scofield, J. E., Buchanan, E. M., and Kostic, B. (2017). A meta-analysis of the survival-processing advantage in memory. *Psychon. Bull. Rev.* 25, 997–1012. doi: 10.3758/s13423-017-1346-0

Serra, M. J., and DeYoung, C. M. (2022). Within-pair factors might explain the inconsistent effects of animacy on paired-associates recall. *Psychon. Bull. Rev.* 30, 688–699. doi: 10.3758/s13423-022-02184-z

Simons, D., Shoda, Y., and Lindsay, D. S. (2017). Constraints on generality (COG): a proposed addition to all empirical papers. *Perspect. Psychol. Sci.* 12, 1123–1128. doi: 10.1177/1745691617708630

Simonsohn, U. (2015). Small telescopes: detectability and the evaluation of replication results. *Psychol. Sci.* 26, 559–569. doi: 10.1177/0956797614567341

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Blunt, J. R. (2013). Adaptive memory: Animacy processing produces mnemonic advantages. *Exp. Psychol.* 60, 172–178. doi: 10.1027/1618-3169/a000186

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Cogdill, M. (2015). Adaptive memory: Animacy effects persist in paired-associate learning. *Memory* 23, 657–663. doi: 10.1080/09658211.2014.916304

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N. S., and Cogdill, M. (2017). A categorical recall strategy does not explain animacy effects in episodic memory. *Q. J. Exp. Psychol.* 70, 761–771. doi: 10.1080/17470218.2016.1159707

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., and Van Der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *J. Pers. Soc. Psychol.* 100, 426–432. doi: 10.1037/a0022790

Xiao, X., Dong, Q., Chen, C., and Xue, G. (2016). Neural pattern similarity underlies the mnemonic advantages for living words. *Cortex* 79, 99–111. doi: 10.1016/j.cortex.2016.03.016

Zimmerman, C. A., and Kelley, C. M. (2010). "I'll remember this!" effects of emotionality on memory predictions versus memory performance. *J. Mem. Lang.* 62, 240–253. doi: 10.1016/j.jml.2009.11.004

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 02 June 2023
DOI 10.3389/fpsyg.2023.1167809

# Perceiving animacy from kinematics: visual specification of life-likeness in simple geometric patterns

Giulia Parovel*

Department of Social, Political and Cognitive Sciences, University of Siena, Siena, Italy

Check for updates

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University, United States

REVIEWED BY
Giorgio Vallortigara,
University of Trento, Italy
Ramon Daniel Castillo Guevara,
University of Talca, Chile
Alon Hafri,
University of Delaware, United States

*CORRESPONDENCE
Giulia Parovel
✉ giulia.parovel@unisi.it

RECEIVED 16 February 2023
ACCEPTED 11 May 2023
PUBLISHED 02 June 2023

CITATION
Parovel G (2023) Perceiving animacy from kinematics: visual specification of life-likeness in simple geometric patterns.
*Front. Psychol.* 14:1167809.
doi: 10.3389/fpsyg.2023.1167809

Since the seminal work of Heider and Simmel, and Michotte's research, many studies have shown that, under appropriate conditions, displays of simple geometric shapes elicit rich and vivid impressions of animacy and intentionality. The main purpose of this review is to emphasize the close relationship between kinematics and perceived animacy by showing which specific motion cues and spatiotemporal patterns automatically trigger visual perceptions of animacy and intentionality. The animacy phenomenon has been demonstrated to be rather fast, automatic, irresistible, and highly stimulus-driven. Moreover, there is growing evidence that animacy attributions, although usually associated with higher-level cognition and long-term memory, may reflect highly specialized visual processes that have evolved to support adaptive behaviors critical for survival. The hypothesis of a life-detector hardwired in the perceptual system is also supported by recent studies in early development and animal cognition, as well as by the issue of the "irresistibility" criterion, i.e., the persistence of animacy perception in adulthood even in the face of conflicting background knowledge. Finally, further support for the hypothesis that animacy is processed in the earliest stages of vision comes from recent experimental evidence on the interaction of animacy with other visual processes, such as visuomotor performance, visual memory, and speed estimation. Summarizing, the ability to detect animacy in all its nuances may be related to the visual system's sensitivity to those changes in kinematics – considered as a multifactorial relational system - that are associated with the presence of living beings, as opposed to the natural, inert behavior of physically constrained, form-invariant objects, or even mutually independent moving agents. This broad predisposition would allow the observer not only to identify the presence of animates and to distinguish them from inanimate, but also to quickly grasp their psychological, emotional, and social characteristics.

KEYWORDS

**animacy, intentionality, perception of causality, perceptual life-detector, life-like kinematics, expressive qualities, Michotte, Heider and Simmel**

## Introduction

In a fairly calm place, a sudden impression of motion immediately attracts our curiosity and awakens in us the impulse to discover the nature of that movement, to see if the moving object is an animated being, e.g., a cat or a fly, or a casual displacement of an inanimate object, e.g., a falling leaf. Motion is one of the key characteristics of animate things and thus, for our ancestors, immediately seeing and reacting to a potential danger or prey certainly played a vital role for survival. Detecting the presence of a living organism solely on the basis of its movement, without having to examine further visual details, can bring a relevant advantage for instance when the visibility conditions are lacking, because the moving object is too far away or blurred, shaded, partially hidden or camouflaged.

The possibility of manipulating the kinematic variables independently from other visual characteristics of the moving objects, and of exploring the relationship between these variables and the meaningful information that emerges, is precisely one of the reasons why the study of *animacy* – i.e., the character of "being alive" – is fascinating for vision scientists. Since the seminal works by Heider and Simmel (1944) and Michotte (1946/1963) to more recent research, many studies demonstrated that, under appropriate conditions, even displays of simple geometrical shapes might give rise to rich and vivid impressions of animacy and intentionality, by virtue of their simple movements and interactions. The main purpose of the present review is to highlight the close relationship between visual kinematics and perceived animacy, by drawing the whole picture of the specific motion cues and spatiotemporal patterns which automatically trigger rich and vivid visual percepts of animacy in geometrical patterns, independently from other appearance-based visual cues. A growing body of evidence for this specific sensitivity to kinematics-based animacy, as we will see, comes from a variety of research fields, ranging from experimental psychophysics to developmental psychology and causal reasoning, as well as animal cognition and neuroscience. It involves different methodologies and theoretical assumptions that are not always easy to compare. The paper then turns to the debate about the processes underlying animacy, contrasting the two main theoretical positions present in the literature: one, that the observer must activate representations – or schemas – and expectations in long-term memory in order to recognize the ongoing events (see Rips, 2011), and the other suggesting that observer can directly perceive high-level properties such as animacy phenomena, which are fairly fast, automatic, irresistible, and highly stimulus-driven (Scholl and Tremoulet, 2000; Scholl and Gao, 2013). The following sections review the increasing empirical evidence that animacy and its related properties – such as intentionality or social causality – are hardwired into the brain (Rutherford, 2013; Hafri and Firestone, 2021; Lemaire and Vallortigara, 2023). In particular, we examine recent findings on sensitivity to kinematic-based animacy cues in newborns; discuss the issue of "irresistibility," i.e., the persistence of animacy perception into adulthood even in the face of conflicting background knowledge and the co-presence of incongruent visual information; and review recent experimental evidence on how animacy interacts with other visual processes, such as visuomotor performance, visual memory, and speed estimation. The final section attempts to synthesize into a unified framework the features that single moving objects – as well as more complex kinematic patterns – must exhibit in order to trigger the animacy response of a hypothetical "life detector" (Vallortigara, 2012). In particular, it takes into account the visual system's predisposition to perceive spatiotemporal relationships between movements that are intrinsically endowed with information about the nature of ongoing events (Hafri and Firestone, 2021).

## Animacy by simple kinetic variations

As observed since Aristotle, a salient property of animated motion is its active character, quite different from the passive motion of a falling body or clouds driven across sky.

"Take, for instance, any animal: the animal moves itself, and we call every movement natural, the principle of which is internal to the body in motion" (Aristotle, Physics, vol. V, p. 307).

Indeed, self-propulsion is considered a strong cue to animacy in experimental psychology. The first systematic exploration of the role of *self-propelled* motion in animacy was carried out by the Belgian psychologist Michotte (1946/1963) in his research into animal locomotion. Living movements, he wrote, "have the appearance of being activities of which the object itself seems to be *the source*." He presented to observers non-rigid extending and contracting rectangles moving like a caterpillar or swimming like a frog (see a demonstration at: https://youtu.be/glEPmTd_EtA). Michotte reported that subjects showed great surprise and, without any prompting, literally described an animal crawling or creeping – an object *which moves of its own accord* (Michotte, 1946/1963, p. 185).

Recent research in this field confirmed the observation that a powerful perceptual cue to convey animacy is to appear self-propelled, i.e., moving by itself in the absence of an external cause, thus implying evidence of an inner energy source. Based on this evidence, many scholars have theorized about the involvement of causal attribution inferences and specialized cognitive processes in animacy recognition (Stewart, 1982; Leslie, 1984; Dasser et al., 1989; Premack, 1990; Scholl and Tremoulet, 2000; Gergely and Csibra, 2003; Csibra, 2008). Developmental studies suggested that infants recognize and distinguish self-moving objects from inert ones as early as 6 or 7 months of age (Leslie and Keeble,1987; Woodward, 1998; Pauen and Träuble, 2009). Di Giorgio et al. (2016) showed that seeing the onset of the self-propelled motion of an object, in contrast to it emerging from behind an occluding rectangle, is a crucial visual cue underlying animacy perception that allows even human newborns to differentiate between self- and non-self-propelled objects. When the onset of the motion is removed, newborns do not manifest any visual preference. Not only humans, but also newly hatched chicks demonstrated having an innate sensitivity to self-produced motion (Mascalzoni et al., 2010). Both these last works support the hypothesis of the presence already at birth of a predisposition to detect specific visual motion cues that might be a precursor to animacy percepts (Di Giorgio et al., 2016; Lemaire and Vallortigara, 2023).

Strictly related to the manifestation of an internal driving force in self-propelled objects is the prerogative of animate beings to not dissipate or even increase their observable kinetic energy. More in general, according to the Newtonian principles violation hypothesis, animacy can be triggered by simple movements that violate energy conservation (e.g., Stewart, 1982; Bingham et al., 1995; Gelman et al., 1995). Moving objects with sudden changes in direction and speed are more likely to appear animated (Tremoulet and Feldman, 2000; Träuble et al., 2014; van Buren et al., 2016).

Recently an animacy pattern has been revealed also in bouncing-like scenarios, obtained with a disk moving vertically downwards and then upwards, one or more times (Parovel et al., 2022; Vicovaro et al., 2023; see here few samples of the stimuli: https://youtu.be/Dt-QyXAjqNk). In specific conditions, depending on the simulated value of the coefficient of restitution, the visual impression of physical bouncing gave way to animate *jumping*. The most compelling animated jumping-like motions tended to occur when the stimuli showed a clear violation of energy conservation as well as multiple bouncing cycles (Bingham et al., 1995; Parovel et al., 2022; Vicovaro et al., 2023).

Speed itself seems to be a relevant factor for inducing animacy; Szego and Rutherford (2007) examined objects moving at constant speed, without changes in trajectory, and revealed that relatively faster objects appear animate, even if the speed difference is illusory. The authors argued that an inanimate object traveling across such a surface would be slowed by friction, so any object able to maintain a constant speed across a surface was likely to be self-propelled.

Also, Szego and Rutherford (2008) found that the visual system is sensitive to changes in the orientation of stimuli relative to gravity. By comparing "rising" vs. "falling" dots moving at the same speed, dots in upward motion were judged as animate more often than those moving downwards.

Consistently, it has been shown that observers are much more sensitive to speed changes in the direction opposite to the direction of gravity when they are required to report whether or not a speed change has occurred (Nguyen and van Buren, 2023).

Regarding the shape of the trajectory, other cues of living motion have been identified: (a) in C-shaped or S-shaped paths (Stewart, 1984; Gelman et al., 1995; Blythe et al., 1996, 1999); (b) in the mimicking of natural stimuli, such as flies, through speed and direction changes (Schultz and Bülthoff, 2013); furthermore, (c) in the alignment of the major axis of the shape of the moving object to its trajectory (Tremoulet and Feldman, 2000). A preference for parallelism between the principal axis of a moving object and its trajectory was observed even in visually-naive newborn chicks (Rosa-Salva et al., 2018).

Other decisive factors for intensifying the perception of animacy are the interactions of movements with other entities – discussed in depth in the next sections – and *goal-directedness*. Opfer (2002) for example demonstrated that unfamiliar shapes, i.e., blobs, with similar trajectories are identified by children and adults as living organisms if their movements are goal-directed (see also Gergely and Csibra, 2003; Schultz et al., 2005; Csibra, 2008). Even in displays where there was no visible goal, subjects often described the self-propelled movements as goal-directed but toward a target outside the observer's visible range, conveying an impression of intentionality (Tremoulet and Feldman, 2000, p. 947).

## Perceiving the relationship between movements: intentions and emotions

As discussed above, the kinematic properties of a single moving object can generate an impression of animacy. However, according to some authors, autonomous motion in itself cannot be a decisive factor in distinguishing animate from inanimate events, because it can be an ambiguous source of information (Gelman et al., 1995; Opfer, 2002). Indeed, a large amount of research has shown how interactions between simple moving shapes can elicit more sophisticated attributions of intentionality and psychological or emotional states, therefore supporting the fundamental role that perception plays in social cognition. Of course, as we will see, phenomenological reports, while essential to attest the emergence of these properties, must be supported by less direct methods of investigation to demonstrate the involvement of genuine visual processing in these scenarios.

In the classic Heider and Simmel (1944) experiment, observers presented with a cartoon-like animation in which two triangles and a disk interacted in and around a rectangular shape, attributed emotions, psychological traits, and intentions to those shapes. They used adjectives such as aggressive, shy, brave, intimidating, chasing, escaping, etc. and described the sequence as an interpersonal story in a remarkably consistent way. From the entire animation the authors extrapolated four basic combinations of movements: *successive* movements with or without spatial contiguity (corresponding for example to launching or joined movements in which one "causes" the other, such as action-at-a-distance), and *simultaneous* movements with or without spatial contact (corresponding to pushing, attracting, chasing and similar events). The movements, conjectured the authors, appeared organized in terms of acts of persons, and the interpretation of these movement-combinations varied according to the unit seen as the causal origin (the original animation is available at: https://youtu.be/8o6d9mUXwtg) (Figures 1).

According to Michotte (1950/1991), the relationship occurring between two or more moving objects within specific kinetic structures gives rise to primitive phenomena. These are to be considered quite different from the meanings which – under the influence of past experience – are attached to simple impressions of motions, merely juxtaposed in space and time. Certain combinations of visual stimuli, defined as to their distance, their speed, etc., cause certain specific impressions, for example, the impression "that an object A goes *toward* an object B," "that A *pursues* B," "that A *bumps* B," "that A *chases* or *repels* B," "that A *goes to find* B and *take it away*," and so on (Thinès et al., 2014, p. 104). These phenomena, stated Michotte, depend essentially on the system of stimulation, so that every notable modification in this system brings about a change in the expressed meaning of the relation. Within a certain distance between the objects, for instance, the impression of "approaching" is much stronger than the simple "shortening of the distance," and can vary in several qualitative ways such as a "friendly" or "aggressive" approach. On the contrary, speed varies only quantitatively.

In Michotte's launching effect, when an object A moves toward and makes contact with another object B, B is perceived *as if it were pushed* by A in a mechanical collision (Michotte, 1946/1963; Thinès

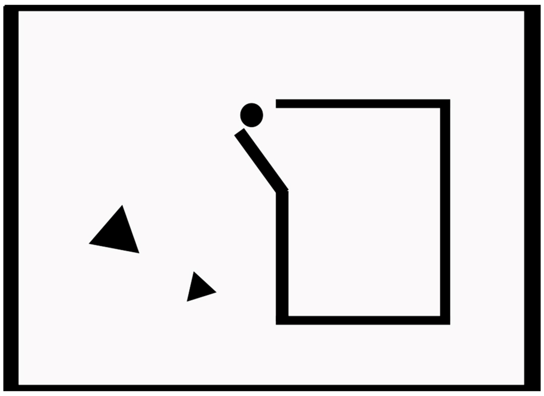


FIGURE 1
A frame of the classic Heider and Simmel (1944) animation.

et al., 2014). According to Michotte, speed ratio and temporal contiguity are crucial factors in the perceptual organization of causality. If an interval is introduced between the two movements, it brings gradual changes in the launching impression; when the delay lasts too long (more than 140 ms), the causal impression disappears (Michotte, 1946/1963, p. 92; see Hubbard, 2013). With regard to the speed ratio, Michotte found that when the speed of the second movement exceeded the speed of the first movement (i.e., about twice as fast), the launching effect gives way to the *triggering effect*. In this case, the motion of the second square was sometimes seen as having an active and self-propelled character (demos are available at https://youtu.be/6r9meK27Tpw) (Figure 2).

Michotte (1946/1963) noted that the triggering effect was surprising – even comic. It was investigated further by Kanizsa and Vicario (1968), who found that when B's motion is much faster than A, and B starts to move before the collision, B is described as if it were "seeing" the approaching object A and intentionally "escaping-away" from it. The authors named this action–reaction event as *intentional reaction* (see a demonstration here: https://youtu.be/BGjY61fzzd0). Observers spontaneously reported that the second object "runs away from the first," such as "to avoid it," or "because afraid" (Kanizsa and Vicario, 1968), thus exhibiting the presence of awareness and mental states even without goal-directedness (unless by goal we mean the intention to avoid a collision with the oncoming object). These and other similar impressions were later considered in infant and developmental research as a form of *psychological* or *social causality* (Schlottmann and Surian, 1999; Schlottmann et al., 2002, 2006, 2012, 2013; Schlottmann and Ray, 2010) or *action at-a-distance* (Spelke et al., 1995).

To investigate the relevant conditions for the perception of intentionality in adults, Dittrich and Lea (1994) designed experimental displays simulating a pattern of "one animal searching for another," that is a pursuing relationship between moving letters in an array of distractors, and collected ratings about the presence of intentions, interactions, and animacy. They found that the impression of intentionality depends directly on variations in motion parameters, such as the direction and speed of the target's motion, and the degree of goal orientation. Regarding the impression of animacy, they conclude that it may depend on both the presence of intentionality (i.e., goal-directedness) and the degree of interaction between the target and its goal.

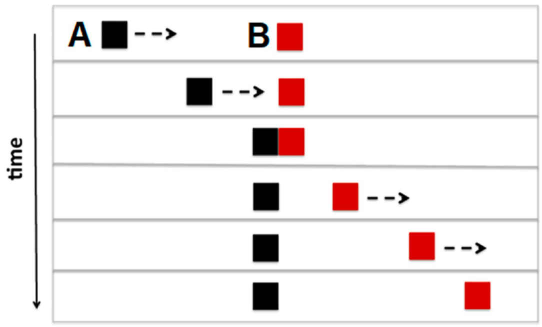


FIGURE 2
Six frames of *launching* or *triggering* animation, according to the speed ratio between square A and square B. Arrows indicate which object is moving in the various stages of the collision event.

A methodological limitation of these studies, concerns the choice of the dependent variables in the measurement of animacy and intentionality. While ratings or free descriptions of the visual stimuli are crucial to highlight the distinct phenomenology of these displays, they do not allow to separate the contribution of automatic visual processing from intervening higher-level reasoning processes based on such kinematic cues (see Firestone and Scholl, 2015; Van Buren and Scholl, 2018). For this reason, a relevant body of research focused on the psychophysics of *chasing* by adopting measures of visual performance that are better insulated from higher-level cognitive factors (Gao et al., 2009, 2010; Gao and Scholl, 2011; Gao et al., 2012) to better understand the interdependency between animacy and intentionality. It was shown that the human visual system is extremely sensitive in detecting chasing between two moving objects (a *wolf* and a *sheep*) in multiple objects displays. For example, Gao et al. (2009) introduced a specific methodological approach by employing visual search tasks in multiple object configurations and interactive displays. To control how directly the wolf approached the sheep, the authors varied the "chasing subtlety" – the maximal angular deviation of the *wolf*'s heading compared to perfect heat-seeking – and identified the optimal range of angular deviation from the straight chasing trajectory. Results indicated that with subtlety values above 30°, even when there was significant actual chasing, participants could not detect it. More generally, the perception of chasing is not a linear function of the degree of statistical correlation between wolf and sheep trajectories, but it depends on specific constraints.

In addition to investigations of chasing, Blythe et al. (1999) explored other basic categories of animate interaction, derived from evolutionary and ecological principles: pursuing, evading, fighting, courting and being courted, and playing. Interestingly, methodologically, the authors isolated the motion patterns of each category by means of a software built to allow subjects, interacting with each other across a computer network, to generate such behavioral trajectories. In this way, they extrapolated specific measures of trajectory parameters (velocity, vorticity, and energy) and plotted the most representative behavioral patterns.

From a theoretical point of view, while the energy conservation hypothesis implies that animacy can be triggered by simple movements that show an increase or change in their kinetic energy, other researchers suggested that the attribution of animacy involves – and may even require – something more than a failure in energy conservation, although this is necessary. Beyond motion, perception of animacy would be elicited by the inferences about the *causes of motion* – i.e., mental states, such as goal-directedness – in contrast to physical forces (Gelman et al., 1995; Premack and Premack, 1995; Opfer, 2002). In the perspective of these authors, even if triggered by particular combinations of visual features (Dittrich and Lea, 1994), the perception of intentionality and goal-directedness are strongly dependent on observer's mental contents and inferential processes. According to Blythe et al. (1999), for example, the motion cues would activate simple heuristic and automatic algorithms necessary for the categorization of agents' intentions. Many theories and computational models have been proposed to understand the development of causal cognition from infancy, based on the early predisposition to distinguish animate from inanimate objects from simple visual displays. In general, these approaches (e.g., Csibra, 2008; Baker et al., 2009; Gergely et al., 1995; Gergely and Csibra, 1997) strongly emphasize the role of rationality in understanding intentional behavior, such as the "teleological instance" or the "rationality principle." They are not discussed here, however, as they are beyond the scope of this study.

For the sake of conceptual clarity, by the way, it may be helpful to emphasize that theoretical or methodological reasons have led many researchers to identify separate constructs that go beyond the main animate-inanimate or physical-social distinctions. These constructs are for instance goal-directedness, intentionality or agency. Nevertheless, one must be aware of the strong interdependence between basic perceptual constraints and higher-level social impressions. In fact, each of these cues may act independently to some extent, but in general they coexist and interact in the natural environment (see Gao and Scholl, 2011; Rosa-Salva et al., 2016). Chasing, for example, combines multiple motion and relational cues related to animacy, such as self-propulsion, acceleration, direction change, and target approach. Even from a phenomenological point of view, being a chaser cannot be separated from being alive, even if the relationship is asymmetrical. It has been suggested that the relationship between animacy attributions and mind attributions is not discrete, but may vary along a continuum from attributions of "physicality" – related to more mechanical characteristics – to attributions of "personhood" – related to human-like behavior (Santos et al., 2008).

Furthermore, as happens in the studies reported in the next section, attributing intentions and mental states to the moving objects would modulate and intensify the impression of animacy itself (Dittrich and Lea, 1994; Gelman et al., 1995; Premack and Premack, 1995; Tremoulet and Feldman, 2006): configurations with two or more moving objects appear more animated than those with a single object (Tremoulet and Feldman, 2006; Heberlein, 2008; Parovel et al., 2018).

For this reason, at least in research that investigates the complex relationship between abstract kinematic patterns and animacy, a number of authors have chosen to use "animacy" to refer to a general perceptual skill. Thus, the term animacy generally refers not only to basic lifelike impressions (i.e., self-propelled locomotion), but also to its related properties, i.e., the infinite nuances that go from animacy to intentional attributions. In most of the papers examined, the term animacy is combined with intentionality or animate agents for greater clarity, and they are often used as interchangeable terms (Gergely et al., 1995; Scholl and Tremoulet, 2000; Opfer, 2002; Tremoulet and Feldman, 2006; Heberlein, 2008; Santos et al., 2008; Visch and Tan, 2009; Gao and Scholl, 2011; Scholl and Gao, 2013; Rosa-Salva et al., 2016; van Buren et al., 2017; Kominsky et al., 2021; Lorenzi and Vallortigara, 2021).

## The influence of context: how the presence of a second object affects animacy judgments about the target object

To better explore the possible common ground between single movement patterns and social displays in the elicitation of animacy and intention, some works added other elements to the trajectory of a single object. Even in simple scenarios with a single moving object, in fact, another simple geometrical shape added to the display is enough to trigger an increase in animacy ratings. Tremoulet and Feldman's (2006) displays showed a single figure (rectangular or round) moving on a screen and changing both speed and trajectory simultaneously while a static object (dot-foil or rectangular paddle) was placed in different positions. In this manner the static object defined several behavioral conditions for the target, such as moving *toward a prey* or *away from a predator* or being an obstacle. Tremoulet and Feldman found a small but significant effect of the context on animacy ratings, particularly in the goal/prey conditions. They suggested that a key factor in the perception of animacy is the attribution of an intention to the object - an intention that can be triggered by speed increase and change in direction alone, but that can also be specified by a supporting context.

More research was conducted to explore further the role of different spatiotemporal configurations on the perception of animacy and related properties, such as emotions and intentionality, in two-dimensional moving objects. In Parovel et al.'s (2018) work, the context consisted of a static or moving object that had the same shape as the target object (i.e., a small black square); in the dynamic conditions, the context object could exhibit either an animate-like (i.e., caterpillar locomotion) or a physical-like trajectory (bouncing event). The experiment was also designed to compare *approaching* vs. *avoiding* displays: it contrasted the relative directions between the target object and the context object. To obtain this, the context object could be located either at the beginning or at the end of the trajectory of the target (a sample of the stimuli is available here: https://youtu.be/4PyfhQoiVdk). Data were collected in both two-alternative forced-choice and Likert-scale rating tasks, and free reports were analyzed too. Results indicated a significant difference between static and dynamic contexts, where dynamic contexts prompted a distinctly clearer impression of animacy than static ones. Moreover, in the dynamic contexts it was consistently found that the impression of animacy was higher when the target was *moving away* from the context element than when it was approaching it. The moving-away behavior could be perceived as more animate for evolutionary reasons because of a higher sensitivity to threat-related events, such as fighting and chasing (Heberlein, 2008).

Psychophysical findings and free reports analysis suggested that there can be different facets to the animacy concept - for instance, an automatic animacy, an instinctive one and a mental/emotional one - and that an additional contextual element plays a crucial role in making them evident (Parovel et al., 2018). Visch and Tan (2009), in a similar perspective, showed that kinematic variations of abstract objects increase animacy attributions as well as specific emotional responses. For example, low velocity generating sadness and high velocity of the objects resulting in more "lively" movements. They speculated that animacy attribution is not only functional for social understanding and other adaptive purposes, but it also confers "reality status" and specific emotion correlations upon percepts of motion pictures. A recent paper confirmed the results of Heider and Simmel's experiment with school-aged children and found that when the rectangular figure, i.e., the house, was present in the display, children produced a higher proportion of animated descriptions (Hofrichter et al., 2021). The overall results corroborated the theoretical assumption which states that intentionality and other emotions are "related properties" of animacy (Visch and Tan, 2009; Gao and Scholl, 2011; Scholl and Gao, 2013).

More generally, the attribution of animacy appears to be significantly influenced by the addition of other elements where the target object is moving: if a single moving object A begins to move, the only frame of reference available is the environmental one, for example the direction up or down or the speed change with respect to the background. If another - static or moving - object is placed next to the first, the same locomotion of A will be better specified by the type of their interaction. It can appear for instance as an intentional approaching behavior - prudent, or determined, according to its speed - or an avoidance behavior – again, more or less quick and compelling. Again, it seems uncertain whether a clear distinction between the animacy cues conveyed by the kinetic variations of single moving objects and those related to the interaction of multiple movements is functional for understanding the nature of the perceptual processes underlying animacy. In both cases, animacy impressions basically depend on a system of spatiotemporal relations: in the case of a single movement, the relationships connect the moving object with its immediate context - such as the environmental coordinates or a static object – while in the case of two or more moving objects, their interplay allows perceptual grouping in space and time.

## The role of spatiotemporal contingencies

In the vein of Michotte's experimental work on perceptual causality (1946/1963), a significant amount of research has focused on the role played by the *spatiotemporal contingencies* between two or more moving objects in generating animacy (see Scholl and Tremoulet, 2000; Wagemans et al., 2006; Thinès et al., 2014). In examining the interdependence and similarity between animacy and causality, the temporal contiguity between the two moving objects is a common crucial variable, plausibly acting as a factor of perceptual grouping (Choi and Scholl, 2006; Duncker, 1935/1969; Schlottmann and Anderson, 1993; van Buren et al., 2017).

Bassili (1976) showed that temporal contingency was the crucial factor for the perception of an interaction between figures. Instead, spatial configuration of motion tended to determine the nature of that interaction, such as patterns of approach and avoidance. Because of the determinant role of temporal contingency, social interaction does not even require spatial contiguity. Schlottmann et al. (2002) and Falmier and Young (2008) found that *causality at a distance* – an action-and-reaction event similar to the intentional reaction- was easier to accept when the agents moved in an animate (caterpillar) manner and when the interaction was labeled as social (or psychological), rather than physical.

Another work, based on the manipulation of temporal contingencies between moving objects, showed that animacy experience increased with the time a moving object paused near a second object as well as with the increasing complexity of the interaction, such as approach and responsiveness, between the objects (Santos et al., 2008). Even a friendly/antagonistic communicative atmosphere can be induced by manipulating synchronous, coincident and not-coincident, movement of two egg shapes, on one side, and forward/backward/parallel tilting movement on the other (Yuasa, 2017).

In these situations, the movement of one object appears causally related to the movement of the other in a meaningful social relationship. In other cases, as it happens when two *casually* concomitant events are perceived *causally* related one to the other (see the example of the perceptual grouping between a door that shuts suddenly and the coming on of a light described by Duncker, 1935/1969), grouping can give rise to "incongruent" or "impossible" events, that may trigger even comic impressions. To assess if paradoxical causal contingencies between two trajectories that are incongruent and differently shaped are effective in evoking comicality, Parovel and Guidi (2015) combined Michotte's launching configuration and locomotion cues. Precisely, they modified the pattern of the second phase of launching in different ways, to obtain animated trajectories, such as a *frog-like* expanding and contracting trajectory or a *rabbit-like* jumping trajectory, as well as physical trajectories, such as *rotating* and *bouncing* squares (a sample of the stimuli is visible at: https://youtu.be/5EeihxEHdiY). The authors found that the paradoxical juxtaposition of animacy cues inside a launching relationship – while not of incongruent physical trajectories – elicited in the participants comical appreciations, in line with the Bergsonian theory of humor (for similar results see also Bressanelli and Parovel, 2012).

Results from Parovel and Guidi (2015) showed that temporal contingency has a crucial influence in evoking comic impressions: scale values and ratings of comicality actually tended to decrease with an increasing delay between the two movements. With a 200 ms temporal delay, it was still possible to get an impression of paradoxical causality between the two movements whilst, with a 1 s delay, perceived causality was disrupted. Interestingly, when spatiotemporal conditions convey an impression of psychological causality (−200 ms delay, and the speed of the first movement lower than the speed of the second movement), even linear trajectory events are judged amusing, confirming the previously quoted Michotte's observation about triggering (Michotte, 1946/1963).

Also, in other combinations of interacting moving objects, it has been observed that a change in the kinetic behavior of one object – i.e., a pause – elicits the perception of animacy only when a second object intercepts its trajectory in coincidence with the pause: the discontinuity in movement is then perceived as an intentional "waiting" (Minguzzi, 1961). Reasonably, kinetic conditions alone are not unequivocal and specific and so easily influenced by some other properties (Gyulai, 2000).

## Interim summary

The reviewed research has revealed the existence of multiple visual parameters inducing the observer to attribute animacy in moving objects. The term animacy, as we have seen, generally refers not to one specific impression, but also to a whole range of nuances of meanings. Consistently, with regard to the complexity of the scenario, these meanings can run from autonomous activity to emotional states such as, for example, fear or curiosity, as well as to psychological intentions, such as aggressive or shy, avoiding or approaching.

Moreover, a large body of psychophysical evidence, since Michotte's and Heider and Simmel's animations, has demonstrated a close dependence of animacy impressions on the spatiotemporal conditions of the stimulation. Therefore, as already seen, minimal variations of physical parameters correspond to discrete differences in the impression of animacy (Santos et al., 2008).

What clearly emerges is that, whilst observing such events, it is almost impossible to perceive the movement as neutral and meaningless, suggesting that the visual system is directly tuned not so much to the "objective" stimulation, but rather to the meaningful information conveyed by these movements – information of high adaptive and behavioral value. Detecting the presence and understanding the intentions of other agents is crucial to survival and reproduction. Thus, it is plausible that humans and other species evolved to be extremely sensitive to signals of animacy and agency, and that they possess fast and unlearnt mechanisms for the detection of them (Mascalzoni et al., 2010; Vallortigara, 2012; Abdai et al., 2017; Lorenzi and Vallortigara, 2021).

Although this paper does not incorporate neuroscientific evidence, it is important to mention that the neural substrates associated with animate motion processing are at least in part distinct from those associated with inanimate motion. The exposure to such visual stimuli elicits strong activation in the temporoparietal cortex, including areas in and near the posterior superior temporal sulcus (pSTS) and angular gyrus, especially in the right hemisphere (Castelli et al., 2000; Blakemore et al., 2003; Heberlein, 2008; Santos et al., 2010). Some data showed that the ability to detect animacy from contingency of objects reacting to other objects is processed by specific networks which are different from brain regions associated to theory of mind tasks. The detection of agency on the basis of cues such as movement and contingency, according to these authors, might be a precursor of our ability to infer other people's mental states (Blakemore et al., 2003). Additionally, to explain the neural substrate underlying the understanding of animacy, two hypotheses have been proposed: the mirror-system hypothesis (Gallese et al., 2004) and the social-network hypothesis (Adolphs, 2003; Wheatley et al., 2007), each engaging anatomically distinct neural substrates (see for a review, for instance, Slotnick, 2013).

## Is animacy directly perceived?

Is animacy a visual property, like shape and size, or the result of automatic reasoning? The "place" of animacy and causality in our perceptual experience is a central theoretical question in experimental psychology and cognitive science. It concerns the complex relationship between perception and cognition.

The nod of debate can be formulated in these terms: (a) do we see low-level cues evoking top-down perceptual judgments about animacy or, (b) interactions themselves are meaningful, as stated by Michotte, because they are the result of bottom-up features within visual processing?

A major obstacle to finding a convergent solution is posed by the fact that the existing positions in cognitive psychology are based on different theoretical premises concerning the nature of the information available to our senses, the so-called "thin" and "rich" views (see Toribio, 2018). According to the "thin" view, the texture - or information - available to the sensory system is limited and insufficient to specify the properties and the events of the world. Thus, animacy impressions - even when elicited by low-level features such as color, shape, texture, or motion - would depend on high-level inferences drawn from information present in long-term memory (see for reviews Rips, 2011; Scholl and Gao, 2013). This is not to say that these properties do not have a compelling appearance as objective properties, but that their phenomenology is insufficient to prove their true low-level nature. Spatial, temporal, and other visual cues would be processed and automatically detected, and this immediacy would be erroneously attributed to visual processing itself. While watching Michotte's like demos, observers would activate representations – or schemas – and expectations in long-term memory relative to the ongoing events in order to recognize them (see Weir, 1978; White and Milne, 1997; Tenenbaum and Griffiths, 2003; White, 2006). These schemas are post-perceptual and, according to these theories, can be acquired with experience and modified by beliefs and expectations.

Otherwise, the so-called "rich" view upholds the hypothesis that the visual system can directly detect meaningful relationships and interactions between objects. This view has its foundations in Michotte's claim, and it was later reformulated as a thesis about a module-based perception of animacy and causality in line with Fodor's perspective, and received much empirical support from, among others, Scholl's research group (see for reviews Scholl and Tremoulet, 2000; Scholl and Gao, 2013). Such perceptual modules would be informationally encapsulated and therefore not shaped by prior knowledge, inference, or expectation (Saxe and Carey, 2006). According to "rich" theories, high level properties are visually represented and not just seemingly represented as a result of a perceptual judgment (Rips, 2011; Toribio, 2018). This position is compatible with the idea that observers have specialized detectors, hardwired in the perceptual system, to take over physical, biological and social interactions (Vallortigara, 2012). Perceptual animacy and causality may occur on first exposure without requiring prior experience with the events. Further learning would take advantage of this elementary, original knowledge and would shape more sophisticated cognitive skills and behaviors (Lorenzi and Vallortigara, 2021).

The main arguments that have been put forward for and against the different positions, i.e., the "perceptual" view and the more widespread "top-down" view, will be briefly mentioned in the following paragraphs. Then, we will address three additional findings in support of the involvement of automatic visual processing in the impressions of animacy: the evidence in favor of a sensitivity to animacy in newborns; the lifelong persistence of animacy through motion despite visual incongruity; the interactions between animacy and other visual processes that have been recently documented.

Scholl and Tremoulet (2000, p. 299) claimed that causal relations and animacy are rich and vivid properties of visual displays, and are "fairly fast, automatic, irresistible, and highly stimulus-driven" phenomena. The phenomenal character of *vividness*, however, is not a valid argument, since it is recognized by both approaches. The same can be said for the apparent effortless and unawareness typically associated with the perception of animacy and causality. Many post-perceptual judgments can also occur effortlessly, automatically, and unintentionally, and even other cognitive processes unrelated to vision (e.g., semantic priming) share these properties without being hardwired in the early stages of visual processing.

On the other hand, *fastness* and *automaticity* can be plausibly related to the adaptive role of this sensitivity, i.e., the satisfaction of vital biological and social needs. Vallortigara (2012) conjectured the existence of a sort of perceptual "life detector" in the brain, inspired by Darwin's suggestion about primitive neural pathways to ensure a bias to attend toward living things. Behavioral and neuroscientific evidence for an innate predisposition to animacy cues comes from research on animal and human newborn cognition. Such data supports the idea that the selective pressure to quickly detect and respond to the presence of other creatures has shaped the brains and behaviors of distant animal species in similar ways throughout evolution (see, for a recent review, Lorenzi and Vallortigara, 2021). It is highly advantageous for animals, the authors argue, to be born with preprogrammed mechanisms for directing attention to salient categories of stimuli, such as animacy and agency, rather than having to learn them through long sequences of trying and failing. This also could lay the groundwork for further refinement as development proceeds.

Furthermore, most of the findings in this area show "dramatic effects of very subtle stimulus manipulations. This is a hallmark of perception" (Scholl and Gao, 2013, p. 198). In other words, the very close *dependence of causality and animacy impressions on the objective spatio-temporal conditions* of the stimulation – such as distance, duration, speed of movement, and so on -, would have been extremely difficult to explain on the basis of previous knowledge and experience (Runeson and Frykholm, 1983; Santos et al., 2010). Similarly, according to Butterfill (2009), observer's sensitivity to some causal interactions and dependency on very brief temporal delays is properly perceptual and categorical, just as phoneme perception is. According to Butterfill, however, the role of causal categorization would not be to convey information about the nature of the event, but only to allow the observer to distinguish different events from each other.

Another classic argument, since the work of Michotte (1946/1963) and Heider and Simmel (1944), concerns the *unanimity of subjective reports* and their inconsistency with individual differences.

Some studies found individual differences in their investigation on causal impressions (Gemelli and Cappellini, 1958; Beasley, 1968; Young and Falmier, 2008). However, many of these studies are difficult to evaluate because they do not clearly separate perceptual processing from other effects, such as those due to task design or other uncontrolled individual dispositions (Rips, 2011). In order to overcome this problem, some authors have preferred to avoid methods like ratings and individual judgments – that appear to be particularly susceptible to being contaminated by post-perceptual judgment (see Firestone and Scholl, 2015; Van Buren and Scholl, 2018), adopting tasks that measure visual-motor performance (Blythe et al., 2009; Gao et al., 2009) or two-alternative forced-choice comparisons (Parovel et al., 2018).

The *irresistibility of the stimulus*, i.e., the cognitive impenetrability of these impressions, is a further criterion that has been proposed to support the hypothesis that animacy is genuinely perceived and not the result of a perceptual judgment (Scholl and Tremoulet, 2000). The impression of animacy is compelling and irresistible; thus, animacy and related properties can be assimilated to standard visual illusions, whose appearance persists even when we know about their objective conditions (see Pylyshyn, 1999). Toribio, in her theoretical work (2015), argues that the irresistibility of the stimulus criterion is the most important evidence supporting the visibility thesis for high-level properties like animacy. Specifically, she refers to the experimental results of Gao and Scholl (2011), which imply the use of information that is not available to the subjects, thus ruling out any effect of top-down inference. More generally, Toribio notes, subjects are well aware that the geometric shapes in motion on the screen are not animate. However, even in the face of conflicting background knowledge, under certain conditions they cannot help but experience such characteristics.

More skeptical is the position of Rips (2011), who provides a theoretical overview of all the work on causal processing along the whole range of conditions, from mechanical launch to animacy and intentionality attributions. Using data from infant and animal studies, cognitive and neuropsychological dissociation experiments, and studies of context effects and individual differences, the author contrasts the two main explanatory hypotheses: the cause-detector hypothesis – a reformulation of the perceptual module hypothesis – and the causal schema hypothesis (Weir, 1978). In the latter, representations or schemas of simple interaction patterns (e.g., launching, triggering, pulling, and so on) are the result of non-modular inferences based on long-term memory information.

The two models, Rips (2011) argues, differ on many assumptions, such as information encapsulation, innateness, the role of development, individual experience and cultural differences, but not on others, such as fastness, automaticity and unanimity. For this reason, much of the evidence does not allow for a distinction to be made between the two competing theories. In any case, in Rips's view, module evidence does not mean that we perceive causal events, but rather spatiotemporal relations that are informed by higher-level knowledge.

A recent study by Morales-Bader et al. (2020) emphasized the role of semantic cues and high-level processes in animacy judgments. They suggested that the tendency to attribute intentionality in Heider and Simmel-like displays can be affected by the interaction between perceptual and semantic cues (i.e., figure shape, label, and apparent speed). Interestingly, by the way, when the authors contrasted the effect of the figure's shape on the attribution of intentionality, they found that triangular shapes were attributed more intentionality than anthropomorphic-stickman figures. This was interpreted as anthropomorphic figures acting as distractors to the type of movement, which was the main cue that led to animacy and intention attributions. Unfortunately, as mentioned above, the methodology adopted by the authors, i.e., categorizing participants' free descriptions, does not allow separating the role of genuine visual processing from the intervention of automatic inferential judgments.

Furthermore, the empirical evidence provided by developmental studies, shows that perceptual causality may be available early, at a time when relevant experience is limited, while the simplest form of causal reasoning develops much later (Schlottmann et al., 2002; Saxe and Carey, 2006). The view of Schlottmann et al. (2006) to explain this gap is that perceptual causality is useful early in children's development because it allows identification of causal events for themselves without need to reason about "why" these events are cause and effect. In fact, children cannot always integrate perceptual constraints with causal mechanisms – the underlying structure of events – until later in development. Generally, according to these authors, developmental evidence cannot exclude a post-perceptual role for higher-level knowledge and learning (see also Vicovaro, 2018).

The presence of animacy-related kinematic constraints from the earliest days of life and their developmental function, in contrast to the appearance-based visual features associated with animate beings, is another developmental issue worth exploring in more detail. It will be partially covered in the next section.

## What newborn babies are attracted to: animated movement and face-like patterns

Ample empirical evidence from infant research supports the notion that from birth the human system is broadly tuned to detect social stimuli on the basis of at least two independent properties: the presence of a face and the way something is moving. Nevertheless, the ontogenetic origin of this sensitivity is still under debate.

Evidence from fMRI literature on functional characterization of cortical responses in infants demonstrates that the cortex of 4–6-month-old human infants is already spatially organized, with distinct regions responding preferentially to human faces versus natural scenes (Deen et al., 2017).

Many data have shown the existence of mechanisms that ensure that newborns' attention is triggered by faces, and that they manifest preference for schematic and real faces. With regard to face-like patterns, it has been shown that newborns look more frequently and longer at geometric stimuli with more elements in the upper part when compared to the inverted version. This preference, would allow newborns to successfully choose faces from other non-face-like stimuli (Simion et al., 2002; Turati et al., 2002; Macchi Cassia et al., 2004).

It has also been recognized that infants are predisposed to attend preferentially to the motion of biological entities, even when presented in the most rudimentary form. These predispositions are thought to be controlled by rapid and automatic subcortical orienting mechanisms, and their presence at birth would contribute to the development through progressive specialization - as a function of experience - of the "social brain" network (Troje and Westhoff, 2006; Yoon and Johnson, 2009; Simion et al., 2011; Di Giorgio et al., 2016).

The first strong evidence for an innate ability to detect biological motion and to respond to it in a specific way came from non-human animal species. Imprinting procedure revealed that newly hatched chicks at their first exposure to point-light displays preferentially approached biological motion compared to nonbiological motion stimuli (Regolin et al., 2000). Moreover, the data suggest a non-species-specific sensitivity to biological motion, as chicks showed no preference for the patterns displaying a walking hen over configurations displaying a walking cat or even scrambled biological motion, all of which are preferred over stimuli that do not display animated motion (Vallortigara et al., 2005; Vallortigara and Regolin, 2006). Similar to adult human observers, visually naive chicks also showed a significant preference for moving stimuli that changed speed relative to stimuli with constant speed, and for stimuli that changed direction (Rosa Salva et al., 2015; Rosa-Salva et al., 2016). Recent neuroscientific evidence has revealed the involvement of subcortical areas of the avian brain in response to stimuli showing speed changes, as compared to those showing constant motion (Lorenzi et al., 2017).

In humans, it has been shown that 4- and 6-month-old infants respond to biological motion stimuli, as they tended to look longer at a point-light display of a walking person than at an array of randomly moving elements (Fox and McDaniel, 1982).

More recent results demonstrated sensitivity for the dynamics of biological motion and to the gravitational forces acting on motion even in newborns. To rule out the possibility of any previous experience, authors adopted hen-walking animations rather than human-walker animations and found that at their first exposure, 2-day-old babies preferred biological motion over random motion point-light displays (Simion et al., 2008; Bardi et al., 2011). Furthermore, newborns choose the upright point-light display of a walking hen over the same display inverted.

Changes in speed seem to be relevant as well, as recent developmental studies demonstrated. When presented with speed changes, newborns showed a preference for a particular speed pattern, i.e., an increase followed by a decrease in speed. In contrast, the reverse sequence pattern or a single speed change do not elicit any visual preference (Di Giorgio et al., 2021).

Regarding the perception of visual features other than the face in the first weeks of life, findings on infant visual categorization development – obtained by comparing behavioral data (i.e., the looking behavior) and brain-activity recording-, suggest that animacy is the earliest categorical distinction of visual objects in infancy (Spriet et al., 2022). However, until 4 months of age, infants' looking behavior when presented with a series of images of real-world objects, shows no evidence of an animate-animate distinction, while revealing a preference for size, elongation, and compactness of objects. Four-month-old infants continue to prefer human and nonhuman faces and bigger objects, but they also show categorizing by animacy. By 10 months, image categorization by animacy emerges despite differences in image size and it is consistent with the cortical organization of object-related information recorded from anterior (temporal) aspects of the visual ventral stream in adults (Spriet et al., 2022).

In summary, it has been theorized that the ability to perceive a wide range of face types, regardless of species specificity, is of great adaptive value for infants (Nelson, 2001). In a similar way, it could be hypothesized that an inborn broad sensitivity to life-like movements in infants – a life detector – could be very advantageous in directing attention toward living things and in differentiating them from inanimate (Vallortigara, 2012). It would allow infants to discriminate between differently shaped entities and patterns, providing crucial support for visual experience in the development of categorical inferences, animacy-related, as well as responses to primitive perceptual features – such as mid-level features. It has been suggested, in fact, that much of object-selective cortical organization can be explained by relatively primitive mid-level features without requiring explicit recognition of the objects themselves (Long et al., 2018).

More generally, a broad sensitivity to animated motion would have the advantage of great flexibility and attention to an equally broad range of possible events. The possibility of committing biases in the sense of attributing life to non-animated objects would be well compensated by avoiding the opposite – and worse – risk.

## Falling leaves or butterflies? An aesthetical side effect of the *irresistibility* criterion

It remains yet to be clarified how these innate mechanisms evolve with the development of more complex and detailed cognitive capacities. According to several authors, perceptual narrowing would occur for instance with increasing experience with certain types of faces and lack of exposure to other types of faces. This would allow the human system to increase its ability to discriminate the highly experienced faces and decrease its ability to discriminate the infrequently experienced faces (Nelson, 2001; Simion et al., 2011). Also, considering perceptual causality, it has been argued that it is an innate tool with the role of supporting learning about the causal texture of the world, and then learning gradually influences perception (Schlottmann et al., 2006).

Nevertheless, much of the literature reviewed in the first sections of this work emphasizes the persistence and automaticity of these impressions into adulthood, despite the knowledge acquired through learning (i.e., the *irrestibility* criterion). If the acquisition of more complex cognitive skills leads to the merging of perceptual constraints and learned mechanisms, then the acquired knowledge about physical mechanisms and social behavior of organisms should prevent adults from juxtaposing mechanical and animate features in perceptual events. On the contrary, animacy impressions take place even in conjunction with relatively incongruent visual information such as geometric moving shapes, remain vivid after many repeated observations, are easily induced and surprise - and fascinate too - the observer. This is well represented by many side effects, other than animated cartoons with non-anthropomorphic forms: we all experience the erroneous attribution of life to inanimate moving objects in the natural environment, e.g., when a leaf blown by the wind seems a living creature - like a butterfly. Or, when a very elastic bouncing ball seems to jump. Similarly, familiar non-living objects can appear alive in movies created by "stop motion" techniques (i.e., a filmmaking technique in which an object is moved in small steps and a photograph is taken at each step), or by performing other video editing operations, such as rewinding playback – which can, for example, make a dropping object appear to rise up against the force of gravity. The pleasure elicited by puppets and marionettes may be another side-effect of the persistence of animacy through motion despite visual incongruity. Actually, many great thinkers have emphasized the importance of juxtaposing, in the same event, mechanical and animated visual qualities to achieve a comic or surreal effect (Bergson, 1900/2013; Ejzenstejn, 2004).

A similar side effect can be speculated to occur in the perception of faces, as suggested by the common phenomenon of pareidolia, the tendency to perceive a face even in a non-living object, such as the moon, clouds, rocks, or the front of a car, a phenomenon described since Leonardo Da Vinci (see for instance Ichikawa et al., 2011). Even in these cases, we see both the objective and inanimate nature of the object and the manifest presence of the face at the same time. Not only that, but we cannot help but even see that the face expresses psychological traits.

This evidence suggests the lifelong persistence of these innate mechanisms and their independence, at least in part, from developmentally acquired inference and categorization processes, thus allowing flexible adaption to changing circumstances (Schlottmann, 2001). In presence of purely casually contingent (or intentionally induced) animacy-related kinematics, learning and reasoning warn us and prevent us from inferring "objective" causality or animacy in the visual scene. On the other hand, reasoning, fortunately, does not prevent the aesthetic enjoyment of surprising and of vivid paradoxical effects when they occur, by choice or by chance. Given the crucial importance for human - and not only human - observers of the detection of life and agency, it is plausible to speculate, as some authors have argued, that we have evolved to be very sensitive, or even *overly sensitive*, to animacy and agency (Vallortigara, 2012).

## Interactions of animacy with other visual processes

As seen in the previous sections, animacy and intentional relationships between moving objects are extracted rapidly and automatically, are sensitive to subtle visual parameters, appear early in development and are present in non-human species. For these reasons, these phenomena would show important hallmarks of automatic visual processing (see Scholl and Gao, 2013). Recently, some authors have further challenged the top-down perspective and claimed that, by virtue of their ecological and adaptive relevance, the perception of animacy and intentionality may be integrated into the mind in ways that are deeper than previously imagined. They hypothesized that animacy may interact with other perceptual processes (van Buren and Scholl, 2017; Hafri and Firestone, 2021). If perceptual animacy can influence other low-level visual features, this should be a further evidence supporting the theory that animacy is processed in the earliest stages of vision and is not a high-level projection added by the observer's mind to neutral stimuli.

In the last two decades abundant empirical evidence has emerged supporting that the perception of launching events can have an influence on other processes in visual cognition. The following paragraphs will first summarize these works and then mention some recent findings on new specific interactions between animacy and other visual processes.

It has been shown that the launching effect can imply: (a) a contraction (two objects appear closer in space when they are causally connected; Buehner and Humphreys, 2010), or (b) an extension of the perceived distance between the colliding squares A and B at the moment of impact, i.e., the degree of overlapping between the two items is underestimated and the degree of underestimation is higher when the causal nature of the event is induced by a surrounding context (Scholl and Nakayama, 2004); (c) a distortion of the perceived trajectory of the apparent motion of A (Kim et al., 2013); (d) larger displacements in the remembered vanishing position for moving targets when the launcher was faster than the launched object (Hubbard et al., 2001; Hubbard and Ruppel, 2002; Choi and Scholl, 2006; De Sá Teixeira et al., 2008); (e) a distortion of the remembered temporal order of the motions of the squares A and B (Bechlivanidis and Lagnado, 2016). Apparent kinematics itself, (f) can be biased in launching events: in certain conditions the perceived speed of B is influenced by the speed of A (Parovel and Casco, 2006; Vicovaro et al., 2020). Causal relations are also visually "prioritized "in the following ways: (g) participants become aware of causal events more rapidly than non-causal events (Moors et al., 2017), (h) launching events are subject to retinotopically specific visual adaptation (Rolfs et al., 2013), and (i) in visual search tasks, adults' causal perception distinguishes between triggering and launching events and this ability cannot be attributed to low-level differences in sensitivity to differences in speed. Instead, according to these authors, this categorical boundary is directly determined by constraints on perception (Kominsky et al., 2017).

Concerning the influence of perceptual animacy on other visual processes, the following paragraphs will discuss, respectively, visuomotor performance, visual memory and speed estimation.

A series of studies has demonstrated that animacy cues influence visuomotor performance (Gao et al., 2010). For example, in various interactive tasks in which participants controlled a disk within a display filled with randomly moving darts, the task of "avoiding" from a pursuing "wolf" disk was less successful if the randomly moving darts remained oriented straight toward the subject's disk (*wolfpack* displays). According to the authors, this happened because in this condition the wolfpack darts were perceived (erroneously) as actively pursuing the subject's disk. Van Buren et al. (2016) explored whether such displays would influence performance even when the putatively animate objects were entirely irrelevant to the task, and subjects were asked to ignore them. Also in this case, subjects took longer to perform their task – to collect dots as quickly as possible – when the irrelevant background darts were always pointing at the disk they were controlling, rather than 90° away from it.

Following the same line of investigation, Van Buren and Scholl (2017) explored the influence of perceived animacy and goal-directedness from simple geometric shapes on spatial memory performance. In particular, they wondered if a matching task between pairs of wolfpack panels, in which participants saw animations with both "darts" and discs with sketched "eyes," would be influenced by animacy cues. Results showed a spatial memory advantage for stimuli that were perceived in animate and intentional terms, and these effects occurred both with "darts" and "eyes." The authors emphasized that the wolfpack panels were prioritized in memory over all other types of panels, showing a robust effect; they suggested that perceiving animacy can really matter for downstream processing.

Finally, a recent work discovered an illusory speed effect in displays conveying animacy (Parovel and Guidi, 2020). A first experiment was based on previous research reported above (Parovel et al., 2018), which found that (a) a moving square created an impression of greater animacy in dynamic contexts than in static ones, and (b) when the target moved away from the context element than when it approached it. In this work, instead, two-alternative forced-choice comparisons were used to test whether the perceived speed of the target square varied across the same set of stimulus conditions. Results showed that an escaping object looked *faster* than an approaching one or neutral one, moving in absence of any context (some demos are available at: https://youtu.be/p17c41B_lq8). In a second experiment, the perceived speed of the escaping black square was psychophysically measured in a condition similar to the *intentional reaction* (Kanizsa and Vicario, 1968), where a two-dimensional square moves toward another square, which gets away before the first square reaches it. The point of subjective equality (PSE) estimates indicated that the speed of the escaping moving object was overestimated between 6.7 and 10.2%, according to the type of motion of the chaser (*linear* vs. *caterpillar-like*). In conclusion, the speed overestimation was found only in the escaping condition and not in the approaching one, and it was stronger when the contextual element, the chaser, moved like a caterpillar.

To summarize, the empirical evidence described in this section - analyzing the influence of animacy on visual performance, visual memory, and speed perception – provides further support to the hypothesis that animacy perception is hardwired in the visual system (see Hafri and Firestone, 2021). It is interesting to note how all the interactions reported above between animacy and other visual processes are concerning "chasing" situations, i.e., potentially threatening events that would therefore require an immediate behavioral reaction. If there exists a perceptual "life detector" hardwired in the brain (Vallortigara, 2012) overall, it seems extremely plausible that it should interact rapidly and efficiently with other visual abilities, favoring appropriate visual-motor skills to quickly react with adaptive behaviors to the surrounding events.

In addition, social relationships involving interacting human figures exhibit further perceptual specificities (see Hafri and Firestone, 2021). To mention just a few findings in this area: (a) extraction of event structure from visual scenes is rapid and spontaneous, as shown in dynamic sequences of two-person scenes, designed to distinguish actors from patients (Hafri et al., 2018); (b) visual search advantage found for face-to-face, relative to back-to-back dyads (Vestner et al., 2020); (c) interacting individuals are remembered as physically closer than are noninteracting individuals (Vestner et al., 2019); (d) meaningful interaction between human agents helps working memory to compress the movements to be stored into a chunk (Ding et al., 2017); (e) visual adaptation aftereffects have been reported suggesting selective coding mechanism for action contingencies (Fedorov et al., 2018).

More generally, all these data are consistent with similar findings on the attentional visual prioritization found in detecting animate objects, using natural looking images (Altman et al., 2016; Bailey and Lang, 2022; Long et al., 2018). The discoveries made in this field are generally interpreted as an additional support for the animate-monitoring hypothesis (New et al., 2007), which suggests that early detection of animacy may have endowed our hunter-gather ancestors with survival advantages, by means of perceptual features that have remained consistent throughout hominid evolution.

## The life-detector's role: a broad-range sensitivity to the ongoing changes in a multiple relational system?

As seen so far, several lines of research support the hypothesis that animacy and its related properties are hardwired in the brain and are automatically processed in the earliest stages of vision. According to these findings, animacy features – originally defined as kinetic structures by Michotte (1946/1963), or spatiotemporal gestalten by Heider and Simmel (1944), likely lend themselves to being conceived as prelinguistic visual primitives (Mandler, 1992) or as social affordances, whose meaning can be directly perceived and solicit the animal's behavior and affect (Gibson, 1977; Withagen et al., 2012; Withagen, 2022), rather than as the top-down result of higher-level processes of recognition and categorization.

In this perspective, what is still unknown is if the numerous kinds of animacy and social interactions are modular specific or depend on a unitary animacy-detector system. In the first case, infants would exhibit separate core systems for animate and inanimate objects (Leslie, 1994; Premack, 1990; Mandler, 2003; Spelke and Kinzler, 2007; Vallortigara, 2012; Scholl and Gao, 2013). In the second case, some theoretical frameworks, such as the cue-based-bootstrapping model, speculate that innate predispositions to low-level visual cues linked to animate beings lead to the development of animacy perception through learning (Schlottmann et al., 2006; Biro et al., 2007; Spriet et al., 2022). Since birth, humans would display some attentional biases toward rudimentary low-level visual cues of motion – such as start from rest by self-propulsion and speed changes - that elicit animacy perception also in adults (Di Giorgio et al., 2021). The exact nature of such low-level cues, i.e., whether the information they carry can be considered only an initial precursor, or whether it is inherently significant information about animacy, is still uncertain.

A further open question concerns whether the cues eliciting the impression of animacy belong to specific categories, such as animate-inanimate (Kominsky et al., 2017). One possibility, proposed by Gao and Scholl (2011), is that different kinematic cues would correspond to specific animacy categories, such as triggering, chasing, approaching, and so on, and all these individual cues could be observed – and investigated – either in isolation or mixed together in common displays. Anyway, although psychophysical experiments usually isolate animacy-inducing parameters, their interaction – for example in more complex displays or in real life situations - can produce results that are not additive and therefore not predictable (Tremoulet and Feldman, 2006). In general, the existence of a specific module for perceiving each form of animacy seems problematic to maintain (see Rips, 2011). Even if perceptual relations generally reflect distinct categories (see Hafri and Firestone, 2021), both phenomenology and experimental psychophysics suggest that the animacy construct must be defined in a broad sense. Given the theoretically infinite range of animacy-related impressions, animated and social events may be better understood not as all-or-nothing properties, but by allowing for the possibility of intermediate categories that are perceptually meaningful.

In this view, the ability to grasp animacy in all its nuances could be understood as a broad-range sensitivity toward those characteristics of kinematics which involve the presence of living beings and agents. This would allow for not only recognition of their presence and of their animate movements, but also to quickly grasp their psychological, emotional and interpersonal characteristics (for instance, being calm, hasty, friendly, avoiding, nervous, unsure, edgy, etc.). Perhaps even moral instances such as helping or hindering, i.e., altruistic versus selfish behavior, can be directly captured and differentiated, as Kuhlmeier et al. (2003) have shown in 12-month-old infants.

This wide-range sensitivity could be rooted in the predisposition of the visual system to perceive spatiotemporal relationships between movements that are intrinsically endowed with information about the nature of the ongoing events. Indeed, across physical and social domains, current findings and theories have reinforced the possibility that meaningful relationships between movements are properly perceived for themselves and reflect highly specialized visual processes (Van Buren et al., 2016; Hafri and Firestone, 2021).

In summary, what characteristics must a moving object, or a global kinematic pattern, have in order to trigger the animacy response of a hypothetical life-detector broadly tuned to meaningful relationship?

As Aristotle wrote, if there is no external force putting it into motion, a moving object appears as having an inner force, i.e., life. Considering all the nuances of animacy that may appear in different types of interaction, we can speculate that besides this biological force, other apparent causes, namely psychological and social, may emerge from kinematic displays. From this perspective, multiple cues must be considered in order to search for a common perceptual sensitivity that might encompass and integrate them together.

As previously seen, a life-detector should be sensitive (a) to the onset/presence, to the changes in speed of already self-propelled moving objects (Lorenzi and Vallortigara, 2021). Furthermore, in animal and human locomotion, (b) a life-detector has to be able to identify the relationships between the constitutive parts of the object, such as the head and tail – e.g., in caterpillar-like non-rigid stretch-and-squeeze motion–, as well as the interaction between several individual points – e.g., in a biological movement pattern. Indeed, in the movement of vertebrates, the spatial relationships between some parts of the body are constantly changing, while the spatial relationships between other parts, which represent connected joints, remain invariant. In other words, the moving object has to be related to structural invariants, such as semi-rigidity principles versus the spatial constancy displayed by rigid inanimate objects (Simion et al., 2011).

Moreover, a moving object, in order to trigger the animacy response, (c) has to be related to physical constraints, such as force of gravity and energy conservation principle (Jörges and López-Moliner, 2017). In addition, (d) a life-detector must keep into account the visible changing relationship between a moving object and its environmental coordinates (movement direction, shape of the trajectory); it must be able (e) to detect the interaction between one moving object and other elements (e.g., avoiding an obstacle); it has to recognize (f) the interaction between two or more moving objects (chasing, approaching, or other social relationship).

Spatiotemporal contingencies (g) are another crucial cue in modulating social and psychological meaningful patterns, and in distinguishing causal from casual interactions. In other words, movements between agents have to look functionally related, that is, the changes of the one must appear as directly dependent on the changes of the other, at least within a specific range of variations (Michotte, 1946/1963; Gao et al., 2009). Very interesting in this regard is the recent work of Lemarie and colleagues (2022), who have emphasized the significant role of a certain degree of unpredictability in the temporal coincidences between interacting moving stimuli in domestic chicks. Animate agents, the authors argue, might require imperfect spatiotemporal contingencies between interactive moving objects – differently to launching events – and might avoid the perception of 'repetitive' or 'mechanical' movements in social aggregation stimuli. Similarly, the irregularity and unpredictability of individual trajectories can be understood as lifelike information that violates Newtonian motion (Mandler, 2012).

As a suggestion, the innate or at least predetermined sensitivity to animacy and life-like movements could be understood as a principle of saliency of the *ongoing changes within a multifactorial relational system,* including variations in relative speed, directions, and/or relative distance. From this perspective, for example, the natural behavior of physically constrained, form-invariant objects, or even mutually independent moving agents, could be seen as a frame of reference, a *neutral level from which non-inertial living forces or social configurations deviate.* In this way, animacy-related events are hardly predictable and become more visually salient, thus attracting the viewer's attention. If this framework is appropriate, then a plausible working hypothesis would be that the more the kinematic changes are sudden, unpredictable, and incongruent with their neutral frame of reference, the more they will evoke impressions of animacy and agency (Tremoulet and Feldman, 2000; Lemaire et al., 2022).

More work is needed to further explore the empirical and theoretical plausibility of a life-detector rooted in the visual system and capable of capturing and integrating all of these relationships, as well as the interaction between perceptual animacy and the acquired knowledge about causal mechanisms (Schlottmann et al., 2006; Simion et al., 2011). Also, the relationship between appearance-based properties of animated objects and pure kinematics animacy-related constraints needs to be better explored in future research (Long et al., 2018; Spriet et al., 2022). For example, it is still unknown how kinematic-based animacy interacts with face-like invariants, and how both interact with other appearance-based animacy features, such as human and animal bodies. Furthermore, if lifelike kinematics easily and automatically induces the perception of animacy in newborns babies, might not the *absence of motion* in a static image itself act as a suppressing factor in the detection of animacy?

## Conclusion

As outlined in this review, the perceptual system seems to be extremely sensitive to the entire range of information conveyed by movement variations and interactions concerning living entities. From a phenomenological point of view, these kinetic configurations are widely evaluated as vivid and meaningful, in an unreflective way, independently from any prior knowledge about the objective nature of the stimuli or any inferential reasoning (Schlottmann and Shanks, 1992). The observer cannot help but see animacy and intentionality as attributes of the objects, even though he knows that it is not the case; the evident inert nature of the moving objects does not hamper this evocative, and quite powerful mechanism, as already noted by Michotte (1950/1991).

Additionally, with respect to the animacy-related visual properties - such as faces or bodies - visual sensitivity to movements brings the relevant advantage of detecting the presence of a living organism even when the visibility conditions are deficient, for instance when the moving object is far, dimly lit, out of focus, partially occluded or camouflaged, or simply unseen before. Even in these conditions, the simple kinetic structure of an event enable us to perceive the nature of the animate or of the social situation.

Many experimental findings have shown that animacy and intentional relationships between moving objects are extracted rapidly and automatically, are sensitive to subtle visual parameters, appear early in development, are present in non-human species and can interact with low level properties - such as visual performance, visual memory, and speed perception.

Moreover, animacy impressions elicited by kinematics and appearance-based animacy features appear dissociate and partially independent from each other (Simion et al., 2011). Observers such as human and non-human newborns - even other vertebrates such as chicks -, are sensitive and pay more attention to lifelike moving objects or patterns (i.e., point-light displays) than to inanimate events, regardless of form. In contrast, it appears that the development of appearance-based visual features associated with animate beings, with the exception of face-like invariants, is not present at birth and requires a period of learning.

In adulthood, this independency between kinematic constraints and appearance-based features still persists and allows a quite interesting side effect, as events might appear ambiguous but also aesthetically rich. In some everyday situations, we can see lifeless objects mimicking living creatures through lifelike movements. Thanks to the autonomy of the impressions of animacy induced by the pure kinematics, many natural events appear to be vitalized and "stop-motion" movies can animate and psychologize geometric shapes and other non-anthropomorphic objects. Actually, incongruity and paradoxicality are important ingredients of visual comedy Bergson, (1900/2013). For example, in Walt Disney's classic movies (e.g., *Steamboat Willie,* 1928; *Fantasia,* 1940), as well as in many of Norman McLaren's shorts (such as *A Chairy Tale,* 1957), co-directed by Norman McLaren and Claude Jutra for the National Film Board of Canada or even in many advertisements, animated agents behave like inanimate ones, and vice versa (see Thomas and Johnston, 1981; Ejzenstejn, 2004). Perhaps the fascinating character of these seemingly alive moving forms lies in this empirical evidence.

These elements support the hypothesis that we can visually shape high-level properties and that the visual system can directly perceive meaningful relationships and interactions between objects. In addition to the distinction between animate and inanimate, a general sensitivity to the ongoing changes in a multiple relational system from which non-inertial living forces or social configurations deviate has been proposed. This would make it possible to rapidly identify psychological, emotional, and social characteristics of lifelike kinematics.

In sum, kinematics appears to be a crucial cue of animacy and agency, even independently from other appearance-based properties. Living-like shaped visual objects can look alive only in virtue of their motion (at least that of breathing), vice versa if still they look dramatically life-less. On the contrary, inanimate life-like moving objects, even in contrast with their other visual features, can appear paradoxically alive.

By way of conclusion, the topic of animacy is rooted in Michotte's experimental phenomenology, that is systematic psychophysical manipulation of stimuli configurations combined with subjective reports (see Costall, 2014; Bianchi and Davies, 2019; Parovel, 2019), and it is triggering a growing corpus of research cutting across several disciplines, including visual perception, developmental psychology, animal cognition, social psychology, cognitive neurosciences and robotics. Animacy, thus, besides being fascinating in itself, represents also a fruitful and challenging subject for empirical intersections and theoretical dialogue among different areas in experimental psychology. These dimensions of the visual scene, such as other expressive qualities of events that are still awaiting to be discovered, should be recognized and explored in all their richness and complexity within a multidisciplinary approach to human perception.

## Author contributions

GP conceived and wrote the entire manuscript.

## Funding

This publication was co-financed by POR CREO FESR Toscana 2014-2020 azione 1.1.5 sub-azione a1 – Bando 1 “Progetti Strategici di ricerca e sviluppo”: ATLAS.

## Acknowledgments

My thanks to the reviewers for their helpful and meaningful comments on the improvement of this paper. Thanks to Michele Vicovaro and Stefano Guidi for their reading of the manuscript and providing valuable observations.

## 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.

## Publisher’s note

## References

Abdai, J., Ferdinandy, B., Terencio, C. B., Pogàny, A., and Miklòsi, A. (2017). Perception of animacy in dogs and humans. *Biol. Lett.* 13:20170156. doi: 10.1098/rsbl.2017.0156

Adolphs, R. (2003). Investigating the cognitive neuroscience of social behavior. *Neuropsychologia* 41, 119–126. doi: 10.1016/s0028-3932(02)00142-2

Altman, M. N., Khislavsky, A. L., Coverdale, M. E., and Gilger, J. W. (2016). Adaptive attention: how preference for animacy impacts change detection. *Evol. Hum. Behav.* 37, 303–314. doi: 10.1016/j.evolhumbehav.2016.01.006

Bailey, R. L., and Lang, A. (2022). The importance of being animate: information selection as a function of dynamic human-environment interactions. *Front. Psychol.* 13:923808. doi: 10.3389/fpsyg.2022.923808

Baker, C. L., Saxe, R., and Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition* 113, 329–349. doi: 10.1016/j.cognition.2009.07.005

Bardi, L., Regolin, L., and Simion, F. (2011). Biological motion preference in humans at birth: role of dynamic and configural properties. *Dev. Sci.* 14, 353–359. doi: 10.1111/j.1467-7687.2010.00985.x

Bassili, J. N. (1976). Temporal and spatial contingencies in the perception of social events. *J. Pers. Soc. Psychol.* 33, 680–685. doi: 10.1037/0022-3514.33.6.680

Beasley, N. (1968). The extent of individual differences in the perception of causality. *Can. J. Psychol.* 22, 399–407. doi: 10.1037/h0082779

Bechlivanidis, C., and Lagnado, D. A. (2016). Time reordered: causal perception guides the interpretation of temporal order. *Cognition* 146, 58–66. doi: 10.1016/j.cognition.2015.09.001

Bergson, H. (1900/2013). *Laughter: An Essay on the Meaning of the Comic*. London: Macmillan Publications.

Bianchi, I., and Davies, R. (Eds.) (2019). *Paolo Bozzi’s experimental phenomenology*. London: Routledge.

Bingham, G. P., Schmidt, R. C., and Rosenblum, L. D. (1995). Dynamics and the orientation of kinematic forms in visual event recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1473–1493. doi: 10.1037//0096-1523.21.6.147

Biro, S., Csibra, G., and Gergely, G. (2007). The role of behavioral cues in understanding goal-directed actions in infancy. *Prog. Brain Res.* 164, 303–322. doi: 10.1016/S0079-6123(07)64017-5

Blakemore, S. J., Boyer, P., Pachot-Clouard, M., Meltzoff, A., Segebarth, C., and Decety, J. (2003). The detection of contingency and animacy from simple animations in the human brain. *Cereb. Cortex* 13, 837–844. doi: 10.1093/cercor/13.8.837

Blythe, P. W., Miller, G. F., and Todd, P. M. (1996). “Human simulation of adaptive behavior: interactive studies of pursuit, evasion, courtship, fighting, and play” in *From animals to animats 4: proceedings of the fourth international conference on simulation of adaptive behavior*. eds. P. Maes, M. J. Mataric, J.-A. Meyer, J. Pollack and S. W. Wilson (Cambridge, MA: MIT press/Bradford books), 13–22.

Blythe, P., Miller, G. F., and Todd, P. M. (1999). “How motion reveals intention: categorizing social interactions” in *simple heuristics that make us smart*. eds. G. Gigerenzer and P. M. Todd, and A. B. C. Research group (New York, NY: Oxford University Press), 257–285.

Bressanelli, D., and Parovel, G. (2012). The emotional effects of violations of causality or how to make a square amusing. *I-Perception* 3, 146–149. doi: 10.1068/i0465sas

Buehner, M. J., and Humphreys, G. R. (2010). Causal contraction: spatial binding in the perception of collision events. *Psychol. Sci.* 21, 44–48. doi: 10.1177/0956797609354735

Butterfill, S. (2009). Seeing causes and hearing gestures. *Philos. Q.* 59, 405–428. doi: 10.1111/j.1467-9213.2008.585.x

Castelli, F., Happé, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuro Image* 12, 314–325. doi: 10.1006/nimg.2000.0612

Choi, H., and Scholl, B. J. (2006). Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. *Perception* 35, 385–399. doi: 10.1068/p5462

Costall, A. (2014). “Phenomenal causality” in *Michotte’s experimental phenomenology of perception*. eds. G. Thinès, A. Costall and G. Butterworth (Hillsdale, NJ: Lawrence Erlbaum Associates), 51–116.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107, 705–717. doi: 10.1016/j.cognition.2007.08.001

Dasser, V., Ulbaek, I., and Premack, D. (1989). The perception of intention. *Science* 243, 365–367. doi: 10.1126/science.2911746

De Sá Teixeira, N. A., De Oliveira, A. M., and Viegas, R. (2008). Functional approach to the integration of kinematic and dynamic variables in causal perception: is there a link between phenomenology and behavioral responses? *Jpn. Psychol. Res.* 50, 232–241. doi: 10.1111/j.1468-5884.2008.00379.x

Deen, B., Richardson, H., Dilks, D., Takahashi, A., Keil, B., Wald, L. L., et al. (2017). Organization of high-level visual cortex in human infants. *Nat. Commun.* 8:13995. doi: 10.1038/ncomms13995

Di Giorgio, E., Lunghi, M., Simion, F., and Vallortigara, G. (2016). Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Dev. Sci.* 20, e12394–e12312. doi: 10.1111/desc.12394

Di Giorgio, E., Lunghi, M., Vallortigara, G., and Simion, F. (2021). Newborns’ sensitivity to speed changes as a building block for animacy perception. *Sci. Rep.* 11:542 (2021). doi: 10.1038/s41598-020-79451-3

Ding, X., Gao, Z., and Shen, M. (2017). Two equals one: two human actions during social interaction are grouped as one unit in working memory. *Psychol. Sci.* 28, 1311–1320. doi: 10.1177/0956797617707318

Dittrich, W. H., and Lea, S. E. G. (1994). Visual perception of intentional motion. *Perception* 23, 253–268. doi: 10.1068/p230253

Duncker, K. (1935/1969). *Psicologia del pensiero produttivo [psychology of productive thinking]*. Firenze: Giunti.

Ejzenstejn, S. M. (2004). *Walt Disney*. Milano: SE.

Falmier, O., and Young, M. E. (2008). The impact of object animacy on the appraisal of causality. *Am. J. Psychol.* 121, 473–500. doi: 10.2307/20445477

Fedorov, L. A., Chang, D. S., Giese, M. A., Bülthoff, H. H., and de la Rosa, S. (2018). Adaptation aftereffects reveal representations for encoding of contingent social actions. *Proc. Natl. Acad. Sci.* 115, 7515–7520. doi: 10.1073/pnas.1801364115

Firestone, C., and Scholl, B. J. (2015). When do ratings implicate perception versus judgment? The “overgeneralization test” for top-down effects. *Vis. Cogn.* 23, 1217–1226. doi: 10.1080/13506285.2016.1160171?journalCode=pvis20

Fox, R., and McDaniel, C. (1982). The perception of biological motion by human infants. *Science* 218, 486–487. doi: 10.1126/science.7123249

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8, 396–403. doi: 10.1016/j.tics.2004.07.002

Gao, T., McCarthy, G., and Scholl, B. J. (2010). The Wolfpack effect: perception of animacy irresistibly influences interactive behavior. *Psychol. Sci.* 21, 1845–1853. doi: 10.1177/0956797610388814

Gao, T., Newman, G. E., and Scholl, B. J. (2009). The psychophysics of chasing: a case study in the perception of animacy. *Cogn. Psychol.* 59, 154–179. doi: 10.1016/j.cogpsych.2009.03.001

Gao, T., and Scholl, B. J. (2011). Chasing vs. stalking: interrupting the perception of animacy. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 669–684. doi: 10.1037/a0020735

Gao, T., Scholl, B. J., and McCarthy, G. (2012). Dissociating the detection of intentionality from animacy in the right posterior superior temporal sulcus. *J. Neurosci.* 32, 14276–14280. doi: 10.1523/JNEUROSCI.0562-12.2012

Gelman, R., Durgin, F., and Kaufman, L. (1995). "Distinguishing between animates and inanimates: not by motion alone" in *Causal cognition: A multidisciplinary debate*. eds. D. Sperber, D. Premack and A. J. Premack (New York, NY: Oxford University Press), 151–184.

Gemelli, A., and Cappellini, A. (1958). The influence of the subject's attitude in perception. *Acta Psychol.* 14, 12–23. doi: 10.1016/0001-6918(58)90003-9

Gergely, G., and Csibra, G. (1997). Teleological reasoning in infancy: the infant's naive theory of rational action. A reply to Premack and Premack. *Cognition* 63, 227–33. doi: 10.1016/s0010-0277(97)00004-8

Gergely, G., and Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends Cogn. Sci.* 7, 287–292. doi: 10.1016/S1364-6613(03)00128-1

Gergely, G. N. Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193. doi: 10.1016/0010-0277(95)00661-h

Gibson, J. J. (1977). "The theory of affordances" in *Perceiving, acting, and knowing: toward an ecological psychology*. eds. R. Shaw and J. Bransford (Hillsdale, NJ: Lawrence Erlbaum), 67–82.

Gyulai, E. (2000). Attribution of meaning from movement of simple objects. *Percept. Mot. Skills* 90, 27–35. doi: 10.2466/pms.2000.90.1.27

Hafri, A., and Firestone, C. (2021). The perception of relations. *Trends Cogn. Sci.* 25, 475–492. doi: 10.1016/j.tics.2021.01.006

Hafri, A., Trueswell, J. C., and Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition* 175, 36–52. doi: 10.1016/j.cognition.2018.02.011

Heberlein, A. (2008). "Animacy and intention in the brain: neuroscience of social event perception, in understanding events" in *From perception to action*. eds. T. F. Shipley and J. M. Zacks (Oxford: Oxford University Press), 363–388.

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259. doi: 10.2307/1416950

Hofrichter, R., Mueller, M. E., and Rutherford, M. D. (2021). Children's perception of animacy: social attributions to moving figures. *Perception* 50, 387–398. doi: 10.1177/03010066211010142

Hubbard, T. L. (2013). Phenomenal causality I: varieties and variables. *Axiomathes* 23, 1–42. doi: 10.1007/s10516-012-9198-8

Hubbard, T. L., Blessum, J. A., and Ruppel, S. E. (2001). Representational momentum and Michotte's "launching effect" paradigm (1946/1963). *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 294–301. doi: 10.1037/0278-7393.27.1.294

Hubbard, T. L., and Ruppel, S. E. (2002). A possible role of naive impetus in Michotte's "launching effect" evidence from representational momentum. *Vis. Cogn.* 9, 153–176. doi: 10.1080/13506280143000377

Ichikawa, H., Kanazawa, S., and Yamaguchi, M. K. (2011). Finding a face in a face-like object. *Perception* 40, 500–502. doi: 10.1068/p6926

Jörges, B., and López-Moliner, J. (2017). Gravity as a strong prior: implications for perception and action. *Front. Hum. Neurosci.* 11, 1–16. doi: 10.3389/fnhum.2017.00203

Kanizsa, G., and Vicario, G. (1968). "The perception of intentional reaction" in *Experimental research in perception*. eds. G. Kanizsa and G. Vicario (Trieste, Italy: University of Trieste), 71–126.

Kim, S.-H., Feldman, J., and Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychol. Sci.* 24, 575–582. doi: 10.1177/0956797612458529

Kominsky, J. F., Shafto, P., and Bonawitz, E. (2021). "There's something inside": Children's intuitions about animate agents. *PLoS One* 16:e0251081. doi: 10.1371/journal.pone.0251081

Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., and Keil, F. C. (2017). Categories and constraints in causal perception. *Psychol. Sci.* 28, 1649–1662. doi: 10.1177/0956797617719930

Kuhlmeier, V., Wynn, K., and Bloom, P. (2003). Attribution of dispositional states by 12-month-old infants. *Psychol. Sci.* 14, 402–408. doi: 10.1111/1467-9280.01454

Lemaire, B., Rosa-Salva, O., Fraja, M., Lorenzi, E., and Vallortigara, G. (2022). Spontaneous preference for unpredictability in the temporal contingencies between agents' motion in naïve domestic chicks. *Proc. R. Soc. Lond. B* 289:20221622. doi: 10.1098/rspb.2022.1622

Lemaire, B. S., and Vallortigara, G. (2023). Life is in motion (through a chick's eye). *Anim. Cogn.* 26, 129–140. doi: 10.1007/s10071-022-01703-8

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception* 13, 287–305. doi: 10.1068/p130287

Leslie, A. M. (1994). "Tomm, toby and agency: Core architecture and domain specificity" in *Mapping the mind*. eds. L. A. Hirschfeld, S. A. Gelman (Cambridge: Cambridge University Press), 119–148.

Leslie, A. M., and Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition* 25, 265–288. doi: 10.1016/S0010-0277(87)80006-9

Long, B., Yu, C.-P., and Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc. Natl. Acad. Sci. U. S. A.* 115, E9015–E9024. doi: 10.1073/pnas.1719616115

Lorenzi, E., Mayer, U., Rosa-Salva, O., and Vallortigara, G. (2017). Dynamic features of animate motion activate septal and preoptic areas in visually naïve chicks (*Gallus gallus*). *Neuroscience* 354, 54–68. doi: 10.1016/j.neuroscience.2017.04.022

Lorenzi, E., and Vallortigara, G. (2021). "Evolutionary and neural bases of the sense of animacy" in *The Cambridge handbook of animal cognition*. eds. A. B. Kaufman, J. Call and J. C. Kaufman ( Cambridge University Press), 295–321.

Macchi Cassia, V., Turati, C., and Simion, F. (2004). Can a non- specific bias toward top-heavy patterns explain newborns' face preference? *Psychol. Sci.* 15, 379–383. doi: 10.1111/j.0956-7976.2004.00688.x

Mandler, J. M. (1992). How to build a baby: II Conceptual primitives. *Psychol. Rev.* 99, 587–604. doi: 10.1037/0033-295X.99.4.587

Mandler, J. M. (2003). "Conceptual categorization" in *Early category and concept development*. eds. D. Rakison and L. M. Oakes (Oxford: Oxford University Press), 103–131.

Mandler, J. (2012). Prelinguistic primitives. *Ann. Meet. Berkeley Linguist. Soc.* 17:414. doi: 10.3765/bls.v17i0.1623

Mascalzoni, E., Regolin, L., and Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proceed. Natl. Acad. Sci. U. S. A.* 107, 4483–4485. doi: 10.1073/pnas.0908792107

Michotte, A. (1946/1963). *The perception of causality*. T. Miles & E. Miles, Trans.. London: Methuen.

Michotte, A. (1950/1991). "The emotions regarded as functional connections" in *Feelings and emotions: the Mooseheart symposium*. ed. M. Reymert (New York: McGraw-Hill), 114–125. Reprinted in Michotte's experimental phenomenology of perception, ed. G. Thinès, A. Costall, and G. Butterworth (Hillsdale, NJ: Erlbaum, 1991), 103–116.

Minguzzi, G. F. (1961). Caratteri espressivi ed intenzionali dei movimenti: "la percezione dell'attesa". *Riv. Psicol.* 55, 157–173.

Moors, P., Wagemans, J., and de-Wit, L. (2017). Causal events enter awareness faster than non-causalevents. *PeerJ.* 5:e2932. doi: 10.7717/peerj.2932

Morales-Bader, D., Castillo, R. D., Olivares, C., and Miño, F. (2020). How do object shape, semantic cues, and apparent velocity affect the attribution of intentionality to figures with different types of movements? *Front. Psychol.* 11:935. doi: 10.3389/fpsyg.2020.00935

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant Child Dev.* 10, 3–18. doi: 10.1017/s1092852900022872

New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceed. Natl. Acad. Sci. U. S. A.* 104, 16598–16603. doi: 10.1073/pnas.0703913104

Nguyen, H. B., and van Buren, B. (2023). May the force be against you: better visual sensitivity to speed changes opposite to gravity. *J. Exp. Psychol. Hum. Percept. Perform.* doi: 10.1037/xhp0001115

Opfer, J. E. (2002). Identifying living and sentient kinds from dynamic information: the case of goal-directed versus aimless autonomous movement in conceptual change. *Cognition* 86, 97–122. doi: 10.1016/S0010-0277(02)00171-3

Parovel, G. (2019). Paolo Bozzi's experimental phenomenology. *Gestalt Theory* 41, 79–84. doi: 10.2478/gth-2019-0006

Parovel, G., and Casco, C. (2006). The psychophysical law of speed estimation in Michotte's causal events. *Vis. Res.* 46, 4134–4142. doi: 10.1016/j.visres.2006.08.005

Parovel, G., and Guidi, S. (2015). The psychophysics of comic: effects of incongruity in causality and animacy. *Acta Psychol.* 159, 22–32. doi: 10.1016/j.actpsy.2015.05.002

Parovel, G., and Guidi, S. (2020). Speed overestimation of the moving away object in the intentional reaction causal effect. *i-Perception* 11:204166952098001. doi: 10.1177/2041669520980019

Parovel, G., Brunello, L., and Vicovaro, M. (2022). Visual perception of bouncing and jumping. In Perception supplement, 44th European Conference of Visual Perception 51: 82.

Parovel, G., Guidi, S., and Kreß, K. (2018). Different contexts change the impression of animacy. *Atten. Percept. Psychophys.* 80, 553–563. doi: 10.3758/s13414-017-1439-x

Pauen, S., and Träuble, B. (2009). How 7-month-olds interpret ambiguous motion events: category-based reasoning in infancy. *Cogn. Psychol.* 59, 275–295. doi: 10.1016/j.cogpsych.2009.06.001

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition* 36, 1–16. doi: 10.1016/0010-0277(90)90051-k

Premack, D., and Premack, A. J. (1995). "Origins of human social competence" in *The cognitive neurosciences*. ed. M. S. Gazzaniga (Cambridge, MA, US: The MIT Press), 205–218.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Brain Sci.* 22, 341–365. doi: 10.1017/s0140525x99002022

Regolin, L., Tommasi, L., and Vallortigara, G. (2000). Visual perception of biological motion in newly hatched chicks as revealed by an imprinting procedure. *Anim. Cogn.* 3, 53–60. doi: 10.1007/s100710050050

Rips, L. (2011). Causation from perception. *Perspect. Psychol. Sci.* 6, 77–97. doi: 10.1177/1745691610393525

Rolfs, M., Dambacher, M., and Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Curr. Biol.* 23, 250–254. doi: 10.1016/j.cub.2012.12.017

Rosa Salva, O., Mayer, U., and Vallortigara, G. (2015). Roots of a social brain: developmental models of emerging animacy-detection mechanisms. *Neurosci. Biobehav. Rev.* 50, 150–168. doi: 10.1016/j.neubiorev.2014.12.015

Rosa-Salva, O., Grassi, M., Lorenzi, E., Regolin, L., and Vallortigara, G. (2016). Spontaneous preference for visual cues of animacy in naïve domestic chicks: the case of speed changes. *Cognition* 157, 49–60. doi: 10.1016/j.cognition.2016.08.014

Rosa-Salva, O., Hernik, M., Broseghini, A., and Vallortigara, G. (2018). Visually-naïve chicks prefer agents that move as if constrained by a bilateral body-plan. *Cognition* 173, 106–114. doi: 10.1016/j.cognition.2018.01.004

Runeson, S., and Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *J. Exp. Psychol. Gen.* 112, 585–615. doi: 10.1037/0096-3445.112.4.585

Rutherford, M. D. (2013). "Evidence for specialized perception of animate motion" in *Social perception: Detection and interpretation of animacy, agency, and intention*. eds. M. D. Rutherford and V. A. Kuhlmeier (Cambridge, MA: MIT Press), 115–138.

Santos, N. S., David, N., Bente, G., and Vogeley, K. (2008). Parametric induction of animacy experience. *Conscious. Cogn.* 17, 425–437. doi: 10.1016/j.concog.2008.03.012

Santos, N. S., Kuzmanovic, B., David, N., Rotarska-Jagiela, A., Eickhoff, S. B., Shah, J. N., et al. (2010). Animated brain: a functional neuroimaging study on animacy experience. *NeuroImage* 53, 291–302. doi: 10.1016/j.neuroimage.2010.05.080

Saxe, R., and Carey, S. (2006). The perception of causality in infancy. *Acta Psychol.* 123, 144–165. doi: 10.1016/j.actpsy.2006.05.005

Schlottmann, A. (2001). Perception versus knowledge of cause and effect in children: When seeing is believing. *Current Directions in Psychological Science* 10, 111–115. doi: 10.1111/1467-8721.00128

Schlottmann, A., Allen, D., Linderoth, C., and Hesketh, S. (2002). Perceptual causality in children. *Child Dev.* 73, 1656–1677. doi: 10.1111/1467-8624.00497

Schlottmann, A., and Anderson, N. H. (1993). An information integration approach to phenomenal causality. *Mem. Cognition* 21, 785–801. doi: 10.3758/BF03202746

Schlottmann, A., Cole, K., Watts, R., and White, M. (2013). Domain-specific perceptual causality in children depends on the spatio-temporal configuration, not motion onset. *Front. Psychol.* 4:365. doi: 10.3389/fpsyg.2013.00365

Schlottmann, A., and Ray, E. (2010). Goal attribution to schematic animals: do 6-month-olds perceive biological motion as animate? *Dev. Sci.* 13, 1–10. doi: 10.1111/j.1467-7687.2009.00854.x

Schlottmann, A., Ray, E., Mitchell, A., and Demetriou, N. (2006). Perceived physical and social causality in animated motions: spontaneous reports and ratings. *Acta Psychol.* 123, 112–143. doi: 10.1016/j.actpsy.2006.05.006

Schlottmann, A., Ray, E., and Surian, L. (2012). Emerging perception of causality in action-and-reaction sequences from 4 to 6months of age: is it domain-specific? *J. Exp. Child Psychol.* 112, 208–230. doi: 10.1016/j.jecp.2011.10.011

Schlottmann, A., and Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *Q. J. Exp. Psychol.* 44, 321–342. doi: 10.1080/02724989243000055

Schlottmann, A., and Surian, L. (1999). Do 9-month-olds perceive causation-at-a-distance? *Perception* 28, 1105–1113. doi: 10.1068/p281105

Scholl, B. J., and Gao, T. (2013). "Perceiving animacy and intentionality: visual processing or higher-level judgment?" in *Social perception: detection and interpretation of animacy, agency, and intention*. eds. M. D. Rutherford and V. A. Kuhlmeier ( MIT Press), 197–229.

Scholl, B. J., and Nakayama, K. (2004). Illusory causal crescents: misperceived spatial relations due to perceived causality. *Perception* 33, 455–469. doi: 10.1068/p5172

Scholl, B. J., and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309. doi: 10.1016/S1364-6613(00)01506-0

Schultz, J., and Bülthoff, H. H. (2013). Parametric animacy percept evoked by a single moving dot mimicking natural stimuli. *J. Vis.* 13:15. doi: 10.1167/13.4.15

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., and Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* 45, 625–635. doi: 10.1016/j.neuron.2004.12.052

Simion, F., Di Giorgio, E., Leo, I., and Bardi, L. (2011). The processing of social stimuli in early infancy: from faces to biological motion perception. *Prog. Brain Res.* 189, 173–193. doi: 10.1016/B978-0-444-53884-0.00024-5

Simion, F., Regolin, L., and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceed. Natl. Acad. Sci. U. S. A.* 105, 809–813. doi: 10.1073/pnas.0707021105

Simion, F., Valenza, E., Macchi Cassia, V., Turati, C., and Umiltà, C. A. (2002). Newborns' preference for up-down asymmetrical configurations. *Dev. Sci.* 5, 427–434. doi: 10.1111/1467-7687.00237

Slotnick, S. D. (2013). *Controversies in Cognitive Neuroscience*. Basingstoke, UK: Palgrave Macmillan

Spelke, E. S., and Kinzler, K. D. (2007). Core knowledge. *Dev. Sci.* 10, 89–96. doi: 10.1111/j.1467-7687.2007.00569.x

Spelke, E. S., Philips, A., and Woodward, A. L. (1995). "Infants' knowledge of object motion and human action" in *Causal cognition*. eds. D. Sperber, D. Premack and A. J. Premack (Oxford, UK: Clarendon), 150–184.

Spriet, C., Abassi, E., Hochmann, J. R., and Papeo, L. (2022). Visual object categorization in infancy. *Proc. Natl. Acad. Sci. U. S. A.* 119:e2105866119. doi: 10.1073/pnas.2105866119

Stewart, J. A. (1982). Perception of animacy. Unpublished doctoral dissertation. Philadelphia: University of Pennsylvania.

Stewart, J. A. (1984). Object motion and the perception of animacy. Paper presented at the meetings of the Psychonomic Society, San Antonio, TX.

Szego, P. A., and Rutherford, M. D. (2007). Actual and illusory differences in constant speed influence the perception of animacy similarly. *J. of Vision* 7, 5–5.7. doi: 10.1167/7.12.5

Szego, P. A., and Rutherford, M. D. (2008). Dissociating the perception of speed and the perception of animacy: a functional approach. *Evol. Hum. Behav.* 29, 335–342. doi: 10.1016/j.evolhumbehav.2008.04.002

Tenenbaum, J. B., and Griffiths, T. L. (2003). "Theory-based causal inference" in *Advances in neural information processing systems*. eds. S. Becker, S. Thrun and K. Obermayer (Cambridge: MIT Press), 35–42.

Thinès, G., Costall, A., and Butterworth, G. (2014). *Michotte's experimental phenomenology of perception*, London: Routledge

Thomas, F., and Johnston, O. (1981). *The illusion of life: disney animation*. New York: Disney Editions.

Toribio, J. (2018). Visual experience: rich but impenetrable. *Synthese* 195, 3389–3406. doi: 10.1007/s11229-015-0889-8

Träuble, B., Pauen, P., and Poulin-Dubois, D. (2014). Speed and direction changes induce the perception of animacy in 7-month-old infants. *Front. Dev. Psychol.* 5:1141. doi: 10.3389/fpsyg.2014.01141

Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–951. doi: 10.1068/p3101

Tremoulet, P. D., and Feldman, J. (2006). The influence of spatial context and the role of intentionality in the interpretation of animacy from motion. *Percept. Psychophys.* 68, 1047–1058. doi: 10.3758/bf03193364

Troje, N. F., and Westhoff, C. (2006). The inversion effect in biological motion perception: evidence for a "life detector". *Curr. Biol.* 16, 821–824. doi: 10.1016/j.cub.2006.03.022

Turati, C., Simion, F., Milani, I., and Umiltà, C. A. (2002). Newborns' preference for faces: what is crucial? *Dev. Psychol.* 38, 875–882. doi: 10.1037/0012-1649.38.6.875

Vallortigara, G. (2012). "Aristotle and the chicken: animacy and the origins of beliefs" in *The theory of evolution and its impact*. ed. A. Fasolo (New York: Springer), 189–200.

Vallortigara, G., and Regolin, L. (2006). Gravity bias in the interpretation of biological motion by inexperienced chicks. *Curr. Biol.* 16, R279–R280. doi: 10.1016/j.cub.2006.03.052

Vallortigara, G., Regolin, L., and Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biol.* 3, e208–e1316. doi: 10.1371/journal.pbio.0030208

van Buren, B., Gao, T., and Scholl, B. J. (2017). What are the underlying units of perceived animacy?: chasing detection is intrinsically object-based. *Psychon. Bull. Rev.* 24, 1604–1610. doi: 10.3758/s13423-017-1229-4

van Buren, B., and Scholl, B. J. (2017). Minds in motion in memory: enhanced spatial memory driven by the perceived animacy of simple shapes. *Cognition* 163, 87–92. doi: 10.1016/j.cognition.2017.02.006

van Buren, B., and Scholl, B. J. (2018). The 'Blindfold Test' for deciding whether an effect reflects visual processing or higher-level judgment. *J. Vis.* 18:56. doi: 10.1167/18.10.56

van Buren, B., Uddenberg, S., and Scholl, B. J. (2016). The automaticity of perceiving animacy: goal-directed motion in simple shapes influences visuomotor behavior even when task-irrelevant. *Psychon. Bull. Rev.* 23, 797–802. doi: 10.3758/s13423-015-0966-5

Vestner, T., Gray, K. L. H., and Cook, R. (2020). Why are social interactions found quickly in visual search tasks? *Cognition* 200:104270. doi: 10.1016/j.cognition.2020.104270

Vestner, T., Tipper, S. P., Hartley, T., over, H., and Rueschemeyer, S. A. (2019). Bound together: social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. *J. Exp. Psychol. Gen.* 148, 1251–1268. doi: 10.1037/xge0000545

Vicovaro, M. (2018). Causal reports: context-dependent contributions of intuitive physics and visual impressions of launching. *Acta Psychol.* 186, 133–144. doi: 10.1016/j.actpsy.2018.04.015

Vicovaro, M., Battaglini, L., and Parovel, G. (2020). The larger the cause, the larger the effect: evidence of speed judgment biases in causal scenarios. *Vis. Cogn.* 28, 239–255. doi: 10.1080/13506285.2020.1783041

Vicovaro, M., Battaglini, L., and Parovel, G. (2023). The psychophysics of bouncing: perceptual constraints, physical constraints, animacy, and phenomenal causality. *PLoS ONE* (in press). doi: 10.1371/journal.pone.0285448

Visch, V. T., and Tan, E. S. (2009). Categorizing moving objects into film genres: the effect of animacy attribution, emotional response, and the deviation from non-fiction. *Cognition* 110, 265–272. doi: 10.1016/j.cognition.2008.10.018

Wagemans, J., van Lier, R., and Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research. *Acta Psychol.* 123, 1–19. doi: 10.1016/j.actpsy.2006.06.003

Weir, S. (1978). The perception of motion: Michotte revisited. *Perception* 7, 247–260. doi: 10.1068/p070247

Wheatley, T., Milleville, S. C., and Martin, A. (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psychol. Sci.* 18, 469–474. doi: 10.1111/j.1467-9280.2007.01923.x

White, P. A. (2006). The role of activity in visual impressions of causality. *Acta Psychol.* 123, 166–185. doi: 10.1016/j.actpsy.2006.05.002

White, P. A., and Milne, A. (1997). Phenomenal causality: impressions of pulling in the visual perception of objects in motion. *Am. J. Psychol.* 110, 573–602. doi: 10.2307/1423411

Withagen, R. (2022). *Affective Gibsonian psychology*. New York: Routledge

Withagen, R., de Poel, H. G., Araújo, D., and Pepping, G.-J. (2012). Affordances can invite behavior: reconsidering the relationship between affordances and agency. *New Ideas Psychol.* 30, 250–258. doi: 10.1016/j.newideapsych.2011.12.003

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34. doi: 10.1016/s0010-0277(98)00058-4

Yoon, J. M., and Johnson, S. C. (2009). Biological motion displays elicit social behavior in 12-month-olds. *Child Dev.* 80, 1069–1075. doi: 10.1111/j.1467-8624.2009.01317.x

Young, M. E., and Falmier, O. (2008). Launching at a distance: the effect of spatial markers. *Q. J. Exp. Psychol.* 61, 1356–1370. doi: 10.1080/17470210701595522

Yuasa, M. (2017). The inference of friendly communicative atmosphere created by geometric shapes. *i-Perception*:8. doi: 10.1177/2041669517744571

Frontiers in Psychology

TYPE Original Research
PUBLISHED 05 June 2023
DOI 10.3389/fpsyg.2023.1145289

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University, United States

REVIEWED BY
Johannes Schultz,
University Hospital Bonn, Germany
Sara B. Felix,
University of Aveiro, Portugal
Aurélia Bugaiska,
Université de Bourgogne, France

*CORRESPONDENCE
Chris Westbury
✉ chrisw@ualberta.ca

RECEIVED 16 January 2023
ACCEPTED 05 April 2023
PUBLISHED 05 June 2023

CITATION
Westbury C (2023) Why are human animacy judgments continuous rather than categorical? A computational modeling approach.
*Front. Psychol.* 14:1145289.
doi: 10.3389/fpsyg.2023.1145289

# Why are human animacy judgments continuous rather than categorical? A computational modeling approach

Chris Westbury*

Department of Psychology, University of Alberta, Edmonton, AB, Canada

**Introduction:** The concept of animacy is often taken as a basic natural concept, in part I because most cases seem unambiguous. Most entities either are or are not animate. However, human animacy judgments do not reflect this binary classification. They suggest that there are borderline cases, such as *virus*, *amoeba*, *fly*, and imaginary beings (*giant*, *dragon*, *god*). Moreover, human roles (*professor, mother, girlfriend*) are consistently recognized as animate by far less than 100% of human judges.

**Method:** In this paper, I use computational modeling to identify features associated with human animacy judgments, modeling human animacy and living/non-living judgments using both bottom-up predictors (the principal components from a word embedding model) and top-down predictors (cosine distances from the names of animate categories).

**Results:** The results suggest that human animacy judgments may be relying on information obtained from imperfect estimates of category membership that are reflected in the word embedding models. Models using cosine distance from category names mirror human judgments in distinguishing strongly between humans (estimated lower animacy by the measure) and other animals (estimated higher animacy by the measure).

**Discussion:** These results are consistent with a family resemblance approach to the apparently categorical concept of animacy.

KEYWORDS

**animacy, word embedding (word2vec), computational modeling methods, human judgment, taxonomy, classification**

## 1. Introduction

The word animacy is defined in the Oxford English Dictionary (2022) as "The quality or condition of being alive or animate; animate existence; an instance of this." This definition seems clear and unambiguous on its surface. However, when humans are asked to make judgments of animacy, they identify many intermediate or anomalous cases. The goal of the present paper is to use computational modeling to shed light on the lack of unanimous binary animacy decisions by English speakers for many words, by modeling the decisions for the 72 words rated for animacy in Radanović et al. (2016) and for 1,200 English words rated living/non-living from VanArsdall and Blunt (2022). I will consider two models with different set of predictors and synthesize their contributions to the understanding human animacy judgments at the end, by considering whether and why the models make the same kinds of errors that humans do.

As examples of the lack of agreement in animacy ratings, Radanović et al. (2016) reported that their university-student judges rated the animacy of giraffes or babies at about 50 (out of

100, where 0 = inanimate and 100 = animate), though we would normally think of the default state of these entities as living. This is approximately the same as the average ratings for balls (49.2) or snow (51.0), though we would not think of these entities as being alive. Other intermediate cases include imaginary beings such as ghosts (rated 41.7) and fairies (49.4); entities that imitate animate entities such as computers (52.2) and robots (33.1); and simple creatures such as amoebae (83.5) and viruses (69.4). Plants are a potentially ambiguous intermediate case, since they are animate by the Oxford English Dictionary's definition, but we often interact with them in inanimate form. This may explain the lack of strong consensus about animacy in the ratings of words referring to plants such as *cabbage* (59.0), *tomato* (38.9), and *orchid* (59.0).

Languages that mark animacy grammatically can add additional complications within specific cultures. For example, in Cree, animal hides, trees (but not pieces of wood), and some (but not all) stones are marked grammatically as animate, perhaps (as suggested by Darnell and Vanek, 1976) reflecting that in Cree "a thing is classified as animate if it has power" (p. 164).

The role of animacy in semantic and lexical processing has been the focus of many studies (e.g., Cappa et al., 1998; Caramazza and Shelton, 1998; Grabowski et al., 1998; Mummery et al., 1998; Moore and Price, 1999; Tyler et al., 2000; Tyler and Moss, 2001; Radanović et al., 2016). Some studies have reported behavioral and/or neurological differences in response to animate and inanimate stimuli (Perani et al., 1995; Martin et al., 1996; Perani et al., 1999). Other studies have failed to replicate these findings (Devlin et al., 2002; Pilgrim et al., 2002; Tyler et al., 2003; Ilić et al., 2013). The linguistic encoding of animacy has been shown to affect many different aspects of psychological functioning, including the processing of relative clauses (Mak et al., 2002; Traxler et al., 2005; Gennari et al., 2012); attentional mechanisms (Bugaiska et al., 2019); the detection of semantic violations in language (Grewe et al., 2006; Szewczyk and Schriefers, 2011); the learning of artificial languages (Vihman et al., 2018); word recognition (Bonin et al., 2019) and the ability to recall words (Bonin et al., 2015; VanArsdall et al., 2015; Bugaiska et al., 2016; Popp and Serra, 2016, 2018; Nairne et al., 2017; Kazanas et al., 2020).

As Radanović et al. (2016) noted, one complication in studies using animacy is how stimuli are selected. Some studies have focused on only a few exemplars (i.e., tools versus animals, as in Perani et al., 1995, 1999; Martin et al., 1996). Others including a wider range of animate and inanimate stimuli.

Animacy ratings have been gathered in many languages (e.g., Serbian/English: Radanović et al., 2016; Portuguese: Félix et al., 2020; Persian: Mahjoubnavaz and Mokhtari, 2022; English: VanArsdall and Blunt, 2022). This study focuses on the two sets of English ratings in this list.

The first set was the set of 72 ratings from Radanović et al. (2016). As noted above, these were rated from 1 (inanimate) to 100 (animate). The authors reported that the English ratings were strongly correlated with independent Serbian ratings of the same words ($r = 0.89$, $p < 0.001$). They included a wide range of words. The ratings are summarized by into categories in Table 1. There is notable variation in ratings within categories of animate things. Contrary to the some claims (see discussion in Radanović et al., 2016, p. 17) human beings are rated as lower in animacy (Average [SD] rating: 60.0 [16.6]) than other animals (Average [SD] rating: 79.8 [22.6]; $t(14.39) = 2.18$; $p = 0.046$). Since human beings are certainly animate, this result is puzzling. I will consider it again in the conclusion section of this paper.

The other set is the recently released set of ratings from VanArsdall and Blunt (2022). They gathered living/non-living ratings from 1 to 7 for 1,200 English words. Each word was rated a minimum of 19 times (average [SD]: 25 [1.62]). The ratings are also summarized by category in Table 1.

These two sets of ratings are along slightly different dimensions. Some things that are clearly non-living (for example, unicorns and Santa Claus) might reasonably be judged animate. However, the larger set of ratings makes it possible to cross-validate the models, which is not possible with the small number of ratings from Radanović et al. (2016). Moreover, the ratings are correlated. The 50 words that appear in both data sets have animacy and living/non-living ratings that correlate at $r = 0.60$ ($p < 0.0001$).

The models use two different sources of data, to allow us to consider the issue from both a bottom-up perspective (to what degree is animacy encoded in semantics/patterns of language use?) and a top-down perspective (to what degree is animacy determined by membership in categories of animate entities?). One model uses the principal components of vector representations of words from a word embedding model (explained in more detail in the next section) to try to predict human ratings. This can give us an idea of the extent which animacy is encoded into language use, a bottom-up approach to animacy. The second model uses the similarity of a word's vector to the vector of the names of definitely animate categories such as *human*, *animal*, and *plant*. This can give us an idea of the extent to which animacy is derivable from the goodness of its categorical membership. For example, though they are animals, humans are generally considered to be poor representatives of that class. It is possible that this is why humans are less likely to be judged as animate than other animals.

## 2. Model 1: introduction

The first model uses generalized additive models (GAMs) across the principal components (PCs) from a word-embedding model to predict human judgments. GAMs are able to capture non-linear relationships between predictors and a dependent measure but can also find linear relationships when they are the best fit for the data.

Word-embedding models are computational models that build vector representations of individual words that represent the average context in which that word appears in a large corpus of language. Perhaps the simplest way to do this is that used in the earliest model, Landauer and Dumais's (1997) Latent Semantic Analysis (LSA), which built a *word x document matrix* in which the individual cells (prior to processing the matrix with singular value decomposition to reduce the dimensionality) recorded how often each word (rows) appeared in each document (columns). Since documents almost always have a semantic focus (they are usually *about* something), we might reasonably expect that words whose untransformed vectors were similar (say, vectors for the words *pet* and *cat*) are words that have similar semantics. Importantly, LSA does not directly measure whether *cat* and *pet* occur together in the same documents, which is what we call *first-order co-occurrence*. It measures whether the documents in which *cat* and *pet* appeared tended to contain the same words (a comparison of *word context* that we call *second-order*

TABLE 1 Animacy ratings from Radanović et al. (2016), living/non-living ratings from VanArsdall and Blunt (2022), and model estimates of the latter, by category, ordered from most animate to least animate by human rating.

| | Radanović et al. | | | VanArsdall and Blunt | | Fitted |
|---|---|---|---|---|---|---|
| Category | Average | SD | *N* | Average | *N* | Model 2 |
| Animal | 79.8 | 22.6 | 9 | 96.2 | 96 | 89.2 |
| Creature | 76.1 | 18.3 | 8 | 94.0 | 33 | 90.1 |
| Imaginary | 64.2 | 18.5 | 7 | 49.1 | 18 | 62.7 |
| Human | 60.0 | 16.6 | 11 | 92.6 | 283 | 55.2 |
| Plant | 51.0 | 10.4 | 12 | 59.7 | 87 | 53.6 |
| Natural | 34.7 | 19.6 | 8 | 19.7 | 29 | 44.9 |
| Artifact | 33.0 | 22 | 17 | 17.1 | 273 | 50.0 |

The category ‘creature’ includes living beings other than mammals and birds, which are in the category ‘animal.’ The category ‘imaginary’ includes names for animate beings that do not actually exist, such as *unicorn* and *ghost*. The ‘natural’ category includes words referring to non-animate entities that occur naturally, such as *mud* and *salt*. The ‘artifact’ category includes words referring to non-animate entities constructed by human beings, such as *guitar* and *hat*. Radanović et al. are average animacy ratings, from 0 ‘Definitely inanimate’ to 100 ‘Definitely animate.’ The VanArsdall and Blunt (2022) living/non-living ratings have been converted from their original 7-point scale to be out of 100 as well. The final column includes fitted values from the living/non-living Model 2.

*co-occurrence*). It is possible for two words to have highly similar LSA vectors without ever appearing in the same document. For example, one can easily imagine that in some set of documents the informal word *cat* and the more formal word *feline* might never appear in the same document, but nevertheless would be likely to occur in documents that share many other words.

The basic principle of constructing vector representations of a word’s context continues in more recent word-embedding models, but the methods of constructing the vectors have been refined. There are two main differences. One difference is that most contemporary models do not construct their vectors from co-occurrence within *documents*, but rather from co-occurrence within some smaller moving window of text (which may be conceived of as very tiny documents, to keep the analogy with LSA precise). The second difference is that contemporary models do not merely count words but rather use more sophisticated computational methods to build the context vectors. In this paper I used a model called *word2vec* (Mikolov et al., 2013a,b,c). Skipping over some minor computational complications, word2vec models use a neural network with a single hidden layer (which is what is used as the vector representation of the word) to either predict the context of a target word (called *CBOW*, for *continuous bag of words*) or the inverse: to use context to try to predict a target word (*skipgram*, because the target word has been ‘skipped’ with context on either side). This paper uses the *skipgram* model with a 300-unit hidden layer and a context defined as two words on either side of the target word. Although these parameters are arbitrary, these values are commonly used in language research. For a corpus, I used the 150,000 most frequent words from a 100 billion words subset of the Google news corpus.[1] To increase the chances that the results might have a clear interpretation, I applied principal components analysis (PCA) to this matrix, retaining all 300 principal components (PCs). The magnitude of the PC can thereby give us an estimate of how much variance in the matrix is accounted for.

Four words of the 1,200 words from VanArsdall and Blunt (2022) were eliminated from this study. The word *bluejay* was eliminated because it is confounded with the name of Canada’s favorite baseball team and appears only in capitalized form in the Google news corpus. The word is also problematic since the name of the bird is normally not considered a compound word but is rather composed of two words. Similarly, the word *hornet* appeared in the Google news corpus only in capitalized form (though it contained the plural form *hornets*), presumably referring to the name of the Marvel comic character. The word *ghoul* did not appear in the corpus, although *ghouls* did. The word *sphinx* did appear in the corpus, but only in capitalized form. The remaining dataset was randomly split into two sets of 598 words, with one half used for model development, and the other for cross-validation.

## 2.1. Model 1: method

All reported analyses were conducted in R 4.2.2 (R Core Team, 2022) using R Studio (2022.12.0 + 353; Posit Software, 2022) for macOS. The GAMs were analyzed using the mgcv package (v. 1.8–41, Wood, 2022; see also Wood, 2017).

Ninety-seven of the 300 PCs were significantly ($p <= 0.05$) correlated by GAM (i.e., possibly non-linearly) with the human animacy ratings from Radanović et al. (2016). This included 13 of the first 20 PCs (but not PC1). Since this provided more reliable predictors than there are data points, I used the 21 PCs that had a GAM whose output correlated with the human estimates at $p <= 0.001$ to construct a full GAM model. All predictors were entered initially. Those with the highest value of $p$ were removed one by one until all remaining predictors entered with $p < 0.05$.

Ninety-two of the of the 300 PCs were significantly ($p <= 0.05$) correlated by GAM with the human living/non-living ratings from VanArsdall and Blunt (2022). This also provided more predictors than datapoints, because each smooth in the GAM has nine parameters using the default rank value (number of possible turning points, or *knots*) of 10 (the tenth is eliminated by centering the predictors). I therefore used the same method as above, initially entering only the 29 PCs that had a GAM whose output correlated with the human

---

[1] https://code.google.com/archive/p/word2vec/

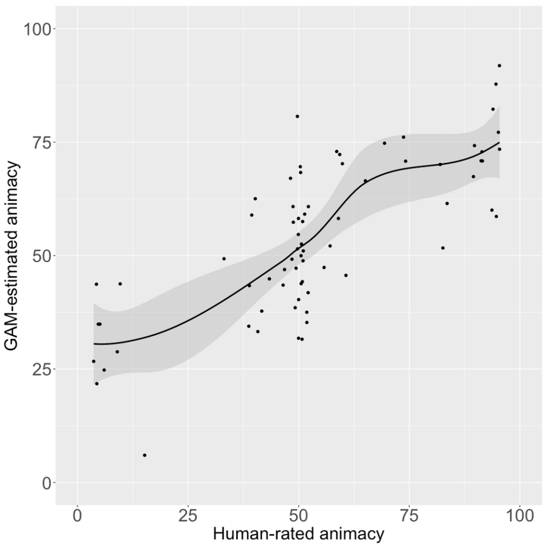


FIGURE 1
Human-rated (*X*-axis) versus GAM-estimated (*Y*-axis) animacy, using word2vec PCs as predictors. The gray-shaded area is the 95% confidence interval.

**TABLE 2 Best GAM model to predict the living/non-living ratings from VanArsdall and Blunt (2022), using word-embedding PCs as predictors.**

| PREDICTOR | RANK | df | *F* | *p* |
|---|---|---|---|---|
| PC30 | 5.00 | 6.09 | 3.25 | 0.004 |
| PC138 | 6.28 | 7.46 | 2.30 | 0.020 |
| PC225 | 8.43 | 8.91 | 3.12 | 0.001 |
| PC248 | 3.06 | 3.93 | 2.78 | 0.026 |

judgments at $p <= 0.001$. This included four of the first 20 PCs, but again, not PC1.

## 2.2. Model 1: result

Only two PCs entered the model of the animacy ratings: PC123 and PC246. Together these PCs accounted for 56.6% of the variance in those ratings ($p < 0.00001$; see Figure 1).

I constructed a dictionary by taking the 75,000 most frequent words from Shaoul and Westbury, 2006. I eliminated words that did not appear in the Google news matrix, which does not include closed class words, as well as compounds words (or phrases) with spaces in them. The final dictionary contains 67,717 words. Applying the GAM to this dictionary suggested that the model may be over-fit to the small data set, since the words estimated most highly animate were not clearly exemplars of any animate category. The top 10 words were *disclaims, threes, clientless, fouling, republication, desegregation, effigies, barriers, reflate*, and *mineralization*.

Four PCs entered the living/non-living model: PC30, PC138, PC248, and PC225 (see Table 2). Together these PCs accounted for 12.5% of the variance in the human ratings ($p < 0.00001$). The model did not cross-validate successfully. Its predictions were unreliably correlated ($r = -0.02$, $p = 0.61$) with the human living/non-living ratings in the validation set.

## 2.3. Model 1: discussion

Although the models did not generalize well to the full dictionary or to a validation dataset, we can draw some tentative conclusions from this initial model.

The lack of good generalization and the lack of concordance between the two models suggests that one conclusion we can draw is that little of the variance in animacy or living/non-living judgments can be derived from the PCs in a word embedding model. The failure of these 'bottom-up' models suggests that animacy or being alive are not strongly encoded in patterns of word use. More speculatively, we can conclude that animacy is not a basic component of lexical semantics, since many components considered to be basic can be well-estimated from the PCs (e.g., see Hollis and Westbury, 2016; Hollis et al., 2017; Westbury and Hollis, 2019).

However, that said, the second conclusion is that animacy may be correlated with other aspects of semantics, since a large number of individual PC GAMs produced estimates that were reliably correlated with the human animacy ratings. The Radanović et al. (2016) are reliably correlated with the extrapolated estimates of human judgments of valence, dominance, and arousal from Hollis et al. (2017). Higher animacy ratings are associated with lower valence ($r = -0.29$, $p = 0.01$), higher arousal ($r = 0.35$, $p = 0.003$), and lower dominance ($r = -0.33$, $p = 0.005$). The negative correlation with dominance reflects the fact that many small (i.e., low dominance) living things such as *amoeba* (83/100), *bacteria* (83/100), *squirrel* (95/100), *worm* (92/100), and *spider* (94/100) receive high animacy ratings. The animacy ratings are also positively correlated ($r = 0.31$, $p = 0.007$) with the measure of self-relevance (how strongly a word is associated with the first-person singular word *I*) that was defined in Westbury and Wurm (2022), where it was shown to strong predictor of the value of early PC values across a large dictionary. A GAM developed with all these values to predict the animacy ratings allowed in only arousal and self-relevance with $p < 0.05$. Together these two measures accounted for 21.2% of the variance in the ratings.

A third conclusion is that (tautologically) most of the error in predicting animacy is seen for words of ambiguous animacy. In Figure 1 there is a wide range of model estimates for words that were rated the mid-range of animacy by humans in Radanović et al. (2016).

## 3. Model 2: introduction

Word2vec vectors for category names (such as the vector representing the word *animal*) usually (though not necessarily) serve as centroids for the category they name. This means that words with vectors that are similar (by cosine distance) to the vector for a category name are often members of that category. For example, the twenty vectors most similar to the vector of the word *vegetable* are the vectors for the words *tomato, potato, tomatoes, broccoli, sweet_potato, onion, onions, cauliflower, mango* (oops!), *kale, potatoes, mangos, cabbage*, and *melons*. We may perhaps forgive the model for sometimes confusing vegetables and fruit, since we ourselves routinely refer to the tomato fruit, the avocado fruit, the olive fruit, the cucumber fruit, the zucchini fruit, and several other fruits (strictly speaking, plants in which the edible part develops from a flower) as vegetables. If humans discuss fruits as if they were vegetables, we must expect that word embedding

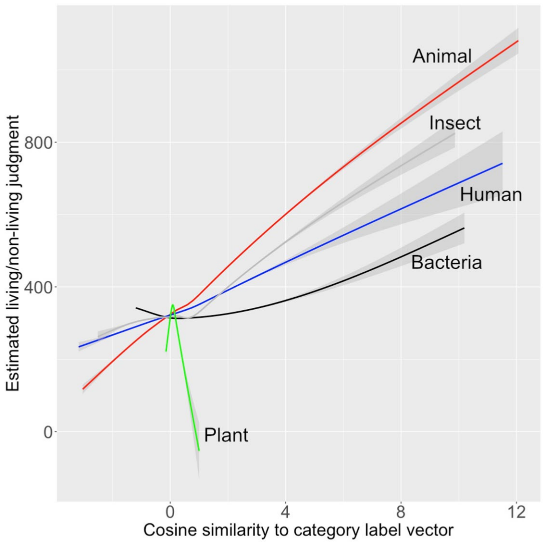


FIGURE 2
Pearson correlations between category-name predictors and human ratings.

**TABLE 3 Best GAM model to predict the animacy ratings from Radanović et al. (2016), using cosine distance to category label vectors as predictors.**

| Predictor | Rank | DF | *F* | *p* |
|---|---|---|---|---|
| Insect | 1.00 | 1.00 | 9.64 | 0.003 |
| Human | 1.00 | 1.00 | 7.24 | 0.009 |

models will reflect that. Of course, fruits and vegetables also do correctly both belong to many other categories: *plant products, things we eat, things that cannot thrive in freezing weather, domesticated products, things you will find at the grocery store, everyday objects, things that can be composted*, and so on. A word-embedding model of categorization may be influenced by all these categories simultaneously since it can only induce the categories from the similarity of the contexts of words as encoded in the words' vectors. It is possible that a super-ordinate category could be better captured by patterns of word use than a more focal category, if people used language in a way that better reflected that super-ordinate category.

In the second model, this categorizing feature of word embedding models is used, by building models of the human ratings based on the distance from the vectors of the names of categories of unambiguously animate things.

## 3.1. Model 2: method

For predictors I used the cosine similarity of each word that had been classified by humans to five main category names of definitely animate and living things: *plant, animal, insect, human,* and *bacteria*. Of the five taxonomic kingdoms, three are captured by these categories (plant, animal, and bacteria/~monera). The other two (funghi and protista) are less relevant kingdoms when it comes to animacy. Insects and humans are broken out of the animalia kingdom to which they belong because they are regularly incorrectly classified as non-animate.

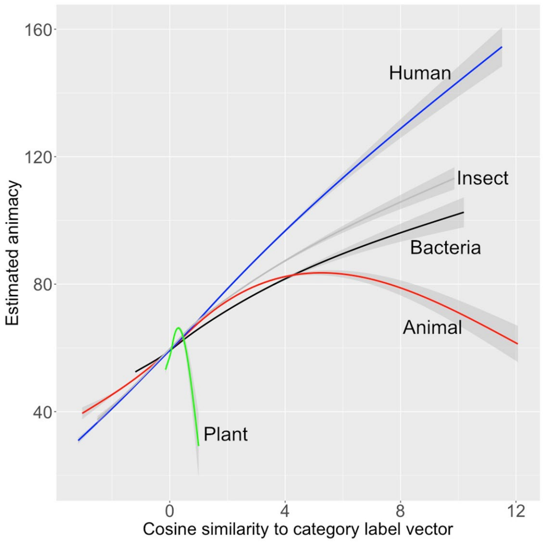


FIGURE 3
Model 2's estimates of animacy (*y*-axis) graphed against the normalized cosine similarity of the vectors for the category labels of 67,717 words (*x*-axis) The gray-shaded area is the 95% confidence interval.

## 3.2. Model 2: results

The Pearson correlations between all the predictors and the ratings from Radanović et al. (2016) are shown in Figure 2. The correlations between the human ratings for each word and the cosine distance of their word2vec vectors from the vectors of the category labels were reliable at $p < 0.001$ for all categories except *plant* ($r = 0.13$, $p > 0.05$).

The best GAM model to predict the Radanović et al. (2016) ratings included only two predictors that entered with $p < 0.05$, *insect* and *human*. This model is summarized in Table 3. It accounted for 28.8% of the variance in the human ratings. The relationship between the predictors and the model estimates for are shown graphically in Figure 3.

When the model was applied to the full dictionary, the 10 words estimated most animate were almost all insects: *beetle, aphid, beetles, moth, pests, pest, aphids, wasps, wasp,* and *fungus*.

The living/non-living judgments from VanArsdall and Blunt (2022) were modeled in the same way. Four predictors entered with $p < 0.05$: *animal, bacteria, insect* and *plant* (see Table 4). The model accounted for 26.7% of the variance in the human ratings in the development set and 23.7% of the variance in the validation set. The model was applied to the full dictionary. The 10 words estimated most animate were *animal, insect, rodent, animals, owl, bird, reptile, critter, feline,* and *elephant*. This list has high face validity, both because it only includes only words that name living things and because it includes many high-level living-thing category names. The relationship between the predictors and the living/nonliving judgments are shown graphically in Figure 4.

The predictions from this model are broken down into categories in the rightmost column of Table 1. The seven average categorical predictions from the model are highly correlated with the average categorical human ratings of both animacy ($r = 0.91$, $p = 0.002$

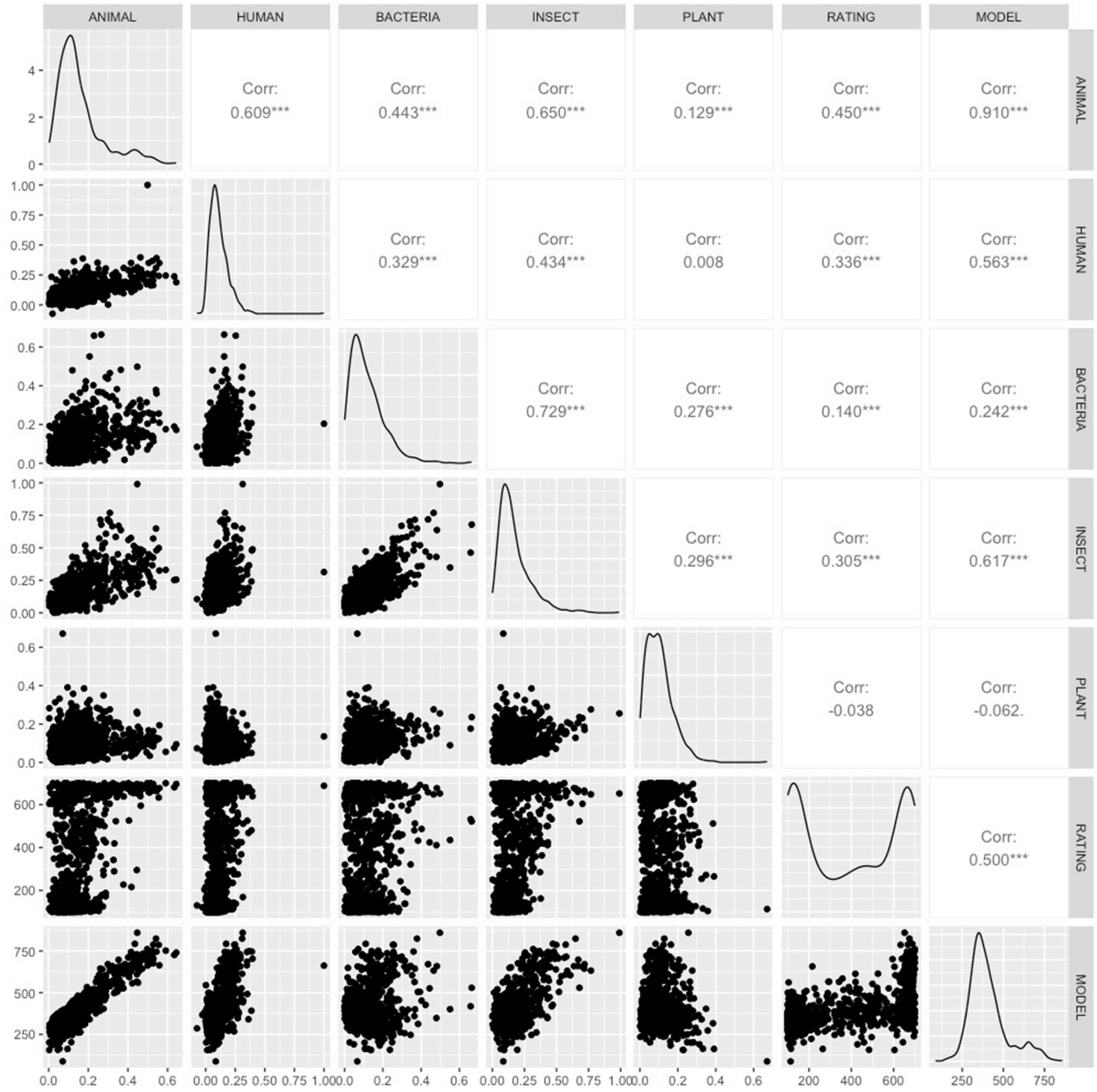


FIGURE 4
Relationships between human living/non-living ratings, Model 2's estimates of those ratings, and vector cosine similarity of each word's vector to five animate category labels.

one-sided) and living/non-living ($r = 0.81$, $p = 0.01$ one-sided) (Table 4).

## 3.3. Model 2: discussion

The results from all models to predict human animacy and living/non-living ratings are summarized in Table 5. There are two main findings.

One is that modeling human animacy and living/non-living judgments using distance from category names is more successful than modeling them using word2vec PCs. Although the word2vec PCs predicted the 72 animacy judgments relatively well ($R^2 = 0.57$), that model had very low face validity when extended to the whole dictionary. Those word2vec vectors were also poor at predicting the living/non-living judgments. The best model accounted for only 12.5% of the variance and failed to cross-validate at all. In contrast, the

**TABLE 4** Best GAM model to predict the living/non-living ratings from VanArsdall and Blunt (2022).

| Predictor | Rank | DF | *F* | *p* |
|---|---|---|---|---|
| Animal | 1.91 | 2.40 | 25.79 | <2e-16 |
| Bacteria | 1.67 | 2.10 | 5.57 | 0.004 |
| Insect | 2.87 | 3.62 | 3.10 | 0.020 |
| Plant | 3.34 | 4.13 | 2.54 | 0.041 |

Using cosine distance to category label vectors as predictors.

TABLE 5 Results for predicting human animacy ratings, using GAM with PCs (Model 1) or cosine similarity to the vectors for label names (Model 2).

| Predictors | Data | $R^2$ |
|---|---|---|
| PCs | Animacy | 0.566 |
| PCs | Living [Dev] | 0.125 |
| PCs | Living [Val] | 0 |
| Cosine | Animacy | 0.288 |
| Cosine | Living [Dev] | 0.267 |
| Cosine | Living [Val] | 0.237 |

model using distances from category names accounted for roughly the same variance in the animacy (28.8%) and living judgments (26.7%), although of course there are many more living judgments. That model cross-validated relatively well, accounting for 23.7% of the variance in the living/non-living judgment validation dataset. It also had high face validity when applied to a larger set of words.

The other finding of interest is that neither of models using categorical distance included distance from the category *human*. This is noteworthy because (as shown in Table 1 and discussed above) human categories tend to be rated low by humans on both animacy and (to a lesser extent) living/non-living judgments, where they received an average rating of 92.6/100, compared to 94.0/100 for mammals and birds and 94.0/100 for other living creatures.

## 4. General discussion

Of course, if we provided a model of animacy with categorical information, it would achieve perfect classification, since the five categories of plant, animal, insects, humans (which are of course also animals, but we generally do not speak of them this way), and bacteria cover the superordinate category of the animate almost perfectly. The fact that humans are not unanimous about their decisions suggests that human beings must not be relying on categorical information, which we already knew from their failure to accept members of these categories as animate with perfect accuracy.

The fact that the pattern of errors in the models is similar to the pattern of errors seen in humans suggests that human may be making animacy decisions based on contextual information (or the categories that may be derived from that information) rather than on category membership.

The model which used cosine distance from category labels performed much better at classifying words as being animate than the analogous model that used PCs. We can roughly conceive of the models as being bottom-up (PC predictors) versus top-down (category label predictors). These results therefore suggest that animacy is unlike valence or arousal, which are usually conceived as being components of semantics (Osgood et al., 1957). It is rather more like *being expensive* or *being soft*, an objectively grounded top-down classification that we learn from experience.

The second noteworthy finding supports this. That is the fact that the models built on cosine distance from the category name vectors make one of the same errors that humans do: they tend to rate humans as lower in animacy than animals. Table 1 shows that human beings rated human words (such as *mother, boy,* and *professor*) as animate at 60/100, compared to 79.8 for animal names. Similarly, the model rates humans at 55.2, compared to 89.2 for animals. This may reflect that humans are not generally conceived of (or, at least, written about) as animate.

The model also replicates humans in (erroneously) classifying plants as moderately animate. Humans rated plants at 51/100 (Table 1). The model rates them at 53.6/100.

The top 200 most animate words according to the final model are reproduced in Supplementary Appendix 1. Animacy ratings for the full dictionary of 67,717 words are available at https://osf.io/k3cn9/.

It is obvious that humans do not make animacy decisions using category membership. If they did their animacy ratings would be unanimously high or low for many words that get intermediate ratings. The success of the category vector distance models at modeling human ratings suggests that humans are instead making animacy judgments by making rough animate category membership judgments (without considering the category of humans, according the best model discussed above). This may have implications for studies looking at animacy effects. The repeated finding that humans and living things outside of the animalia kingdom are poorly classified as animate by models using cosine distance from the vectors of category names suggests that language use does not present humans and living things outside of the animalia kingdom in contexts that highlight their animacy. These results suggest that humans make animacy ratings not by considering the category of each item, but rather by making family resemblance judgments to animate categories. The nature and direction of those judgments are reflected in word-embedding models.

## 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.

## Ethics statement

The studies involving human participants were reviewed and approved by the University of Alberta Research Ethics Board. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

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

## Funding

This work was funded by the Natural Sciences and Engineering Research Council of Canada, grant#RGPIN-2018-04679.

## Acknowledgments

Thanks to the editor and reviewers for their helpful feedback on an earlier version of this manuscript.

## 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.

## Publisher's note

## Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1145289/full#supplementary-material

## References

Bonin, P., Gelin, M., Dioux, V., and Méot, A. (2019). "It is alive!" evidence for animacy effects in semantic categorization and lexical decision. *Appl. Psycholinguist.* 40, 965–985. doi: 10.1017/S0142716419000092

Bonin, P., Gelin, M., Laroche, B., Méot, A., and Bugaiska, A. (2015). The "how" of animacy effects in episodic memory. *Exp. Psychol.* 62, 371–384. doi: 10.1027/1618-3169/a000308

Bugaiska, A., Grégoire, L., Camblats, A. M., Gelin, M., Méot, A., and Bonin, P. (2019). Animacy and attentional processes: evidence from the stroop task. *Q. J. Exp. Psychol.* 72, 882–889. doi: 10.1177/1747021818771514

Bugaiska, A., Méot, A., and Bonin, P. (2016). Do healthy elders, like young adults, remember animates better than inanimates? An adaptive view. *Exp. Aging Res.* 42, 447–459. doi: 10.1080/0361073X.2016.1224631

Cappa, S. F., Perani, D., Schnur, T., Tettamanti, M., and Fazio, F. (1998). The effects of semantic category and knowledge type on lexical-semantic access: a PET study. *NeuroImage* 8, 350–359. doi: 10.1006/nimg.1998.0368

Caramazza, A., and Shelton, J. R. (1998). Domain-specific knowledge systems in the brain the animate-inanimate distinction. *J. Cogn. Neurosci.* 10, 1–34. doi: 10.1162/089892998563752

Darnell, R., and Vanek, A. L. (1976). The semantic basis of the animate/inanimate distinction in Cree. *Res. Lang. Soc. Interact.* 9, 159–180.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, M. J., et al. (2002). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia* 40, 54–75. doi: 10.1016/S0028-3932(01)00066-5

Félix, S. B., Pandeirada, J., and Nairne, J. S. (2020). Animacy norms for 224 European Portuguese concrete words. *Análise Psicológica* 38, 257–269. doi: 10.14417/ap.1690

Gennari, S. P., Mirković, J., and MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cogn. Psychol.* 65, 141–176. doi: 10.1016/j.cogpsych.2012.03.002

Grabowski, T. J., Damasio, H., and Damasio, A. R. (1998). Premotor and prefrontal correlates of category-related lexical retrieval. *NeuroImage* 7, 232–243. doi: 10.1006/nimg.1998.0324

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., and Schlesewsky, M. (2006). Linguistic prominence and Broca's area: the influence of animacy as a linearization principle. *NeuroImage* 32, 1395–1402. doi: 10.1016/j.neuroimage.2006.04.213

Hollis, G., and Westbury, C. (2016). The principals of meaning: extracting semantic dimensions from co-occurrence models of semantics. *Psychon. Bull. Rev.* 23, 1744–1756. doi: 10.3758/s13423-016-1053-2

Hollis, G., Westbury, C., and Lefsrud, L. (2017). Extrapolating human judgments from skip-gram vector representations of word meaning. *Q. J. Exp. Psychol.* 70, 1603–1619. doi: 10.1080/17470218.2016.1195417

Ilić, O., Ković, V., and Styles, S. J. (2013). In the absence of animacy: superordinate category structure affects subordinate label verification. *PLoS One* 8:e83282. doi: 10.1371/journal.pone.0083282

Kazanas, S. A., Altarriba, J., and O'Brien, E. G. (2020). Paired-associate learning, animacy, and imageability effects in the survival advantage. *Mem. Cogn.* 48, 244–255. doi: 10.3758/s13421-019-01007-2

Landauer, T. K., and Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* 104, 211–240. doi: 10.1037/0033-295X.104.2.211

Mahjoubnavaz, F., and Mokhtari, S. (2022). Animacy in Persian language: Animacy norms for 401 Persian words. *J. Neurodevelop. Cogn.* 5, 21–37.

Mak, W. M., Vonk, W., and Schriefers, H. (2002). The influence of animacy on relative clause processing. *J. Mem. Lang.* 47, 50–68. doi: 10.1006/jmla.2001.2837

Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature* 379, 649–652. doi: 10.1038/379649a0

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. ICLR. Available at: https://arxiv.org/abs/1301.3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*. 26:3111–3119.

Mikolov, T., Yih, W. T., and Zweig, G. (2013c). "Linguistic regularities in continuous space word representations,"in *Proceedings of the 2013 conference of the North American chapter of the Association for Computational Linguistics: Human language technologies*. 746–751.

Moore, C. J., and Price, C. J. (1999). A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122, 943–962. doi: 10.1093/brain/122.5.943

Mummery, C. J., Patterson, K., Hodges, J. R., and Price, C. J. (1998). Functional neuroanatomy of the semantic system: divisible by what? *J. Cogn. Neurosci.* 10, 766–777. doi: 10.1162/089892998563059

Nairne, J. S., VanArsdall, J. E., and Cogdill, M. (2017). Remembering the living: episodic memory is tuned to animacy. *Curr. Dir. Psychol. Sci.* 26, 22–27. doi: 10.1177/0963721416667711

Oxford English Dictionary (2022). Oxford University Press, September 2022. Available at: www.oed.com/view/Entry/248721 (Accessed December 1, 2022).

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The measurement of meaning*. Chicago, USA: University of Illinois Press.

Perani, D., Cappa, S. F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., et al. (1995). Different neural systems for the recognition of animals and man-made tools. *Neuroreport* 6, 1637–1641. doi: 10.1097/00001756-199508000-00012

Perani, D., Schnur, T., Tettamanti, M., Gorno, M., Cappa, S. F., and Fazio, F. (1999). Word and picture matching: a PET study of semantic category effects. *Neuropsychologia* 37, 293–306. doi: 10.1016/S0028-3932(98)00073-6

Pilgrim, L. K., Fadili, J., Fletcher, P., and Tyler, L. K. (2002). Overcoming confounds of stimulus blocking: an event-related fMRI design of semantic processing. *NeuroImage* 16, 713–723. doi: 10.1006/nimg.2002.1105

Popp, E. Y., and Serra, M. J. (2016). Adaptive memory: animacy enhances free recall but impairs cued recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 186–201. doi: 10.1037/xlm0000174

Popp, E. Y., and Serra, M. J. (2018). The animacy advantage for free-recall performance is not attributable to greater mental arousal. *Memory* 26, 89–95. doi: 10.1080/09658211.2017.1326507

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radanović, J., Westbury, C., and Milin, P. (2016). Quantifying semantic animacy: how much are words alive? *Appl. Psycholinguist.* 37, 1477–1499. doi: 10.1017/S0142716416000096

Shaoul, C., and Westbury, C. (2006) *USENET orthographic frequencies for 111,627 English words. (2005–2006)*. Edmonton, AB: University of Alberta.

Szewczyk, J. M., and Schriefers, H. (2011). Is animacy special?: ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Res.* 1368, 208–221. doi: 10.1016/j.brainres.2010.10.070

Traxler, M. J., Williams, R. S., Blozis, S. A., and Morris, R. K. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *J. Mem. Lang.* 53, 204–224. doi: 10.1016/j.jml.2005.02.010

Tyler, L. K., Bright, P., Dick, E., Tavares, P., Pilgrim, L., Fletcher, P., et al. (2003). Do semantic categories activate distinct cortical regions? Evidence for a distributed neural semantic system. *Cogn. Neuropsychol.* 20, 541–559. doi: 10.1080/02643290244000211

Tyler, L., and Moss, H. (2001). Towards a distributed account of conceptual knowledge. *Trends Cogn. Sci.* 5, 244–252. doi: 10.1016/S1364-6613(00)01651-X

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., and Levy, J. P. (2000). Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. *Brain Lang.* 75, 195–231. doi: 10.1006/brln.2000.2353

VanArsdall, J. E., and Blunt, J. R. (2022). Analyzing the structure of animacy: exploring relationships among six new animacy and 15 existing normative dimensions for 1,200 concrete nouns. *Mem. Cogn.* 50, 997–1012. doi: 10.3758/s13421-021-01266-y

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Cogdill, M. (2015). Adaptive memory: Animacy effects persist in paired-associate learning. *Memory* 23, 657–663. doi: 10.1080/09658211.2014.916304

Vihman, V. A., Nelson, D., and Kirby, S. (2018). Animacy distinctions arise from iterated learning. *Open Linguist.* 4, 552–565. doi: 10.1515/opli-2018-0027

Westbury, C., and Hollis, G. (2019). Conceptualizing syntactic categories as semantic categories: unifying part-of-speech identification and semantics using co-occurrence vector averaging. *Behav. Res. Methods* 51, 1371–1398. doi: 10.3758/s13428-018-1118-4

Westbury, C., and Wurm, L. H. (2022). Is it you you're looking for? Personal relevance as a principal component of semantics. *Ment. Lexicon* 17, 1–33. doi: 10.1075/ml.20031.wes

Wood, S. (2017). *Generalized additive models: An introduction with R, 2nd*. London, England: Chapman and Hall/CRC.

Wood, S. (2022). mgcv: Mixed GAM computation vehicle with automatic smoothness estimation. Available at: https://cran.r-project.org/web/packages/mgcv/index.htm

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 August 2023
DOI 10.3389/fpsyg.2023.1173352

OPEN ACCESS

EDITED BY
Josefa N. S. Pandeirada,
University of Aveiro, Portugal

REVIEWED BY
Helena Mendes Oliveira,
University of Minho, Portugal
Mallory Stites,
Sandia National Laboratories (DOE),
United States

*CORRESPONDENCE
Anna Czypionka
✉ anna.czypionka@uni-konstanz.de

RECEIVED 24 February 2023
ACCEPTED 31 July 2023
PUBLISHED 17 August 2023

CITATION
Czypionka A, Kharaman M and Eulitz C (2023) *Wolf-hound* vs. *sled-dog*: neurolinguistic evidence for semantic decomposition in the recognition of German noun-noun compounds. *Front. Psychol.* 14:1173352.
doi: 10.3389/fpsyg.2023.1173352

# *Wolf-hound* vs. *sled-dog*: neurolinguistic evidence for semantic decomposition in the recognition of German noun-noun compounds

Anna Czypionka*, Mariya Kharaman and Carsten Eulitz

Department of Linguistics, University of Konstanz, Konstanz, Germany

Animacy is an intrinsic semantic property of words referring to living things. A long line of evidence shows that words with animate referents require lower processing costs during word recognition than words with inanimate referents, leading among others to a decreased N400 amplitude in reaction to animate relative to inanimate objects. In the current study, we use this animacy effect to provide evidence for access to the semantic properties of constituents in German noun-noun compounds. While morphological decomposition of noun-noun compounds is well-researched and illustrated by the robust influence of lexical constituent properties like constituent length and frequency, findings for semantic decomposition are less clear in the current literature. By manipulating the animacy of compound modifiers and heads, we are able to manipulate the relative ease of lexical access strictly due to intrinsic semantic properties of the constituents. Our results show additive effects of constituent animacy, with a higher number of animate constituents leading to gradually attenuated N400 amplitudes. We discuss the implications of our findings for current models of complex word recognition, as well as stimulus construction practices in psycho-and neurolinguistic research.

KEYWORDS

**animacy, compound, N400, word recognition, semantic decomposition**

## 1. Introduction

Animacy describes the property of certain things that we perceive as “having a soul”[1] or more simply put as being alive; the least ambiguous examples involving (vertebrate) animals including fellow humans. The distinction between animate and inanimate entities shapes many different areas of cognition.

With respect to language, animacy belongs to the semantic properties making up the meaning of a word, and is arguably one of the most striking and influential of its semantic properties. In natural language, words referring to animate entities are highly salient compared to words referring to inanimate entities. This is visible in the special “treatment” that animate entities get in the world’s languages. In sentence production, speakers go out of their way to produce animates in early sentence positions and make them the subjects, rather than the objects, of sentences, often at the cost of syntactic simplicity. In sentence comprehension, the animacy of the arguments is one of the crosslinguistically most robust cues for understanding who did what to whom. Even at the single-word level, words referring

---

[1] From Latin *anima*, “the soul”.

to animate entities are recognized more quickly than words referring to inanimate entities, thanks to their high saliency.

In this paper, we present an EEG study using this processing benefit of animates over inanimates to answer a long-standing question in single-word recognition research. Our question concerns the recognition of compound words consisting of two nouns, and whether the semantic properties of both constituent nouns are routinely accessed during compound recognition.

In the following, we provide an overview of the role of animacy in language processing, followed by a literature overview of the processing of noun-noun compounds. We will then give an outline of how we use animacy to monitor semantic decomposition in compound recognition, and formulate our research questions and hypotheses in more detail before presenting the results of our study.

## 2. Background

### 2.1. Animacy

The semantic property of animacy is a strong influence in many languages of the world, both at the single-word and sentence level. Animates and inanimates are referred to with different interrogative pronouns (*who* and *what*, respectively, in English), and differ in number marking (Croft, 1990; Corbett, 2000; Haspelmath, 2013). In Differential Object Marking (DOM) languages (like Hindi or Spanish), overt object case marking is only obligatory for a particular semantic class of nouns, with animacy or even humanness being a frequent classification (Bossong, 1985, 1991; Næss, 2004; Malchukov, 2008). Furthermore, the morphological makeup of nouns via case syncretism is shaped by animacy in a complex interaction with agentivity, biological and grammatical gender (Krifka, 2009). A thorough overview of crosslinguistic animacy effects can be found in Yamamoto (1999).

The special status of the animate-inanimate distinction in human language is mirrored in language processing.[2] During language acquisition, the distinction between animates and inanimates develops early in life (Opfer and Gelman, 2010), an observation that also holds for autistic children (Rutherford et al., 2006). In the case of language deficits, different studies report category-specific deficits affecting only one semantic subclass, while other semantic subclasses are spared. Animacy is one of the relevant semantic subclasses; for example, a patient may exhibit impaired naming for animals, but not for fruit and vegetables (Caramazza and Shelton, 1998) or artifacts (see Capitani et al., 2003; Caramazza and Mahon, 2003 for overviews).

In adult language processing, there is also a host of evidence for a distinction between animates and inanimates. In general, findings suggest that lexical access is less costly for animates than for inanimates. Behavioral studies in multiple languages show that reaction times are shorter for animates than inanimates in word and picture naming (Janyan and Andonova, 2011), semantic categorization and lexical decision tasks (Bonin et al., 2019), and for ink color naming in a Stroop task adaptation (Bugaiska et al., 2019). In addition, animates are remembered better than inanimates, both in free recall and paired-associate tasks (Nairne et al., 2013; VanArsdall et al., 2015); these findings are unlikely to be reduced to categorical recall strategies (VanArsdall et al., 2017). Further support comes from Bonin et al. (2014), who found that animates are remembered better than inanimates, both for word and picture stimuli, and word recall and recognition tasks (see also Bonin et al., 2015 for replication and added detail). Neurolinguistic evidence[3] for the animate-inanimate distinction includes differential BOLD responses for animals relative to manipulable objects (Anzellotti et al., 2011) and differences in the EEG spectral power (Verkhlyutov et al., 2014). In EEG studies, Sitnikova et al. (2006) found an increased anterior negativity for animals relative to tools, and left-posterior negativity for tools relative to animals, between 200 and 600 ms. They interpret their findings as evidence for feature-based organization of semantic knowledge. Proverbio et al. (2007) investigated images of animate and inanimate stimuli in a non-verbal categorization task. Stimulus pairs were presented and participants had to judge if they belonged to the same or to different semantic categories (animals or artifacts). Compared to artifacts, animates showed shorter reaction times, higher accuracy, a larger P300 amplitude and a reduced N400 amplitude. The authors conclude that in contrast to animates, manipulable objects lead to the activation of areas associated with motor representation (see, however, findings by Ković et al., 2009, suggesting no N400 amplitude differences for animates relative to inanimates.).

In the processing of sentences, the animacy of arguments is the central semantic cue for argument role assignment (e.g., MacDonald et al., 1994; Trueswell et al., 1994; Weckerly and Kutas, 1999; Frisch and Schlesewsky, 2001; Kuperberg, 2007; Branigan et al., 2008; Bornkessel-Schlesewsky et al., 2011; Paczynski and Kuperberg, 2011; Czypionka, 2014), interacting with the processing of number agreement (Bamyacı et al., 2014) and case marking (Verhoeven, 2014; Czypionka and Eulitz, 2018). The prominent role of argument animacy in sentence processing is reflected in its central role in models of sentence processing, where it is associated with the assignment of thematic roles (see, among others, Levelt, 1993; Bornkessel-Schlesewsky and Schlesewsky, 2006, 2009, 2013; Hagoort, 2007, 2016; Kuperberg, 2007 for different approaches to sentence comprehension and production).[4]

In sum, animacy is an intrinsic semantic property of a word's referent that influences all levels of language processing. Words with animate referents are highly salient in the sentence and discourse context. Lexical access is less costly for words with animate referents than for words with inanimate referents. This is reflected in shorter reaction times and reduced N400 amplitudes for

---

Abbreviations: BOLD, Blood Oxygenation Level Dependent; EEG, Electroencephalogram; ERP, Event-related potential; MMN, Mismatch negativity.

[2] For the sake of readability, we will use "animate/inanimate" to also refer to "words with animate/inanimate referents". In case a clear distinction between referents and words referring to them is necessary, we will be more specific.

[3] Please refer to the Supplementary material for background information on the ERP literature and on the P300 and N400, the two ERP components that are most relevant in the context of the current study.

[4] For the sake of readability, we do not give an in-depth overview of the role of animacy in sentence processing at this point. We will discuss the implications of our findings for sentence processing research in the discussion.

animates relative to inanimates. This makes animacy a useful tool for investigating the role of semantics in single-word processing, all the more as it is an intrinsic property that does not depend on context.[5] In the following, we will outline how these properties can be informative for questions related to compound processing, in particular with respect to semantic decomposition.

## 2.2. Compounds

Compounds are words consisting of more than one constituent; in the context of this paper, we will refer exclusively to noun-noun compounds unless specifically mentioned otherwise.[6] These are words like *gunpowder* or *garden hose*. In English, these words appear both as a single orthographic unit (*gunpowder*) and as two adjacent nouns (*garden hose*), with little semantic difference between both options. In other languages like German, however, orthographic rules for noun-noun compounds demand that they appear as one orthographic unit (*Schieß.pulver* "gunpowder", *Garten.schlauch* "garden hose", dots marking the constituent boundaries are not part of the German orthography and are only inserted here for clarity).

The lexical category, syntactic features, and main semantic properties of the compound depend on the lexical head, which is always the last (in our case, second) constituent in German compounds (as it is in most English compounds): *Schlittenhund*, "sled dog" is a kind of dog, not of sled, whereas *Hundeschlitten*, "dog sled", is a kind of sled, not a kind of dog. The first constituent is the modifier, extending and modifying the meaning of the lexical head: *Schlittenhund* is the specific kind of dog that pulls sleds, whereas *Hundeschlitten* is the specific kind of sled that is pulled by dogs.

Compounds can be semantically transparent or opaque. With semantically transparent compounds like *Pferdedecke* "horse rug", the full-form meaning can easily be inferred from combining the meanings of its constituents in a straightforward manner—a horse rug is a rug or blanket used to cover a horse. With semantically opaque compounds, the full-form meaning cannot be inferred by simply combining the constituent meanings (*Windbeutel*, literally "wind bag", is not a bag full of wind, but rather a cream-puff-like pastry).

In psycho- and neurolinguistic research, compounds are mostly studied with the focus on the nature of their lexical entries and lexical access. The main overarching research question in this literature is the amount and nature of compound decomposition, i.e., whether compounds are stored and accessed via their full-form meaning, or whether this meaning is calculated from the constituents when compounds are encountered. Related questions are concerned with whether decomposition occurs in a similar manner for all kinds of compounds, and which kinds of constituent information is accessed during decomposition.[7]

According to full-listing models (Butterworth, 1983; Bybee, 1995), known words are always stored and accessed in their full form in the lexicon. In contrast, full-parsing models (Taft and Forster, 1975; Libben et al., 1999; McKinnon et al., 2003; Taft, 2004; Taft and Ardasinski, 2006; Taft and Nguyen-Hoan, 2010) propose morphological decomposition for all complex words. Another proposal is that decomposition and full-form access are both a part of complex word recognition, but happen at different points in time (e.g., according to the supralexical model by Giraudo and Grainger, 2001, where full-form access precedes decomposition). Finally, dual-route models allow for both full-form access and decomposition before access, (e.g., Augmented Addressed Morphology Model by Caramazza et al., 1988, or the Morphological Race Model by Schreuder and Baayen, 1995; Baayen and Schreuder, 1999); the multiple-route model by Kuperman et al. (2009) also allows for parallel access via multiple and interactive routes. Which route ultimately leads to identification depends on the words' familiarity, its semantic transparency, and the frequencies of its constituents and full-form, among other factors.

The number and variety of different accounts of lexical access already hints at the very different findings with respect to compound recognition in the literature.[8] In general, a strong point in favor of decomposition is when properties of the modifier (in addition to full-form and head properties) can be shown to influence compound recognition.

While especially some earlier work argues against automatic decomposition of complex words (e.g., Sandra, 1990), the picture has become more nuanced over time, highlighting the important role of experimental paradigm, linguistic context and stimulus properties for eliciting compound decomposition. In an EEG study monitoring the processing of compounds in sentence reading, Stites et al. (2016) report enhanced late positivities for letter transpositions relative to non-transposed baselines. The effects of letter transpositions did not differ for transpositions within constituents and across constituent boundaries, suggesting that in this paradigm, full-form access offers the best explanation for the findings. Huang et al. (2020) report findings from a cross-modal priming study in Chinese. Primes were opaque compounds in a sentence context. Morphological priming was observed with neutral sentences, but not with sentences biasing toward the opaque meaning. These findings suggest that the extent to which compound constituents are accessed during sentence processing is influenced by the sentence context.

---

[5] While some studies deal with changes of perceived animacy/agency, e.g., Nieuwland and Van Berkum (2006), these studies are specifically designed to override a noun's intrinsic animacy via an elaborate discourse context.

[6] Compounding is not limited to two nouns; examples for longer German compounds range from everyday words like *Rechts.schutz.versicherungs.makler*, law.protection.insurance.broker, 'legal defense insurance broker'; to unusually long examples like '*Weihnachts.baum.schmuck.ausstattungs.spezial.geschäft*', literally 'christmas.tree.decoration.equipment.special.shop', to be found on a shopfront in the Swiss city of Basel; see https://de.wikipedia.org/wiki/Johann_Wanner_Unternehmer for illustration.

[7] Our work is rooted in the combinatorial tradition of research on the mental lexicon. Please refer to the Supplementary material for a brief discussion of links to learning-based models of lexical access.

[8] For the sake of readability and brevity, we limit this literature overview mainly to work on compound recognition, omitting the vast literature on the processing of derived complex words. An overview of the general parallels and differences is given in Leminen et al. (2019).

In contrast, many studies have shown at least some amount of decomposition for compounds during word recognition, often as a function of semantic transparency.

Libben et al. (2003) report repetition priming for both first and second constituents as speeding up compound recognition times, arguing for routine decomposition in both transparent and opaque compounds. Lemhöfer et al. (2011) showed that reaction times for Dutch compounds with (orthographically incorrect) orthotactic cues between morphemes were faster than those for (orthographically correct) compounds without these cues, further supporting the notion that lexical access via constituents is routinely performed, and that it eases processing, rather than adding to processing cost [for details concerning similar findings in other languages, see Dronjic (2011) and Cui et al. (2013) for Chinese, Bertram and Hyönä (2003) and Hyönä (2012) for Finnish, and Smolka and Libben (2017) for German, a.o.]. A recent line of research has made use of reduction or enhancement of the mismatch negativity (MMN) amplitude in an oddball paradigm. MMN amplitude is reduced during combinatorial processing, but enhanced during full-form lexical access, making it a valuable tool for researching decomposition. For Chinese, Tsang et al. (2022) report a reduction of the MMN amplitude for transparent compounds relative to pseudocompounds, but equal MMN amplitudes for opaque and pseudocompounds. Zou et al. (2023) report an MMN amplitude reduction relative to the pseudoword baseline for low-frequency compounds, but not for high-frequency compounds. The authors of the respective studies explain their findings as showing that Chinese compounds are routinely decomposed. For transparent and for low-frequency compounds, combinatorial processing seems to be the dominant way of lexical access. For opaque and for high-frequency compounds, the MMN reduction from combinatorial processing is canceled out by the MMN enhancement due to full-form access, with both effects canceling each other out and leading to similar MMN amplitudes as in the processing of pseudowords.

One point supporting the idea of early decomposition is the fact that constituent frequency has an impact on compound recognition in a number of different languages [Juhasz et al. (2003), Andrews et al. (2004), Fiorentino and Poeppel (2007), Wang et al. (2010), and MacGregor and Shtyrov (2013) for English, Duñabeitia et al. (2007) for Basque and Spanish, Kuperman et al. (2009) for Dutch, Bronk et al. (2013) for German, and Hyönä and Pollatsek (1998) and Pollatsek et al. (2000) for Finnish]. The general direction of effects is a compound benefit, i.e., a processing advantage for compounds relative to simple words matched for full-form length and frequency. This strongly suggests that compounds are routinely decomposed during word recognition. It also (again) supports the idea that morphological decomposition is not *per se* costly, at least not so much as to override the processing benefit from more easily accessible (highly-frequent) constituents.

The above literature illustrating processing benefits for compounds depending on lexical properties of the constituents draws on data from a variety of languages and methods. Fiorentino and Poeppel (2007) found faster reaction times and different MEG signatures for compounds relative to simple words matched for length and frequency [see Crepaldi et al. (2013) and Fiorentino et al. (2014) for additional behavioral and EEG evidence in favor of compound decomposition]. MacGregor and Shtyrov (2013) manipulated frequency and transparency to investigate whether constituents are accessed during compound recognition, concluding that transparent compounds are accessed combinatorially with constituent and full-form properties both influencing lexical access, while high-frequency opaque compounds are accessed via their full form.

Some of the studies supporting morphological decomposition suggest a special role for the second constituent. Duñabeitia et al. (2007) monitored the processing of Basque compounds, manipulating compound headedness (unlike Germanic languages, but like, e.g., Italian, Basque allows both right- and left-headed compounds, allowing to disentangle effects of position from those of headedness). The authors found facilitation only for second, but not first constituents, interpreting their findings as showing routine decomposition which is however blind to semantics. In another study on Basque compounds, Vergara-Martínez et al. (2009) used EEG measurements, manipulating the frequency of constituents and the compound headedness. They found that the N400 amplitude was larger for low- than high-frequency second constituents, while evidence for an influence of first constituent frequency was less clear.

Additional evidence for a privileged position of the second constituent in German compounds (where position is confounded with headedness) is provided by Holle et al. (2010), who report larger N400 amplitudes when heads (rather than modifiers) are exchanged for non-words.

Strong evidence in favor of lexical access to both heads and modifiers is presented by Bronk et al. (2013). In a series of lexical decision task experiments, they tested the recognition of German compounds against simple words matched for full-form length and frequency. Compounds came in two conditions, one with a highly frequent modifier, and the other with a low-frequency modifier. Results showed that compounds with high-frequency modifiers elicited shorter reaction times than compounds with low-frequency constituents or simple words. This finding was robust for both semantically transparent and opaque compounds; however, for opaque compound only, the constituent benefit was lost in the presence of difficult rather than easy pseudowords (i.e., with nonexisting combinations of two existing nouns). The authors describe this as evidence for early morphological decomposition, before access to the semantics of the full form, arguing against models assuming full-form access instead of or before decomposition. The findings also strongly support lexical access to modifiers.

In sum, there is ample evidence for morphological decomposition of noun-noun compounds, beginning early during word recognition. Lexical constituent properties like frequency and length influence processing cost, showing that the lexical entries of the constituents are accessed during compound recognition. However, the question remains whether semantic constituent properties are routinely accessed during compound recognition in a similar way to lexical constituent properties, and whether they have an influence on compound processing.

Compared to the vast literature on morphological decomposition of complex words, the literature on semantic decomposition is still smaller, and studies tend to focus on different aspects of semantic constituent properties. One approach is to focus on the influence of semantic transparency, comparing the processing of transparent vs. opaque complex words. Early priming studies (Sandra, 1990; Zwitserlood, 1994 for Dutch) report semantic priming of constituent meanings for transparent, but not for opaque compounds. While this suggests some amount of semantic access to constituents for transparent compounds, the findings were also interpreted as evidence against automatic full decomposition for all types of compounds, since opaque compounds seem to not be connected to their constituents at the semantic level (see also Pratarelli, 1995 for additional influences of length in English).

To assess the role of semantic transparency in derived words, Smolka et al. (2014) and Smolka and Eulitz (2018) used German complex verbs in a series of priming experiments. Verbs included both non-separable prefix verbs like *ver.stehen* ("to understand") and separable particle verbs like *auf.stehen* ("to stand up"); like compounds, these complex verbs can be semantically transparent or opaque. They consistently found that priming from the verb stems was comparable for semantically opaque and transparent complex verbs, suggesting that the lexical representation of complex verbs is accessed via the verb base, irrespective of whether this verb base contributes to the full-form semantics of the complex verb. Koester et al. (2007) investigated the processing of acoustically presented German compounds using EEG. In their stimuli, the gender of the full form and the first constituent were either congruent or incongruent (in the German three-gendered system); this was manipulated for semantically transparent and opaque compounds. For incongruent gender only, they found an increase in the amplitude of the left anterior negativity (LAN), interpreted as evidence of morphological decomposition. Relative to opaque compounds, transparent compounds showed an increased negativity with a centroparietal maximum that occurred during the presentation of the head constituent. The authors interpreted their findings as showing semantic integration of constituents that had previously been accessed separately, arguing that transparent, but not opaque compounds need to be semantically integrated, which incurs additional processing costs. (These additional processing costs due to semantic integration for transparent compounds are not usually discussed in the literature reporting benefits for compounds relative to simple words; see above).

In a follow-up study, Koester et al. (2009) used German compounds consisting of three constituents, manipulating the plausibility of the second and third constituents. Implausible third constituents led to increased N400 amplitudes, as did implausible second constituents. The authors interpret these increased N400 amplitudes as showing the difficulty of lexical integration for implausible constituent combinations; furthermore, they argue that their findings show incremental lexical integration as morphologically complex words unfold. In a series of six lexical decision tasks, Ji et al. (2011) monitored the processing of English compounds that were semantically transparent (e.g., *rosebud*) or opaque (e.g., *hogwash*). In their experiments, they manipulated the likelihood of semantic decomposition (e.g., by adding easy or difficult to spot pseudowords, or by separating the two constituents by empty spaces or color markings). Like preceding studies, they found a compound processing advantage relative to length- and frequency-matched simple words. This advantage was initially visible for both transparent and opaque compounds, but held up only for transparent compounds when decomposition was encouraged. The authors interpret their findings as supporting semantic composition, with the opacity disadvantage showing a conflict between different potential meanings of opaque words.

In two behavioral experiments, Marelli and Luzzatti (2012) investigated the processing of Italian compounds, manipulating headedness, semantic transparency, and constituent frequency. Their results show that both constituent frequency influences recognition, and interacts with full-form properties and semantic transparency. The authors argue for an extension of multiple-route models to include explicit pathways for early semantic processing. Their findings were supported by Arcara et al. (2014) reporting increased processing cost for head-final compared to head-initial Italian compounds, visible in an enhanced LAN component.

In sum, the literature points to a certain amount of routine decomposition, or put differently, to direct access to the constituents of complex words during word recognition. This is strikingly visible in the compound benefit, i.e., a processing advantage of compounds relative to simple words matched for length and frequency, if the compounds contain highly accessible constituents. This has been shown for lexical constituent properties like frequency (Bronk et al., 2013). However, it is still unclear if semantic constituent properties are also accessed during compound recognition. The existing literature on semantic decomposition deals with semantic properties that only apply in the context of the compound word, like headedness (in languages like Italian or Basque) or semantic transparency. While this line of research has added important insight to our understanding of compound processing, the manipulations in the stimulus material always concerned semantic contributions of constituents to the full-form meaning. This type of semantic property is not a semantic property of the constituent noun *per se*, and is unlikely to be part of its lexical entry. Therefore, if we aim to answer the question whether semantic constituent properties play a role during compound recognition (in parallel to the role played by lexical constituent properties like frequency), we need to manipulate an intrinsic semantic constituent property that is independent of the compound context and allows us to measure semantic constituent access directly.

This is where the animacy effects outlined above can add important insight: Animacy is an intrinsic semantic property of the constituents, and can be manipulated independently for constituents and the full form. To illustrate, the compounds *Wolfshund* ("wolfhound") and *Schlittenhund* ("sled dog") both refer to animates, as their lexical heads refer to animates. These words should be expected to have a processing advantage over compounds referring to inanimates, like *Pferdedecke* ("horse rug") and *Tischdecke* ("table cloth"). However, for *Wolfshund*, both the modifier and the lexical head are animate, whereas for *Schlittenhund*, the modifier is inanimate. This leads us to formulate the following general research hypotheses:

1. If semantic constituent properties play a role in compound recognition, we would expect a processing advantage for *Wolfshund* (full-form animate with animate modifier) over

*Schlittenhund* (full-form animate with inanimate modifier). In a similar vein, we would expect a processing advantage of *Pferdedecke* (full-form inanimate with animate modifier) over *Tischdecke* (full-form inanimate with inanimate modifier).

2. However, if lexical, but not semantic constituent properties play a role in compound recognition, we would expect no processing advantage for *Wolfshund* over *Schlittenhund*, since both full forms refer to animates. Neither would we expect a processing advantage for *Pferdedecke* over *Tischdecke*, since both full forms refer to inanimates. Instead, we would expect to see a clear processing advantage of full-form animates (*Wolfshund* and *Schlittenhund*) over full-form inanimates (*Pferdecke* and *Tischdecke*), without any influence of modifier animacy.

## 2.3. Research questions and hypotheses

The current study is designed to answer the research question whether semantic constituent properties are accessed during compound recognition. To this end, we monitor single word recognition of simple words and compounds in a lexical decision task using EEG measurements.[9] Based on the literature, we assume that the most reliable indicator of the processing cost associated with lexical accessibility (in general and to compound constituents) is the N400 amplitude [see Kutas and Federmeier (2000, 2011) for lexical accessibility in general and Vergara-Martínez et al. (2009) and Holle et al. (2010) for constituent accessibility in particular]:

- For single words, we assume a straightforward link between animacy and N400 amplitude - animate simple words should elicit reduced N400 amplitudes than inanimate simple words. This comparison serves as our control to replicate basic findings from the literature and ensure that our measurements are sensitive enough to spot processing differences between existing words brought on by semantic factors.
- For compounds, we assume a link between lexical accessibility and N400 amplitude.
    - If lexical constituent properties are *not* accessed during compound recognition, we expect the N400 amplitude to reflect full-form animacy, which is identical with the animacy of the head.
    - If lexical constituent properties are accessed during compound recognition, we expect the N400 amplitude to reflect both the animacy of the lexical head (identical with full-form animacy) and the modifier.

## 3. Language materials

Language materials consisted of one set of simple words (the control conditions) and another set of compounds (the critical conditions). All words were German nouns. Words were interspersed with non-words resembling simple words and compounds. Non-words followed the rules of German phonotactics and orthography, but at the same time were not designed to be particularly difficult to spot or to contain existing words as their constituents. Examples of simple pseudowords include *Schapf* or *Lofer*; examples of compound pseudowords include *Bopfhalz* or *Pluserfeun* (none of these words have a meaning in German, and neither do the pseudo-constituents *Bopf, Halz, Pluser,* and *Feun*). The simple word set had 40 simple words per condition (80 in total) interspersed with 80 simple pseudowords. The compound word set had 40 compounds per condition (160 in total) interspersed with 160 compound pseudowords.

Simple words came in two conditions, inanimate or animate. Animates referred to animals, but not to humans or professions. Inanimates referred to concrete objects, never to abstract concepts. Compounds came in four conditions, named for the animacy of the modifier and the animacy of the head (in this order): inanimate-inanimate, animate-inanimate, inanimate-animate, animate-animate. Full-form animates (conditions inanimate-animate and animate-animate) referred to animals, but not to humans or professions. Full-form inanimates (conditions animate-inanimate and animate-animate) referred to concrete objects, never to abstract concepts. Likewise, animate constituents always referred to animals. Inanimate constituents referred to concrete objects.[10]

Results for simple words were not meant to be compared directly to results for compound words. For this reason, the matching described below was performed for both stimulus sets separately.

### 3.1. Length and frequency matching

Frequencies were accessed from the dlexdb corpus described in Heister et al. (2011) (access: April 2022).[11]

---

[9] First tentative evidence in favor of semantic constituent access comes from behavioral studies in the context of two qualification theses at the University of Konstanz (Strinzel, 2014; Gozebina, 2017). These studies used a visual lexical decision task, following the general approach that was used in Bronk et al. (2013) to monitor lexical decomposition. The general direction of findings was that both modifier and head animacy influenced reaction times, supporting the idea that semantic constituent properties are accessed during recognition. However, due to issues with stimulus control and the small scale of the studies, the findings are not conclusive. The goal in our current study is to monitor the extent to which semantic constituent properties have an influence on lexical accessibility during compound recognition, using a carefully controlled stimulus set and EEG measurements.

[10] We avoided words with abstract referents, since contrasts between abstract- and concrete-referring words are known to influence N400 amplitude (West and Holcomb, 2000; Huang and Federmeier, 2015).

[11] We chose to not control for family size, i.e., the numbers of nouns that modifiers and heads combined with to form other noun-noun compounds. Certain types of noun-noun compounds are either over- or underrepresented in the current version of the DWDS corpus (to our knowledge the most extensive German corpus for our purposes), making the available data unreliable. Since family size is unlikely to have a strong influence on N400 amplitudes (Kwon et al., 2012), we are confident that this will not be a major concern.

### 3.1.1. Simple words

Simple words and non-words were matched for length in characters [words = 6.60, non-words = 6.65, $t_{(79)} = 0.52$, $p > 0.6$]. Animate and inanimate words were matched for length in characters [animate = 6.65, s.d. = 1.59, inanimate = 6.55, s.d. = 1.48, $t_{(39)} = 0.22$, $p > 0.8$], and lemma frequency [animate = 250.28, s.d. = 238.06, inanimate = 245.20, s.d. = 183.20, $t = -0.62$, $p > 0.5$].

### 3.1.2. Compounds

Full forms of compound words and non-words were matched for length in characters [words = 10.19, s.d. = 1.66, non-words = 9.81, s.d. = 1.38, $t_{(159)} = 2.50$, $p > 0.01$].[12]

Matching for compound words was performed using 2 × 2 ANOVAS with the factors MODIFIER and HEAD animacy. Compound words were matched for full-form length in characters (animate-animate = 10.05, s.d. = 1.55, animate-inanimate = 10.50, s.d. = 1.26, inanimate-animate = 10.20, s.d. = 1.94, inanimate-inanimate =10.03, s.d. = 1.82, no statistically significant differences). They were also matched for full-form lemma frequency [animate-animate = 26.15, s.d. = 38.08, animate.inanimate = 24.27, s.d. = 38.12, inanimate-animate = 30.82, s.d. = 43.84, inanimate-inanimate = 33.23, s.d. = 41.51; MODIFIER $F_{(1,56)} = 1.09$, $p > 0.2$; HEAD $F_{(1,56)} = 2.64$, $p > 0.1$; MODIFIER:HEAD $F_{(1,56)} = 0.17$, $p > 0.6$].

In addition, compound words were matched for lengths and lemma frequencies of heads and modifiers. For modifiers, there were no significant effects and interactions of MODIFIER and HEAD on length [mean values: animate-animate 4.75, s.d. = 1.08, animate-inanimate = 4.90, s.d. = 1.08, inanimate-animate = 5.05, s.d. = 1.13, inanimate-inanimate = 5.00, s.d. =1.06; MODIFIER $F_{(1,56)} = 1.50$, $p > 0.2$; HEAD $F_{(1,56)} = 0.17$, $p > 0.6$; MODIFIER:HEAD $F_{(1,56)} = 0.31$, $p > 0.5$] and lemma frequency [mean values: animate-animate = 3905.07, s.d. = 13684.31, animate-inanimate = 3078.18 s.d. = 3408.76, inanimate-animate = 3282.30, s.d. = 2740.30, inanimate-inanimate = 3240.22, s.d. = 3467.06; MODIFIER $F_{(1,56)} = 2.30$, $p > 0.1$; HEAD $F_{(1,56)} = 2.35$, $p > 0.1$; MODIFIER:HEAD $F_{(1,56)} = 2.58$, $p > 0.1$].

For heads, there were no significant main effects or interactions of MODIFIER and HEAD on length [mean values: animate-animate = 4.75, s.d. = 1.08, animate-inanimate = 4.92, s.d. = 0.97, inanimate-animate = 4.95, s.d. = 1.58, inanimate-inanimate = 5.03, s.d. = 1.13; MODIFIER $F_{(1,56)} = 0.29$, $p > 0.5$; HEAD $F_{(1,56)} = 0.85$, $p > 0.3$; MODIFIER:HEAD $F_{(1,56)} = 0.01$, $p > 0.9$] and lemma frequency [mean values: animate-animate = 1741.03, s.d. = 2217.59, animate-inanimate = 1748.03, s.d. = 2205.65, inanimate-animate = 2158.55, s.d. = 3219.34, inanimate-inanimate = 1945.70, s.d. = 1679.16; MODIFIER $F_{(1,56)} = 1.86$, $p > 0.1$; HEAD $F_{(1,56)} = 1.97$, $p > 0.1$; MODIFIER:HEAD $F_{(1,56)} = 0.12$, $p > 0.7$].

In addition, compounds were matched for Levenshtein neighborhood sizes, extracted from the dlexdb corpus (see Laszlo and Federmeier, 2011 for the link between neighborhood sizes and N400 amplitude). We collected the numbers of higher-frequency neighbors (HF neighbors) and the total number of neighbors (all neighbors) for constituents and full-forms. For full-forms of compounds, there were only 27 items of the 160 which had 1 higher-frequency neighbor. The remaining had none. We therefore refrained from an analysis of full-form neighborhood sizes.

For modifiers, there were no significant effects or interactions of head animacy and modifier animacy for the mean number of all neighbors. For the mean number of more highly frequent neighbors, there was a main effect of modifier animacy [$F_{(1,156)} = 5.7$, $p < 0.5$]. For noun-noun compounds with animate modifiers, the mean number of HF neighbors was 2.4. For noun-noun compounds with inanimate modifiers, the mean number of HF neighbors was 1.4 (numbers only take into account the modifiers that did have HF neighbors). We are confident that this small numerical difference does not put us at risk of a systematic confound. (Inanimate-inanimate: mean HF neighbors = 1.6, mean all neighbors 32.7, 24 items had more highly frequent neighbors; animate-inanimate: mean HF neighbors = 2.0, mean all neighbors = 33.0, 28 items had more highly frequent neighbors; inanimate-animate: mean HF neighbors = 1.2, mean all neighbors = 2.8, 29 items had more highly frequent neighbors; animate-animate: mean HF neighbors = 2.8, mean all neighbors = 36.7, 20 items had more highly frequent neighbors).

For heads, there were no significant effects or interactions of head animacy and modifier animacy. This held for both the mean numbers of more highly frequent neighbors and the mean numbers of all neighbors. (Inanimate-inanimate: mean HF neighbors = 2.3, mean all neighbors = 36.1, 28 items had HF neighbors; animate-inanimate: mean HF neighbors = 3.1, mean all neighbors = 36.4, 31 items had HF neighbors; inanimate-animate: mean HF neighbors = 3.3, mean all neighbors = 34.8, 34 items had HF neighbors; animate-animate: mean HF neighbors = 2.9, mean all neighbors = 36.5, 34 items had HF neighbors).

## 3.2. Familiarity

In a prestudy, the stimuli were rated for familiarity by 10 monolingually raised native German speakers (seven male, three female, mean age = 26.5 years, s.d. = 5.64 years, max = 34 years, min = 19 years). Ratings were elicited on a four-paint scale containing the ratings *kenne ich* ('I know [this word]'), *verstehe ich/habe ich schon einmal gehört* ('I understand [this word]/I have heard [this word] before') and *kenne ich nicht* ('I do not know [this word]').

### 3.2.1. Simple words

For words, the mean number of 'I know' ratings was 9.7 (minimum number of 'I know' ratings for any single word was 8); mean number of 'I understand' ratings was 0.1 (max. was 1); mean number of 'I don't know' ratings was 0.2 (maximum number for a single word was 2). For non-words, the mean number of 'I don't know' ratings was 9.2 (minimum number for a single non-word was 6), mean number of 'I understand' ratings was 0.6 (maximum 4), mean number of 'I know' ratings was a 0.3 (maximum 2).

---

[12] The mean length difference between compound words and non-words was smaller than one character, but still reached statistical significance. Since non-words were only added to the experiment to provide a task in lexical decision, but were not meant to be analyzed or compared to words, we decided to accept this small length difference.

### 3.2.2. Compounds

For words, the mean number of "I know" ratings was = 9.5 (with the minimum of a single word being 4); mean number of "I understand" ratings was 0.3 (max = 4); mean number of "I don't know" ratings was = 0.15 (max = 3). For non-words, the mean number of "I know" ratings was 0.1 (max = 2), the mean number of "I understand" ratings was 0.3 (max = 3), and the mean number of "I don't know" ratings was 9.6 (min = 6). This indicates that simple and compound words were familiar to participants, and that simple and compound non-words clearly recognizable as non-words and did not resemble existing words.

## 3.3. Semantic transparency

To ensure that our compound nouns could truly be considered semantically transparent, we conducted a transparency rating study. Compounds were interspersed with 40 filler items, namely, compounds that we expected to be semantically intransparent, to provide participants with a contrast between transparent and opaque compounds (remember that our stimulus set did not contain semantically opaque compounds). Intransparent compound nouns included *Muskelkater* (lit. 'muscle cat', a muscle ache after exercise), *Schlafmütze* (lit. 'sleep hat', i.e., a sleepy head) or *Milchstraße* (lit. 'milk street', i.e., the Milky Way). Transparency ratings were given by 10 monolingually raised native German speakers (mean age 25.4 years, s.d. = 3.9 years, min = 18, max = 30; five male, five female). Ratings were given on a three-point scale with the points 'transparent', 'unsure', and 'not transparent'. For transparent compounds, the mean across conditions for 'transparent' ratings was 8.7 (out of 10 raters). In comparison, for opaque compounds, the mean rating for 'transparent' was 1 and 'not transparent' was 8.05. Thus, all our compounds are clearly rated differently from truly opaque compounds.

# 4. Experiment

## 4.1. Methods

### 4.1.1. Participants

Forty participants were recruited via SONA systems database at the University of Konstanz. All of them were monolingually raised native speakers of German. They had normal or corrected to normal vision and reported no history of psychological or neurological illness. All participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The measurement was conducted in accordance with COVID-19 health safety regulations. All parties involved in the experiment were wearing medical masks and had negative antigen test results not older than 24 h before the arrival to the laboratory. Participants signed an informed consent form and received 25 € compensation for their time. The research was approved by the Research Ethics Committee of the University of Konstanz under the ethics approval number 05/2021. During data analysis, one participant was excluded due to poor data quality. The remaining 39 participants had ages ranging from 18 to 31 years (mean = 23.2 years, s.d. = 2.8 years). 18 participants were male, 22 participants were female.

### 4.1.2. Procedure

The EEG was recorded with BrainVision Recorder (version 1.24.0001, Brain Products GmbH), with 64 EEG actiCAP slim electrodes, attached to an elastic cap with actiCAP SNAP holders and connected to BrainAmp DC amplifiers. The electrode arrangement was based on the equidistant M43-V1 layout as provided by Easycap GmbH. Horizontal and vertical eye movements were registered by four EOG Ag/AgCl sintered passive ring electrodes, connected to BrainAmp ExG bipolar amplifier. Data were recorded in the frequency range 0.016–250 Hz. Impedance values below 20 kΩ were accepted. The signal was digitized with a sampling rate of 500 Hz. Participants were comfortably seated in a sound-attenuated room in front of a monitor at approximately two meters. They were asked to avoid excessive eye and body movements during the EEG recording. They were instructed to press the right button if a word was presented on the screen, and the left one if there was a non-word. During the measurement, no feedback regarding the correctness of their response was given. The Presentation software by Neurobehavioral Systems Inc. (version 20.2) was used for delivering stimuli and trigger codes. Depending on the button press reaction time, the presentation of stimuli took approximately 18–20 min. It was divided into three runs with short breaks between them.

### 4.1.3. Data preprocessing and analysis

Data were processed using the Brain Vision Analyzer 2 software (Brain Products, Gilching). Raw data were inspected visually, and time windows including strong visible artifacts, as well as breaks, were manually removed. Next, we performed an ICA blink correction using the slope algorithm, followed by filtering (low cutoff 0.5 Hz, high cutoff 40 Hz, 50 Hz notch filter) and topographic interpolation via triangulation for channels showing long stretches of noisy data. After interpolation, all electrodes were re-references to average reference. An Automatic Raw Data Inspection was performed for the re-referenced data (maximal allowed voltage step: 50 $\mu$/ms; maximal allowed difference: 100 $\mu$V/200 ms; minimal/maximal allowed amplitudes 200 $\mu$V/$-$200 $\mu$V; lowest allowed activity: 0.5 $\mu$V/100 ms). Data were segmented starting at 100 ms before stimulus onset and ending at 800 ms after stimulus onset. A baseline correction was performed for 100 ms before stimulus onset. Averages were calculated per participant for all conditions. Participants with less than 35 trials in one of the six experimental conditions were excluded from the data analysis, leading to the exclusion of one participant.

We exported mean amplitudes per condition for each participant from the time window from 350 to 450 ms. This time window was chosen based on the literature, as well as the fact that the peak of the N400 component was close to or around 400 ms, supporting our assumption that the classical N400 is a relevant component to monitor lexical access.

For analysis a subset of 25 electrodes was selected. Electrode position was coded by assigning electrodes to five medial-lateral as well as five anterior-posterior positions. Medial-lateral positions were: lateral-left (front to back: F7, FC5, C5, P5, PO7), medial-left (front-to-back: F3, FC3, C3, CP3, PO3), midline (front-to-back: Fz, FCz, Cz, CPz, Pz), medial-right (front-to-back: F4, FC4, C4, CP4, PO4), and lateral-right (front-to-back: F8, FC6, C6, P6, PO8). Anterior-posterior positions were: anterior (left to right: Fz, F3, Fz, F4, F8), medial-anterior (left to right: CF5, FC3, FCz, FC4, FC6); medial (left to right: C5, C3, Cz, C4, C6), posterior-medial (P5, CP3, CPz, CP4, P6), posterior (left to right: PO7, PO3, Pz, PO4, PO8).

We performed a repeated measures ANOVA of the mean voltages in the selected electrode sites. Only voltages elicited by words were analyzed. For simple words, we performed a repeated-measures ANOVA monitoring the main effects and interactions of within-subjects factors ANIMACY, ANTERIOR-POSTERIOR position, and MEDIAL-LATERAL position. For compounds, we performed a repeated-measures ANOVA monitoring the main effects and interactions of within-subjects factors MODIFIER animacy, HEAD animacy, ANTERIOR-POSTERIOR position, and MEDIAL-LATERAL position. Statistical analyses were performed in a hierarchical fashion, i.e., only statistically significant interactions were pursued, unless specifically mentioned otherwise. Interactions were resolved if they included at least one of the experimental factors (ANIMACY for simple words, MODIFIER or HEAD for compounds). A Huyhn-Feldt correction was performed when the degree of freedom in the numerator was higher than 1. Original degrees of freedom and corrected probability levels are reported. Analyses were performed in R (R Development Core Team, 2019) using the ezANOVA function of the ez package (Lawrence, 2011).

## 4.2. Results

For the sake of readability, we only report the highest interactions involving the experimental factors, followed by the resolution of these interactions.

### 4.2.1. Simple words

Descriptively speaking, waveforms were more negative-going for the inanimate than for the animate condition. This was most visible over central and posterior sites. The negativity for inanimate conditions was clearly visible around 400 ms, and persisted up until around 600 ms; the later negativity was more pronounced in posterior than central sites.

In the time window from 350 to 450 ms, there was an interaction of ANIMACY and MEDIAL-LATERAL [$F_{(4,152)} = 1.31$, $p < 0.01$, HF $\varepsilon = 0.59$; ANIMACY significant in medial-lateral regions medial-left ($p < 0.001$), midline ($p < 0.001$), medial-right ($p < 0.05$)].

A graphic depiction of grand averages for selected electrode sites and voltage difference maps is given in Figures 1, 2 shows mean voltage amplitudes for both conditions at electrode site Cz.

### 4.2.2. Compounds

Descriptively speaking, waveforms for the inanimate-inanimate condition were more negative-going than for the animate-animate condition. This was most visible over central and posterior sites; the negativity was clearly visible around 400 ms and persisted until about 600 ms. While the general pattern was similar to findings for the simple nouns, the amplitude differences for compounds were rather smaller. The waveforms for the mixed conditions animate-inanimate and inanimate-animate ran mostly together, between the inanimate-inanimate and animate-animate waveforms. The general pattern was that the number of animate constituents was reflected in the amplitude, with waveforms going more positive for each animate constituent. There was no visible influence of the type of constituent that was animate (no stronger influence of head and thereby full-form animacy compared to modifier animacy).

In the time window from 350 to 450 ms, there was a significant main effect of HEAD [$F_{(1,38)} = 11.20$, $p < 0.01$], and an interaction of MODIFIER and ANTERIOR-POSTERIOR position [$F_{(4,152)} = 3.85$, $p < 0.05$, $\varepsilon = 0.39$; MODIFIER significant in anterior-posterior regions medial ($p < 0.01$), posterior-medial ($p < 0.01$), and posterior ($p < 0.05$)].[13]

A graphic depiction of grand averages for selected electrode sites and voltage difference maps is given in Figures 3, 4 shows mean voltage amplitudes for all four conditions at electrode site Cz.

## 5. Discussion and conclusion

For simple words, the N400 had a larger amplitude for inanimates than for animates. This fits findings from the literature indicating that lexical access is more costly for inanimate compared to animate nouns that are matched on lower-level factors like frequency and length (Janyan and Andonova, 2011; Nairne et al., 2013; Bonin et al., 2014, 2015, 2019; VanArsdall et al., 2015, 2017; Bugaiska et al., 2019), and that this reduced processing cost for animates surfaces as a reduced N400 amplitude (Proverbio et al., 2007). The difference in N400 amplitudes already becomes visible in a simple lexical decision task, without additional tasks like semantic categorization (as in Proverbio et al., 2007) needed. This shows that the influence of animacy on lexical accessibility is robust even in routine single-word processing, and that the N400 amplitude is an informative measure to tap into this.

For compounds, both head animacy (which corresponds to full-form animacy) and modifier animacy influence the N400 amplitude. The amplitude differences are smaller than for simple words, but reach statistical significance. Generally speaking, the N400 amplitude is least negative-going for animate-animate compounds, and most negative-going for inanimate-inanimate compounds. Compounds with one animate and one inanimate constituent show an N400 that tends to run between these two extremes. Neither a descriptive overview nor the statistical analysis suggest an interaction between modifier and head animacy. At this point, the facilitating effect of constituent animacy seems to be additive - the higher the proportion of animate constituents, the less negative-going the N400 amplitude will be. Descriptively, the

---

[13] Following the recommendation of an anonymous reviewer, we reanalyzed the data with a reduced participant set, excluding the data from three participant with a high number of interpolated channels (8 or 9 out of 64 channels). While reducing the participant set reduces the statistical power and turned the interactions with the multi-level topographical factors into marginally significant ones, the main effect of modifier now became significant. This shows that the effect of modifier animacy cannot be reduced to artifacts introduced by overly interpolated data.

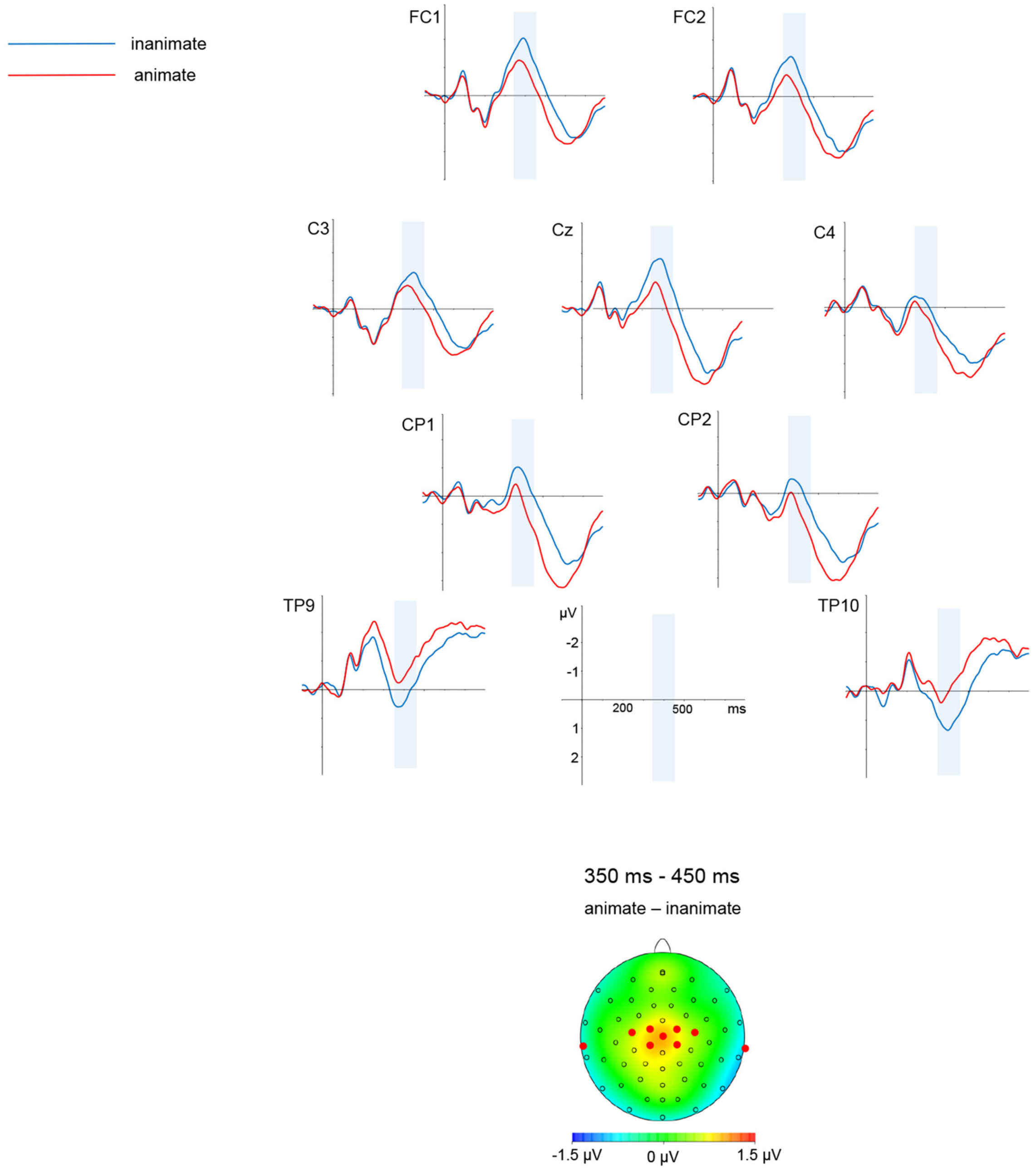


FIGURE 1
Simple nouns. Grand average ERPs for selected electrode sites and a difference map are shown. A mean voltage difference map (animate minus inanimate) for the marked time window from 350 to 450 ms is given on the left side. The electrodes selected for illustration are marked in the maps.

distribution both head and modifier effects fits with the usual N400 topography. However, the interaction with topographical factors was only significant for modifier animacy.

Our findings strongly support the idea that constituent properties of (transparent) compounds are routinely accessed during compound recognition [see, a.o., Pollatsek et al. (2000), Juhasz et al. (2003), Andrews et al. (2004), Fiorentino and Poeppel (2007), Kuperman et al. (2009), Wang et al. (2010), and in particular Bronk et al. (2013)].

Unlike some earlier studies (Duñabeitia et al., 2007; Vergara-Martínez et al., 2009; Holle et al., 2010), our current findings do not support a privileged role for the second constituent, which in German always coincides with the lexical head. While the influence of head animacy was more widespread than the role of modifier animacy, both were present and statistically significant in the N400 time window. The absence of an interaction between modifier and head animacy further supports the idea that during the investigated time window, N400 amplitude simply reflects the added accessibility, which is mainly influenced by

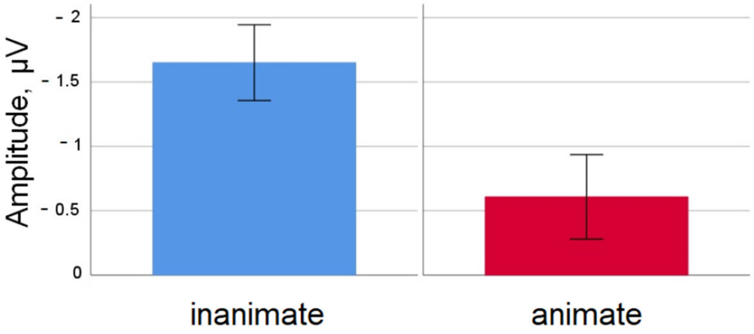


FIGURE 2
Simple nouns. Mean amplitude difference across participants for the inanimate and animate condition at electrode Cz in the 350–450 ms time window. Error bars depict standard errors of the mean.

the accessibility of the constituents, without visible overriding influences of head/full-form animacy. Discrepancies to these earlier studies may stem from the differences in possible headedness in Basque and Spanish [in the case of Duñabeitia et al. (2007) and Vergara-Martínez et al. (2009)], or due to differences in stimulus design (our own stimuli did not contain opaque compounds, while there is no mention for control of constituent properties across conditions in Holle et al., 2010).

Most importantly in the context of our research question, our findings strongly support the idea that semantic properties of both heads and modifiers influence lexical accessibility. This is in contrast to some earlier studies arguing against semantic access to compound constituents (Duñabeitia et al., 2007) during decomposition, but fits the wider literature describing some amount of semantic (de)composition for compounds [see Sandra (1990), Zwitserlood (1994), and Koester et al. (2007, 2009) for transparent compounds, Smolka et al. (2014) and Smolka and Eulitz (2018) for both transparent and opaque compound verbs, and Ji et al. (2011) for a nuanced discussion of semantic composition in transparent and opaque noun-noun compound recognition]. Unlike these early studies, our findings provide evidence for the direct influence of intrinsic semantic constituent properties via our manipulation of constituent animacy, allowing a direct comparison to manipulations of lexical constituent properties (like e.g., frequency in Bronk et al., 2013). Our studies thus support the earlier findings on lexical decomposition, but circumvent some of the problems stemming from the indirect approach to semantics (via, e.g., semantic transparency; see MacGregor and Shtyrov, 2013 and also Koester et al., 2007 for insightful discussions on the difficulties of interpreting absent priming effects for opaque compounds, since they could either reflect the fact that semantic constituent priming does not happen in opaque compounds because they are accessed differently, or that it does not happen because prime and target are not semantically related by virtue of the compound being opaque.).

Importantly, our results were found in a context that did not particularly encourage semantic or morphological decomposition, and in comparison of words to "easy" to spot pseudowords that respected the rules of German orthography and phonotactics, but did not need careful reading and did not combine existing with non-existing constituents. We therefore interpret our findings as showing that access to semantic constituent properties is an automatic and routine process during the recognition of semantically transparent compounds (in contrast, to e.g., Stites et al., 2016 in the more recent literature).

With respect to models of word recognition, our findings would fit into full-parsing models (e.g., Taft and Forster, 1975; Libben et al., 1999 and follow-up models), but also into dual/multiple route models (Caramazza et al., 1988; Baayen and Schreuder, 1999; Kuperman et al., 2009), since they allow constituent access for transparent compounds. A prerequiste is that the models allow for early access to semantic constituent properties in addition to lexical ones, supporting the demands in Marelli and Luzzatti (2012) for formulating explicit extensions for early semantic processing pathways in the context of existing multiple-route models.

Future studies should address issues of timing differences between full-form and constituent access to elucidate if one precedes the other. It would also be informative to monitor semantic composition in semantically opaque compounds using the N400 amplitude as a direct measure of the ease of lexical access, and constituent animacy manipulations to influence the relative difficulty of said access. However, this would also imply a careful control of potential confounding factors like the semantic relation between modifier and head across conditions, and elegant solutions to disentangle absent effects from multiple overlaying effects for semantically opaque compounds (see Ji et al., 2011 for a detailed discussion of this issue).

For future studies on the effects of headedness and constituent position (in continuation of Duñabeitia et al., 2007; El Yagoubi et al., 2008; Vergara-Martínez et al., 2009; Marelli and Luzzatti, 2012; Arcara et al., 2014), animacy manipulations may provide a useful tool to monitor semantic access and its interactions with headedness in languages like Basque or Italian, allowing a more direct approach than the one taken in previous studies via manipulations of semantic transparency.

In general, our findings show the strong influence that constituent animacy has on lexical accessibility in compound recognition, visible in related neurolinguistic

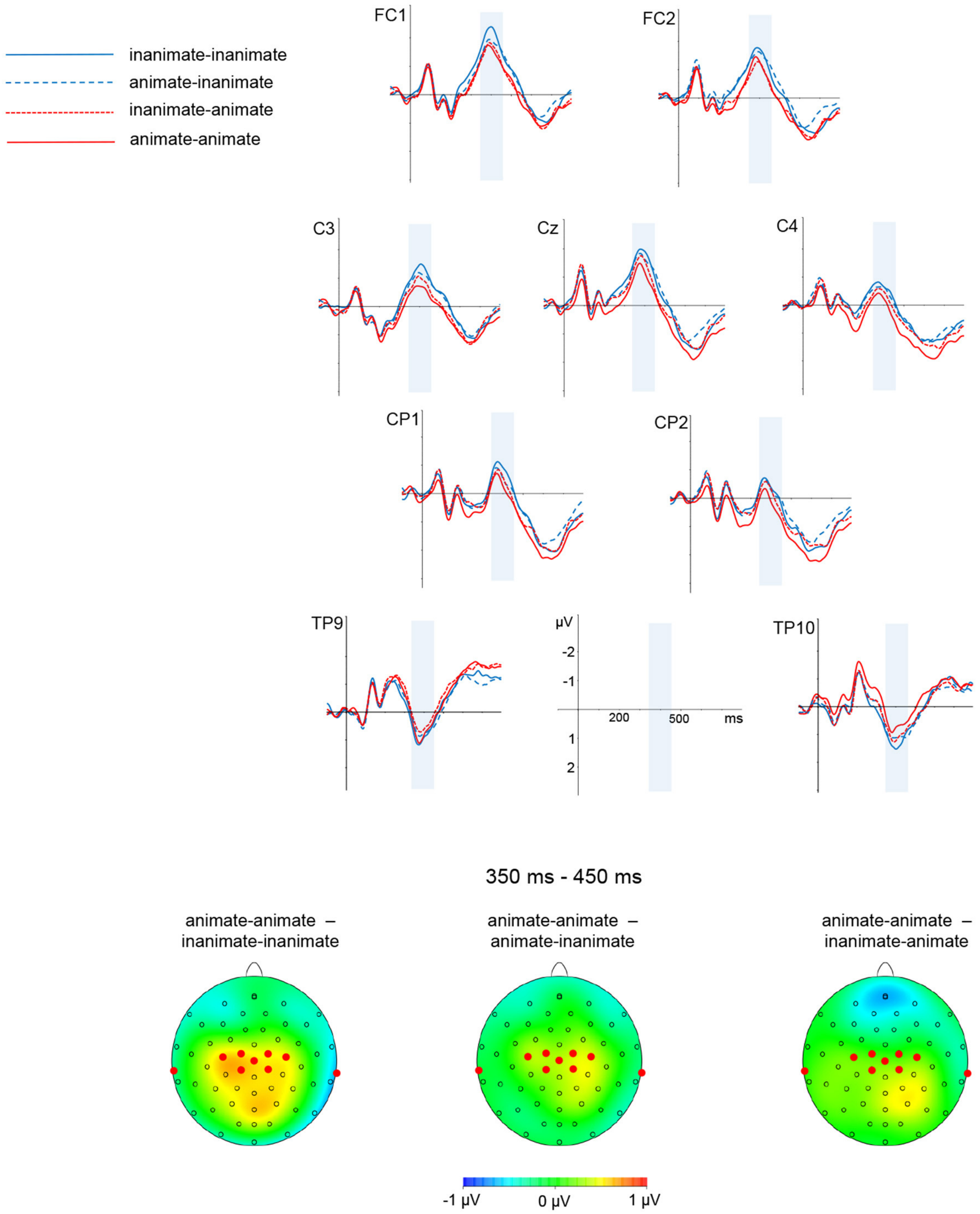


FIGURE 3
Noun-noun compounds. Grand average ERPs for selected electrode sites and difference maps are shown. Mean voltage difference map (animate-animate minus each of the other conditions) for the marked time window from 350 to 450 ms is given on the left side. The electrodes selected for illustration are marked in the maps.

measures. Irrespective of specific research questions, this shows that along with lexical constituent properties like length and frequency, semantic constituent properties like animacy need to be carefully controlled and/or balanced in studies investigating compound processing to avoid losing significant effects, or even end up with spurious effects in severely unbalanced stimulus sets. In a similar vein, in sentence processing research, it is advisable to control for morphological complexity of words in comparable positions across conditions, taking into account full-form and constituent

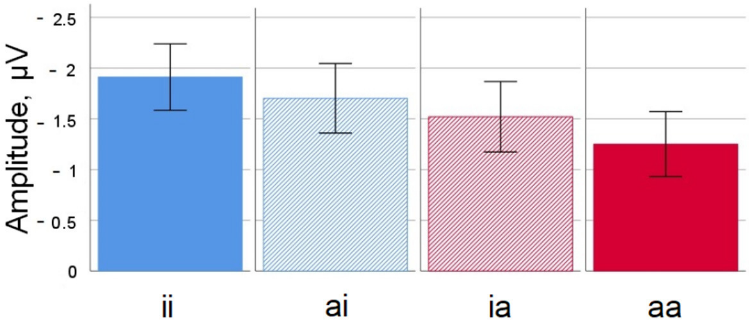


FIGURE 4
Noun-noun compounds. Mean amplitude difference across participants for all four conditions at electrode Cz in the 350–450 ms time window. Error bars depict standard errors of the mean. ii, inanimate-inanimate; ai, animate-inanimate; ia, inanimate-animate; aa, animate-animate.

properties alike, to avoid contamination from complex single-word recognition effects.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors upon request, without undue reservation.

## Ethics statement

The studies involving humans were approved by Ethikkommission der Universität Konstanz (Ethics Committee of the University of Konstanz). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. No potentially identifiable images or data are presented in this study.

## Author contributions

AC was responsible for experiment conception and design, running data preprocessing and statistical analysis, as well as for overseeing data acquisition. MK was responsible for discussing choices in data preprocessing and analysis and for overseeing data acquisition. CE was responsible for experiment conception, discussing choices in data preprocessing and analysis, as well as for providing laboratory infrastructure. All authors contributed to writing the manuscript.

## Funding

AC and MK were supported by a Grant from the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) within projects EU 39/8-2 and BA 1178/10-2, FOR2111. Student assistant support and participant compensation were supported by the Grant EU 39/9-1 awarded to CE, as well as Grants from the University of Konstanz awarded to CE and AC.

## Acknowledgments

The authors wish to thank Philine Link, Daniela Kaupp, Kornelia Nowińska, Gizem Aslan, and Zacharias Satow for student assistant support during data acquisition and special thanks are due to Zacharias Satow for student assistant support with stimulus construction and pretests. We are grateful to Oleksiy Bobrov for technical and programming support. This work is part of an ongoing series of studies at our department that were informed by work in a number of qualification theses and our thanks are due to our former students Michaela Strinzel, Christina Gozebina, and Ayfer Uygun.

## 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.

## Publisher's note

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1173352/full#supplementary-material

# References

Andrews, S., Miller, B., and Rayner, K. (2004). Eye movements and morphological segmentation of compound words: there is a mouse in mousetrap. *Eur. J. Cogn. Psychol.* 16, 285–311. doi: 10.1080/09541440340000123

Anzellotti, S., Mahon, B. Z., Schwarzbach, J., and Caramazza, A. (2011). Differential activity for animals and manipulable objects in the anterior temporal lobes. *J. Cogn. Neurosci.* 23, 2059–2067. doi: 10.1162/jocn.2010.21567

Arcara, G., Marelli, M., Buodo, G., and Mondini, S. (2014). Compound headedness in the mental lexicon: an event-related potential study. *Cogn. Neuropsychol.* 31, 164–183. doi: 10.1080/02643294.2013.847076

Baayen, R. H., and Schreuder, R. (1999). War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain Lang.* 68, 27–32.

Bamyacı, E., Häussler, J., and Kabak, B. (2014). The interaction of animacy and number agreement: an experimental investigation. *Lingua* 148, 254–277. doi: 10.1016/j.lingua.2014.06.005

Bertram, R., and Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: evidence from eye movements when reading short and long Finnish compounds. *J. Mem. Lang.* 48, 615–634. doi: 10.1016/S0749-596X(02)00539-9

Bonin, P., Gelin, M., and Bugaiska, A. (2014). Animates are better remembered than inanimates: further evidence from word and picture stimuli. *Mem. Cogn.* 42, 370–382. doi: 10.3758/s13421-013-0368-8

Bonin, P., Gelin, M., Dioux, V., and Méot, A. (2019). "It is alive!" Evidence for animacy effects in semantic categorization and lexical decision. *Appl. Psycholinguist.* 40, 965–985. doi: 10.1017/S0142716419000092

Bonin, P., Gelin, M., Laroche, B., Méot, A., and Bugaiska, A. (2015). The "how" of animacy effects in episodic memory. *Exp. Psychol.* 62, 371–384. doi: 10.1027/1618-3169/a000308

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., et al. (2011). Think globally: cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain Lang.* 117, 133–152. doi: 10.1016/j.bandl.2010.09.010

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2006). The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychol. Rev.* 113, 787–821. doi: 10.1037/0033-295X.113.4.787

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: a cross-linguistic approach. *Lang. Linguist. Compass* 3, 19–58. doi: 10.1111/j.1749-818X.2008.00099.x

Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal–ventral stream model of sentence comprehension. *Brain Lang.* 125, 60–76. doi: 10.1016/j.bandl.2013.01.010

Bossong, G. (1985). *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, G. (1991). "Differential object marking in Romance and beyond," in *New Analyses in Romance Linguistics*, eds D. A. Kibbee and D. Wanner (Amsterdam; Philadelphia, PA: Benjamins), 143–170.

Branigan, H. P., Pickering, M. J., and Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118, 172–189. doi: 10.1016/j.lingua.2007.02.003

Bronk, M., Zwitserlood, P., and Bölte, J. (2013). Manipulations of word frequency reveal differences in the processing of morphologically complex and simple words in German. *Front. Psychol.* 4, 546. doi: 10.3389/fpsyg.2013.00546

Bugaiska, A., Grégoire, L., Camblats, A.-M., Gelin, M., Méot, A., and Bonin, P. (2019). Animacy and attentional processes: evidence from the Stroop task. *Q. J. Exp. Psychol.* 72, 882–889. doi: 10.1177/1747021818771514

Butterworth, B. (1983). "Lexical representation," in *Language Production II: Development, Writing and Other Language Processes*, ed B. Butterworth (London: Academic Press), 257–294.

Bybee, J. (1995). Regular morphology and the lexicon. *Lang. Cogn. Process.* 10, 425–455.

Capitani, E., Laiacona, M., Mahon, B., and Caramazza, A. (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cogn. Neuropsychol.* 20, 213–261. doi: 10.1080/02643290244000266

Caramazza, A., Laudanna, A., and Romani, C. (1988). Lexical access and inflectional morphology. *Cognition* 28, 297–332.

Caramazza, A., and Mahon, B. Z. (2003). The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361. doi: 10.1016/S1364-6613(03)00159-1

Caramazza, A., and Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *J. Cogn. Neurosci.* 10, 1–34.

Corbett, G. G. (2000). *Number*. Cambridge: Cambridge University Press.

Crepaldi, D., Rastle, K., Davis, C. J., and Lupker, S. J. (2013). Seeing stems everywhere: position-independent identification of stem morphemes. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 510. doi: 10.1037/a0029713

Croft, W. (1990). *Typology and Universals*. Cambridge: Cambridge University Press.

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., and Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: an eye-movement-contingent display change study. *Q. J. Exp. Psychol.* 66, 527–547. doi: 10.1080/17470218.2012.667423

Czypionka, A. (2014). *The Interplay of Object Animacy and Verb Class in Representation Building* (dissertation). Humboldt-Universität zu Berlin, Berlin, Germany.

Czypionka, A., and Eulitz, C. (2018). Lexical case marking affects the processing of animacy in simple verbs, but not particle verbs: evidence from event-related potentials. *Glossa J. Gen. Linguist.* 3. doi: 10.5334/gjgl.313

Dronjic, V. (2011). Mandarin Chinese compounds, their representation, and processing in the visual modality. *Writ. Syst. Res.* 3, 5–21. doi: 10.1093/wsr/wsr005

Duñabeitia, J. A., Perea, M., and Carreiras, M. (2007). The role of the frequency of constituents in compound words: evidence from Basque and Spanish. *Psychon. Bull. Rev.* 14, 1171–1176. doi: 10.3758/BF03193108

El Yagoubi, R., Chiarelli, V., Mondini, S., Perrone, G., Danieli, M., and Semenza, C. (2008). Neural correlates of Italian nominal compounds and potential impact of headedness effect: an ERP study. *Cogn. Neuropsychol.* 25, 559–581. doi: 10.1080/02643290801900941

Fiorentino, R., Naito-Billen, Y., Bost, J., and Fund-Reznicek, E. (2014). Electrophysiological evidence for the morpheme-based combinatoric processing of English compounds. *Cogn. Neuropsychol.* 31, 123–146. doi: 10.1080/02643294.2013.855633

Fiorentino, R., and Poeppel, D. (2007). Compound words and structure in the lexicon. *Lang. Cogn. Process.* 7, 953–1000. doi: 10.1080/01690960701190215

Frisch, S., and Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *Neuroreport* 12, 3391–3394. doi: 10.1097/00001756-200110290-00048

Giraudo, H., and Grainger, J. (2001). Priming complex words: evidence for supralexical representation of morphology. *Psychon. Bull. Rev.* 8, 127–131. doi: 10.3758/BF03196148

Gozebina, C. (2017). *Der Einfluss der Konstituentenbelebtheit auf die Verarbeitung deutscher Nominalkomposita unter Berücksichtigung der semantischen Relation, [The influence of constituent animacy on the processing of German nominal compounds, considering the role of semantic relation]* (M.A. thesis). University of Constance, Konstanz, Germany.

Hagoort, P. (2007). "The memory, unification, and control (MUC) model of language," in *Automaticity and Control in Language Processing*, eds A. Meyer, L. Wheeldon, and A. Krott (Hove; New York, NY: Psychology Press), 261–288.

Hagoort, P. (2016). "MUC (Memory, Unification, Control): a model on the neurobiology of language beyond single word processing," in *Neurobiology of Language*, eds G. Hickock and S. L. Small (Cambridge, MA: Elsevier), 339–347.

Haspelmath, M. (2013). "Occurrence of nominal plurality," in *The World Atlas of Language Structures Online*, eds M. Dryer and M. Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology).

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., and Geyken, A. (2011). dlexDB — eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau* 32, 10–20. doi: 10.1026/0033-3042/a000029

Holle, H., Gunter, T. C., and Koester, D. (2010). The time course of lexical access in morphologically complex words. *Neuroreport* 21, 319–323. doi: 10.1097/WNR.0b013e328335b3e0

Huang, H.-W., and Federmeier, K. D. (2015). Imaginative language: what event-related potentials have revealed about the nature and source of concreteness effects. *Lang. Linguist.* 16, 503–515. doi: 10.1177/1606822X15583233

Huang, J., Tsang, Y.-K., Xiao, W., and Wang, S. (2020). Morphosemantic activation of opaque Chinese words in sentence comprehension. *PLoS ONE* 15, e0236697. doi: 10.1371/journal.pone.0236697

Hyönä, J. (2012). The role of visual acuity and segmentation cues in compound word identification. *Front. Psychol.* 3, 188. doi: 10.3389/fpsyg.2012.00188

Hyönä, J., and Pollatsek, A. (1998). Reading Finnish compound words: eye fixations are affected by component morphemes. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1612.

Janyan, A., and Andonova, E. (2011). "Word naming and picture naming in Bulgarian: contribution of experiential and semantic variables to the naming RTs," in *European Perspectives on Cognitive Science*, eds B. Kokinov, A. Karmiloff-Smith, and N. Nersessian (New Bulgarian University Press).

Ji, H., Gagné, C. L., and Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *J. Mem. Lang.* 65, 406–430. doi: 10.1016/j.jml.2011.07.003

Juhasz, B. J., Starr, M. S., Inhoff, A. W., and Placke, L. (2003). The effects of morphology on the processing of compound words: evidence from naming, lexical decisions and eye fixations. *Brit. J. Psychol.* 94, 223. doi: 10.1348/000712603321661903

Koester, D., Gunter, T. C., and Wagner, S. (2007). The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain Lang.* 102, 64–79. doi: 10.1016/j.bandl.2006.09.003

Koester, D., Holle, H., and Gunter, T. C. (2009). Electrophysiological evidence for incremental lexical-semantic integration in auditory compound comprehension. *Neuropsychologia* 47, 1854–1864. doi: 10.1016/j.neuropsychologia.2009.02.027

Ković, V., Plunkett, K., and Westermann, G. (2009). Shared and/or separate representations of animate/inanimate categories: an ERP study. *Psihologija* 42, 5–26. doi: 10.2298/PSI0901005K

Krifka, M. (2009). "Case syncretism in German feminines: typological, functional and structural aspects," in *On Inflection*, eds M. Krifka and P. Steinkrüger (Berlin: Mouton de Gruyter), 141–172.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Res.* 1146, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kuperman, V., Schreuder, R., Bertram, R., and Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 876. doi: 10.1037/a0013484

Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Kwon, Y., Nam, K., and Lee, Y. (2012). ERP index of the morphological family size effect during word recognition. *Neuropsychologia* 50, 3385–3391. doi: 10.1016/j.neuropsychologia.2012.09.041

Laszlo, S., and Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology* 48, 176–186. doi: 10.1111/j.1469-8986.2010.01058.x

Lawrence, M. A. (2011). *ez: Easy Analysis and Visualization of Factorial Experiments*. R package version 3.0-0. Available online at: https://www.rdocumentation.org/packages/ez/versions/3.0-1/topics/ezANOVA

Lemhöfer, K., Koester, D., and Schreuder, R. (2011). When bicycle pump is harder to read than bicycle bell: effects of parsing cues in first and second language compound reading. *Psychon. Bull. Rev.* 18, 364–370. doi: 10.3758/s13423-010-0044-y

Leminen, A., Smolka, E., Dunabeitia, J. A., and Pliatsikas, C. (2019). Morphological processing in the brain: the good (inflection), the bad (derivation) and the ugly (compounding). *Cortex* 116, 4–44. doi: 10.1016/j.cortex.2018.08.016

Levelt, W. J. (1993). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Libben, G., Derwing, B. L., and de Almeida, R. G. (1999). Ambiguous novel compounds and models of morphological parsing. *Brain Lang.* 68, 378–386.

Libben, G., Gibson, M., Yoon, Y. B., and Sandra, D. (2003). Compound fracture: the role of semantic transparency and morphological headedness. *Brain Lang.* 84, 50–64. doi: 10.1016/S0093-934X(02)00520-5

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676.

MacGregor, L. J., and Shtyrov, Y. (2013). Multiple routes for compound word processing in the brain: evidence from EEG. *Brain Lang.* 126, 217–229. doi: 10.1016/j.bandl.2013.04.002

Malchukov, A. L. (2008). Animacy and asymmetries in differential case marking. *Lingua* 118, 203–221. doi: 10.1016/j.lingua.2007.02.005

Marelli, M., and Luzzatti, C. (2012). Frequency effects in the processing of Italian nominal compounds: modulation of headedness and semantic transparency. *J. Mem. Lang.* 66, 644–664. doi: 10.1016/j.jml.2012.01.003

McKinnon, R., Allen, M., and Osterhout, L. (2003). Morphological decomposition involving non-productive morphemes: ERP evidence. *Neuroreport* 14, 883–886. doi: 10.1097/00001756-200305060-00022

Næss, Å. (2004). What markedness marks: the markedness problem with direct objects. *Lingua* 114, 1186–1212. doi: 10.1016/j.lingua.2003.07.005

Nairne, J. S., VanArsdall, J. E., Pandeirada, J. N., Cogdill, M., and LeBreton, J. M. (2013). Adaptive memory the mnemonic value of animacy. *Psychol. Sci.* 24, 2099–2105. doi: 10.1177/0956797613480803

Nieuwland, M. S., and Van Berkum, J. J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *J. Cogn. Neurosci.* 18, 1098–1111. doi: 10.1162/jocn.2006.18.7.1098

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Opfer, J. E., and Gelman, S. A. (2010). "Development of the animate-inanimate distinction," in *The Wiley-Blackwell Handbook of Childhood Cognitive Development, 2nd Edn.*, ed U. Goswami (Chichester: Wiley-Blackwell), 213–238.

Paczynski, M., and Kuperberg, G. R. (2011). Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb argument processing. *Lang. Cogn. Process.* 26, 1402–1456. doi: 10.1080/01690965.2011.580143

Pollatsek, A., Hyönä, J., and Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 820. doi: 10.1037/0096-1523.26.2.820

Pratarelli, M. E. (1995). Modulation of semantic processing using word length and complexity: an ERP study. *Int. J. Psychophysiol.* 19, 233–246.

Proverbio, A. M., Del Zotto, M., and Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neurosci.* 8, 24. doi: 10.1186/1471-2202-8-24

R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna.

Rutherford, M., Pennington, B. F., and Rogers, S. J. (2006). The perception of animacy in young children with autism. *J. Autism Dev. Disord.* 36, 983–992. doi: 10.1007/s10803-006-0136-8

Sandra, D. (1990). On the representation and processing of compound words: automatic access to constituent morphemes does not occur. *Q. J. Exp. Psychol.* 42, 529–567.

Schreuder, R., and Baayen, H. R. (1995). "Modeling morphological processing," in *Morphological Aspects of Language Processing*, ed L. Feldman (Hillsdale, NJ: Lawrence Erlbaum), 131–154.

Sitnikova, T., West, W. C., Kuperberg, G. R., and Holcomb, P. J. (2006). The neural organization of semantic memory: electrophysiological activity suggests feature-based segregation. *Biol. Psychol.* 71, 326–340. doi: 10.1016/j.biopsycho.2005.07.003

Smolka, E., and Eulitz, C. (2018). Psycholinguistic measures for German verb pairs: semantic transparency, semantic relatedness, verb family size, and age of reading acquisition. *Behav. Res. Methods* 50, 1540–1562. doi: 10.3758/s13428-018-1052-5

Smolka, E., and Libben, G. (2017). "Can you wash off the hogwash?"–semantic transparency of first and second constituents in the processing of German compounds. *Lang. Cogn. Neurosci.* 32, 514–531. doi: 10.1080/23273798.2016.1256492

Smolka, E., Preller, K. H., and Eulitz, C. (2014). "Verstehen" ("understand") primes "stehen" ("stand"): morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *J. Mem. Lang.* 72, 16–36. doi: 10.1016/j.jml.2013.12.002

Stites, M. C., Federmeier, K. D., and Christianson, K. (2016). Do morphemes matter when reading compound words with transposed letters? Evidence from eye-tracking and event-related potentials. *Lang. Cogn. Neurosci.* 31, 1299–1319. doi: 10.1080/23273798.2016.1212082

Strinzel, M. (2014). *Der Einfluss der Kategorie Belebtheit auf die Worterkennung von Komposita, 'The influence of the category of animacy on word recognition of compounds'* (B.A. thesis). University of Konstanz, Konstanz, Germany.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Q. J. Exp. Psychol. Sec. A* 57, 745–765. doi: 10.1080/02724980343000477

Taft, M., and Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *Mental Lexicon* 1, 183–199. doi: 10.1075/ml.1.2.02taf

Taft, M., and Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *J. Verb. Learn. Verb. Behav.* 14, 638–647.

Taft, M., and Nguyen-Hoan, M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Lang. Cogn. Process.* 25, 277–296. doi: 10.1080/01690960903043261

Trueswell, J. C., Tanenhaus, M. K., and Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *J. Mem. Lang.* 33, 285–318.

Tsang, Y.-K., Zou, Y., and Tse, C.-Y. (2022). Semantic transparency in Chinese compound word processing: evidence from mismatch negativity. *Neuroscience* 490, 216–223. doi: 10.1016/j.neuroscience.2022.03.007

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Cogdill, M. (2015). Adaptive memory: animacy effects persist in paired-associate learning. *Memory* 23, 657–663. doi: 10.1080/09658211.2014.916304

VanArsdall, J. E., Nairne, J. S., Pandeirada, J. N., and Cogdill, M. (2017). A categorical recall strategy does not explain animacy effects in episodic memory. *Q. J. Exp. Psychol.* 70, 761–771. doi: 10.1080/17470218.2016.1159707

Vergara-Martínez, M., Duñabeitia, J. A., Laka, I., and Carreiras, M. (2009). Erp correlates of inhibitory and facilitative effects of constituent frequency in compound word reading. *Brain Res.* 1257, 53–64. doi: 10.1016/j.brainres.2008.12.040

Verhoeven, E. (2014). Thematic prominence and animacy asymmetries: evidence from a cross-linguistic production study. *Lingua* 143, 129–161. doi: 10.1016/j.lingua.2014.02.002

Verkhlyutov, V., Mar'ina, I., and Strelets, V. (2014). Event-related spectral EEG power on presentation of pseudowords and words of the "animate" and "inanimate" categories. *Neurosci. Behav. Physiol.* 44, 16–23. doi: 10.1007/s11055-013-9866-1

Wang, M., Lin, C. Y., and Gao, W. (2010). Bilingual compound processing: the effects of constituent frequency and semantic transparency. *Writing Syst. Res.* 2, 117–137. doi: 10.1093/wsr/wsq012

Weckerly, J., and Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology* 36, 559–570.

West, W. C., and Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *J. Cogn. Neurosci.* 12, 1024–1037. doi: 10.1162/089892900511 37558

Yamamoto, M. (1999). *Animacy and Reference: A Cognitive Approach to Corpus Linguistics*. Amsterdam; Philadelphia, PA: John Benjamins Publishing Company.

Zou, Y., Tsang, Y.-K., Shum, Y.-H., and Tse, C.-Y. (2023). Full-form vs. combinatorial processing of Chinese compound words: evidence from mismatch negativity. *Int. J. Psychophysiol.* 187, 11–19. doi: 10.1016/j.ijpsycho.2023.02.004

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Lang. Cogn. Process.* 9, 341–368.

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 07 September 2023
DOI 10.3389/fpsyg.2023.1168739

OPEN ACCESS

EDITED BY
Michael J. Serra,
Texas Tech University, United States

REVIEWED BY
Johannes Schultz,
University Hospital Bonn, Germany
Giorgio Vallortigara,
University of Trento, Italy
Josefa N. S. Pandeirada,
University of Aveiro, Portugal
Sara Felix,
University of Aveiro, Portugal, in collaboration with reviewer JP

*CORRESPONDENCE
Sajjad Torabian
✉ torabias@uci.edu

RECEIVED 18 February 2023
ACCEPTED 25 July 2023
PUBLISHED 07 September 2023

CITATION
Torabian S and Grossman ED (2023) When shapes are more than shapes: perceptual, developmental, and neurophysiological basis for attributions of animacy and theory of mind. *Front. Psychol.* 14:1168739.
doi: 10.3389/fpsyg.2023.1168739

# When shapes are more than shapes: perceptual, developmental, and neurophysiological basis for attributions of animacy and theory of mind

Sajjad Torabian* and Emily D. Grossman

Visual Perception and Neuroimaging Lab, Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, United States

Among a variety of entities in their environment, what do humans consider alive or animate and how does this attribution of animacy promote development of more abstract levels of mentalizing? By decontextualizing the environment of bodily features, we review how physical movements give rise to perceived animacy in Heider-Simmel style animations. We discuss the developmental course of how perceived animacy shapes our interpretation of the social world, and specifically discuss when and how children transition from perceiving actions as goal-directed to attributing behaviors to unobservable mental states. This transition from a teleological stance, asserting a goal-oriented interpretation to an agent's actions, to a mentalistic stance allows older children to reason about more complex actions guided by hidden beliefs. The acquisition of these more complex cognitive behaviors happens developmentally at the same time neural systems for social cognition are coming online in young children. We review perceptual, developmental, and neural evidence to identify the joint cognitive and neural changes associated with *when* children begin to mentalize and *how* this ability is instantiated in the brain.

KEYWORDS

**social cognition, cognitive development, animacy, agency, theory of mind, motion perception, Heider and Simmel, default-mode network**

## 1. Introduction

In their seminal work on apparent behavior, Heider and Simmel (1944) showed that when humans viewed a two-dimensional animation of simple geometric shapes, their interpretations of the movements tended not toward a physical story. Instead, people perceived the shapes as animated beings and agents, and described their observation in rather abstract terms. For example, when a triangle vibrated in proximity to another triangle, people saw the two as agents who engaged in a social interaction such as fighting. A line of studies followed the work of Heider and Simmel, showing how motion alone can turn objects into living beings. A single frame of an example Heider-Simmel like animation with two interacting shapes is depicted in Figure 1. Recently Ratajska et al. (2020) designed an extended range of social plots to demonstrate that simple shapes of various types, not just triangles and circles, can depict rich narratives beyond conflict interactions, even on a brief timescale (13–23 s compared to 2 1/2 min).

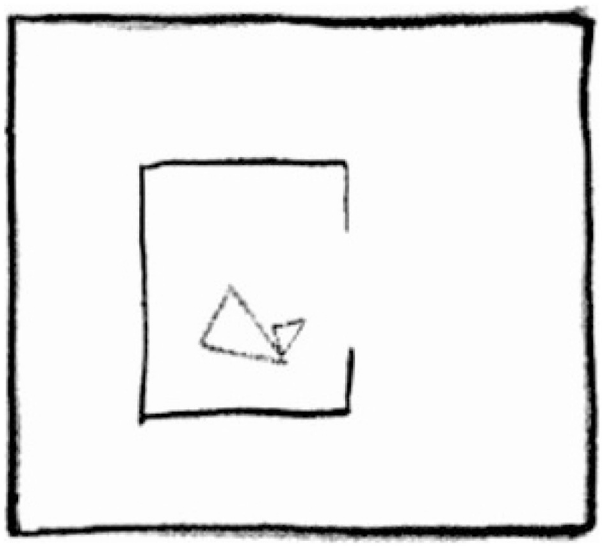


FIGURE 1
Single frame of Heider-Simmel animation designed by Castelli et al. (2000), depicting a mother who persuades child to go out, conveying theory of mind. Adapted with permission.

Often cast within the broad framework of theory of mind (ToM, the ability to attribute mental states to others, which are inferred and therefore unobservable, and can be used to make predictions about the behaviors of others, Premack and Woodruff, 1978), the perception of animacy, interactivity and goal directed behaviors derived from Heider-Simmel type animations reflect the human tendency to construct social interpretations and derive inferences about beliefs and desires from movement patterns alone (Baker et al., 2017). In this review, we will discuss perceptual, developmental, and neural underpinnings of perceived animacy and social attributions. Specifically we seek to link the development of neural systems to the ability to draw upon perceptual cues for animacy in order to establish more complex beliefs about the goals of others. We also discuss the evidence that detection of animacy, and to some extent the ability to discern goal-directed behaviors, is not uniquely human. We couch our discussion within the framework that the behavior of animate objects can be broadly categorized into goal-directed and mentalistic (see Gergely and Csibra, 2003, but also Schaafsma et al., 2015 for a systematic deconstruction of theory of mind), motivated by the psychological principle of rational action. This principle states that a bias exists to interpret behaviors as goal-oriented, guided by environmental constraints and mental states, the latter of which will be inferred under the assumption that the agent is performing efficient actions.

To better disentangle goal-directed and mentalistic representations, we review (1) the developmental literature as it offers clear perspectives into how children acquire rich mental representations of the social world around them, as well as (2) findings in monkeys, apes, and chicks. We will discuss neural systems supporting goal-directed and mentalistic representations in adults and the development of those brain systems in children under age two when these cognitive systems come online. In this review we focus on research that employs Heider-Simmel type animations that are deprived of many of the explicit cues that typically signal animacy, the determination that it is appropriate to apply psychological reasoning to a given entity (Csibra et al., 1999), and agency, the capacity to engaged in intrinsically motivated (goal-directed) behavior. This approach is particularly valuable because it is accessible to adults, children and non-humans alike, while also decoupling animacy from the perception of species-specific cues, such as faces and eye gaze. In the following section we first lay the groundwork for studying social behavior, by discussing perceptual cues that give rise to animacy.

## 2. The perceptual determinants of perceived animacy

There are many cues in our environment that signal animacy, intention and goals, including eye gaze, head tilt, facial expressions and body movements (Chang and Troje, 2008). Heider-Simmel type animations are devoid of all of these cues, and nonetheless give rise to the perception of animacy, which lies at the foundation of mental state attributions (Schultz and Frith, 2022). Initiation of movement, change in speed, and change in direction (particularly to avoid a barrier) are all examples of such cues that are readily and reflexively interpreted as signaling animacy (Stewart, 1982; Tremoulet and Feldman, 2000). Each of these features shares the property of self-propulsion, velocity changes that are initiated without physical contact, which manifests perceptually as a property of an animate creature. Stewart (1982) describes this core factor as motion that violates Newtonian laws, which is described by Scholl and Tremoulet (2000) as "hidden energy" possessed by animate bodies. In contrast, if an object travels in a consistent direction with sustained movement, or changes direction as a consequence of contact with another object, observers are typically not left with the impression of animacy (Stewart, 1982).

It is important to note that a single object on a featureless background was used in Tremoulet and Feldman (2000), which shows that animacy does not require the presence of other entities. Even with a single object, stimulus changes that are self-induced—and therefore consistent with a hidden energy—can trigger the attribution of animacy. Tremoulet and Feldman (2000) showed that when a short line segment travels along a straight line and changes direction without realigning its orientation to its new path, it is less likely to be perceived as animate compared to when it does realign. The same researchers demonstrated that circles, or more generally non-pointed shapes, are similarly perceived as less animate than shapes that are able to exhibit rotations, even if they traverse the same trajectory.

Much in the same way that eye gaze signals the intentional state of others, Gao et al. (2010) demonstrated the power of oriented features in shapes to convey complex mental states such as predatory desires. In these "wolfpack" demonstrations, arrows oriented toward a target are perceived as having intent directed at the target (as in wolves toward a sheep), even when the movements of the objects themselves were completely random. Computational modeling indicates that the attributions adults make when viewing these chasing animations reflects super-additive gains from the integration of high-level attentive tracking with salient perceptual cues for animacy (Gao et al., 2019). In the coming section we will discuss how an object, after showing cues of animacy, can behave in meaningful social ways.

## 3. Toward attributions of social behavior

An animate object can interact with the environment, for instance by wandering around another animate object, at varying levels of complexity. It has been debated whether understanding social interactivity requires high-level reasoning. Shu et al. (2018) addressed this question with decontextualized stimuli from real-life aerial videos of moving people. Observers more often rated the dynamic, decontextualized scenes as interactive rather than non-interactive or unsure, indicating there are critical visual motion cues between items that give rise to the perception of interactivity. Consistent with that hypothesis, the authors developed a computational model that lacked explicit high-level intentions and goals that nonetheless accurately predicted human judgements. This finding is consistent with the notion of directedness of interactions in driving perception in simple animations.

In contrast, Rasmussen and Jiang (2019) maintained that both low-level motion characteristics *and* high-level reasoning contribute to people's judgements of social interaction in Heider-Simmel animations. They based this conclusion, in part, on the observation that perceived interactivity differs when viewing the vignettes in forward vs. reverse. The ability to capture the influence of higher-order inferences in the forward-played movies, which was weaker when viewed in reverse, indicates an important factor of extended time-dependent, narrative-like contextual cues present in Heider-Simmel animations. Confirmation that more elaborate narratives are associated with more abstract inferences also comes from computational work in which models that incorporate contextual information in addition to object trajectory cues better fit measures of human action recognition (Roemmele et al., 2016).

Simple shapes can also elicit more complex attributions about thematic content of events, and the animations themselves may evoke emotional states in the viewers. When asked to categorize the narratives depicted by simple 3-dimensional animations of moving objects into film genres, people can consistently do so, identifying themes of non-fiction, comedy, drama, and action (Visch and Tan, 2009). Observers also report experiencing sympathy and rooting, for example toward struggling circles—or "underdogs"—that move uphill (Kim et al., 2008).

Because Heider-Simmel animations have the potential to engage more complex mentalizing, these movies have also been considered for use in assessing social intelligence, as an alternative to traditional written tests (Brown et al., 2022). This is particularly valuable to studies of cognitive development, in which children do not yet have the ability to read narratives. This also makes the study of social inferences derived from Heider-Simmel animations particularly valuable for comparative study of theory of mind abilities in non-human species. In the following section we discuss the developmental and comparative evidence for mentalizing abilities in children, non-human primates, monkeys and chicks. In the second half of the review, we will discuss neural evidence to support the behavioral findings.

## 4. Attributions of animacy, goals, and beliefs: a developmental approach

Babies are born with preferentially looking patterns directed toward socially meaningful features, including faces (Morton and Johnson, 1991; Buiatti et al., 2019), the eyes (Farroni et al., 2002), direction of gaze (Batki et al., 2000), biological motion (Simion et al., 2008), and animated shapes that move in accordance with cues for animacy, such as self-propelled motion and speed changes (Di Giorgio et al., 2017, 2021). Orienting toward simple shapes that convey animacy is apparent after only a couple of days of birth. This very early social orienting system is believed to reflect the function of a subcortical and more rudimentary orienting system at birth, which is subject to refinement over the next few years (i.e., Di Giorgio et al., 2016).

### 4.1. Attributions of goals

In parallel to the development of social cognitive systems, infants also possess an intuitive physics (Hespos and vanMarle, 2012) which after 2 months enables them to understand the basic properties of objects, such as solidity, cohesion, and invariance in object size, shape, pattern and color (Baillargeon, 2008). For example, continuity gives infants the expectation that a moving ball will stop when it comes in contact with a wall, and infants will look significantly longer in surprise if the ball passes through the wall (Spelke et al., 1992). With the development of an intuitive physics children understand the physical interactions between non-agent entities. At the same time an intuitive psychology helps infants understand the behaviors of agents. Around the age of 2 months, infants begin to react differently—through smiles and vocalizations—to the facial movements of people vs. to the facial movements of a doll (Legerstee et al., 1987), and by 5 months infants' looking patterns are consistent with attributing goals to the movements of human hands (Woodward, 1998).

Infants can also make goal attributions solely based on variability of behavior, without explicit cues for animacy. For example, when viewing a self-propelled box persistently moving toward a cone, 3-month-olds identify the cone as the goal of the box and will show heightened interest if the box approaches a newly introduced object (Luo, 2011). Around the same age, children can discern the social goals of a simple shape as it facilitates or impedes another shape's goals (Hamlin et al., 2010), and 3 months later show a preference toward the helper as compared to the hinderer (Hamlin et al., 2007). Older studies have also argued for children's ability to perceive goals in Heider-Simmel style animations, although those findings suggested later onset of this competence, at 6.5 (Csibra, 2008) and 9 months (Csibra et al., 1999).

As proposed by Csibra et al. (1999), infants younger than one year are confined to a *teleological* stance by which they see phenomena in terms of purposes. In a review by Saxe et al. (2004) and inspired by Flavell (1988), this stance is described as making direct *connections* between objects. Importantly, in the teleological framework, infants utilize the psychological principle of rational action to understand goal-directed behavior such that an actor will approach its goal through the most efficient means as imposed by the physical environment ["situational" constraints; Gergely and Csibra (2003)]. Figure 2 shows examples of rational and irrational actions as studied by Gergely et al., 1995, with one shape reaching a goal either in presence of a wall or in its absence.

Constraints on rational actions can include the hidden beliefs of agents, which are not directly observable but nonetheless have the potential to guide more complicated actions that may otherwise be perceived as irrational. It is with the development of the *mentalistic* stance that more sophisticated mental reasoning is constructed, which allows the interpretations of more complex actions. This stance is termed as *representations* in Saxe et al. (2004)'s view or subjective experiences attributed to others. Gergely and Csibra (2003) and Saxe et al. (2004) agree in that the teleological stance/connections precede the mentalistic stance/representations in the course of development. They, however, differ in what they regard as "mentalistic". Gergely and Csibra (2003) believe that mental state attributions emerge only after the teleological stance, while Saxe et al. (2004) consider both connections and representations to be mentalistic. We will continue our discussion using the teleological/mentalistic model as it offers a less ambiguous framework.

## 4.2. Attributions of beliefs

Later in development children are able to attribute mental states to agents and understand that they hold subjective experiences of their own, which in turn enables the child to reason about complex actions driven by beliefs (Gergely and Csibra, 2003). Compared to the teleological stance with its components of goals, physical constraints, and actions, this more sophisticated mentalistic stance includes desires, beliefs, and intentions. Moreover, these internal states are interwoven such that desires define goals, beliefs shape implied constraints, and intentions lie behind actions. For example, in a study by Berry and Springer (1993) using motion pictures similar to the original Heider-Simmel animation, a 3-year-old girl reported the following description: "The daddy is chasing the little one around the house. He'll catch him. Well, he didn't catch him, so he got mad and broke the house and that's the end".

As proposed by Gergely and Csibra (2003), the mentalistic stance is also guided by the principle that agents will strive to achieve their desires through the most efficient means.[1] One can therefore make inferences about intentions through mentalistic reasoning when observing behaviors, when desires and beliefs are known. Indeed, given any pair of the triple components of mental states, a prediction on the unknown one can be made.

It was traditionally believed that the mentalizing aspect of theory of mind, and in particular false belief representations, develops around the age of 3.5 (Wellman et al., 2001). However, it has more recently been argued that younger toddlers also possess an understanding of false beliefs. This was demonstrated in a help task experiment by Buttelmann et al. (2009), in which 18-month-olds observed an actor placing a toy under one of two boxes. Another actor then moved the toy to the other box, either in presence of the first actor (true belief condition) or in the actor's absence (false belief condition). In both conditions, the first actor subsequently reached for the empty box. Whereas in the true belief condition the young toddlers helped the actor open the empty box, in the false belief condition the toddlers guided the actor to the correct box. This implies the belief that the actor was seeking the toy but held a false belief about its location. This main effect of underlying belief was demonstrated in another study using the same paradigm in 15-month-olds and measuring looking time and violation of expectation (Onishi and Baillargeon, 2005). Toddlers in this experiment looked longer when the actor returned and reached for the box where the toy actually hid, showing that 15-month-olds expected the actor to choose the box based on her false belief.

There is further evidence that children reason about the internal states of others' minds very early after the first birthday (Surian et al., 2007). The researchers measured the looking time of 13-month-olds as they watched animations of an agent looking for food and found longer fixations toward actions that violated false beliefs. Interestingly, however, a caterpillar played the main role in the animations rather than a human actor, indicating that children's ascription of complex actions to minds is not restricted to humans. Therefore, similar to animacy and goal attributions, higher level attributions of mental states can also occur toward a variety of objects.

More recently, evidence shows that even around 10 months children are capable of representing mental states of others to distinguish between pro- and anti-social behaviors (Hamlin et al., 2013). In a social evaluation task, children observed a puppet show in which a lion showed a preference toward one of two objects, either in the presence or in the absence of two elephants. The elephants then lifted doors to give the lion access to an object. Children preferred the prosocial agent (the elephant that lifted the door to the preferred toy) only when the elephants had seen the lion's initial preference. Otherwise, if the elephants were not present to see the lion's preference, children did not evaluate their subsequent door lifting as pro- or anti-social, and chose one elephant randomly. The 10-month-olds therefore showed preference based on the match between the implied desires and actions of the puppets, showing some understanding of the mental states of others.

Hamlin et al. (2013)'s finding brings theory of mind to the first year of life, although it is not the earliest evidence to do so. We will discuss in the coming section how causality studies might have suggested even an earlier age for the emergence of theory of mind, possibly around 8–10 months (Rochat et al., 2004). As we touched on earlier, it is important to note that these earlier findings of theory of mind at 13 (Surian et al., 2007), 10 (Hamlin et al., 2013), and 8–10 months— discussed next—all involve non-human protagonists, i.e., an animated caterpillar, animal puppets, and

---

[1] It should be noted that rationality as being discussed here pertains to mental models of agency rather than the theory of the rational decision-maker as proposed in other fields including behavioral economics. The traditional rational-agent model in economics, in particular, assumed that human decisions are rational. Prospect theory (Kahneman and Tversky, 1979), on the other hand, has argued that human preferences are often *frame-bound* as opposed to *reality-bound*, and deviations from reality toward subjective frames result in inconsistencies in behavior, and therefore irrationality.

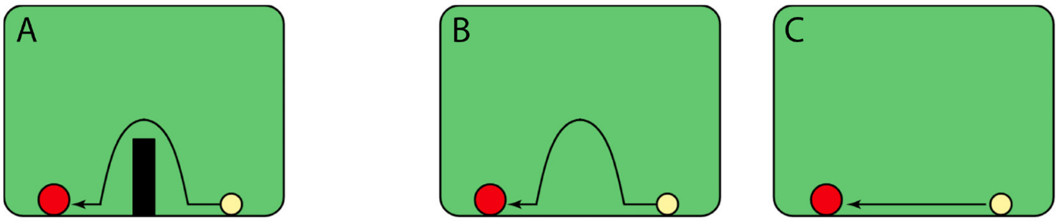


FIGURE 2
Rational goal-directed behavior of a shape moving towards the other by avoiding a barrier (familiarization) **(A)**. **(B, C)** Depict test trials where in the absent of the barrier, the animate object either travels along the same but now inefficient trajectory, or on a straight path to the goal, as expected from a rational teleological stance. Frames adapted with permission from redrawings of Gergely et al. (1995) by Gergely and Csibra (2003).

circles. Whether or not attributing mental states to simpler agents develops earlier compared to attributions toward humans requires further investigation (Carey, 1985). Perceiving goals, however, has been shown to occur earlier with simple geometric shapes (Hamlin et al., 2010; Luo, 2011) than with a human hand (Woodward, 1998) as reviewed in the previous section.

## 4.3. Inferred physical and social causality

In this section, we discuss interactions between objects and focus on events that involve causal inference, as well as on factors that break causal links. Drawing on the teleological-mentalistic framework, we investigate causality as it occurs either within physical constraints (i.e., physical causality) or according to mental states (i.e., social causality). Note that causality can be studied in various forms, for example as sunshine causes a plant to grow, but our focus remains on animacy and proximal and immediate interactions.

We previously discussed how an object is perceived as animate and moves purposefully, which are important first steps in studying causality. Consider an event with circle A moving from rest on a straight path toward a second circle, B. As A hits B, B starts moving on about the same direction, resulting in a causal interaction. This is a case of physical causality and specifically an example of a launching effect, as illustrated in the classic works of Michotte (1963) on perceptual causality. Importantly, the link between the two objects will break if they violate physical laws, for instance if they leave temporal or spatial gaps between them. That is to say, in our launching example, if B begins moving not immediately after the moment of impact, or if there is distance between the stopping point of A and starting point of B, then A is not perceived as physically causing the movement of B. Under these circumstances, the principle of rationality would be unable to explain the event in teleological terms.

What do developmental studies teach us about the attribution of causation by contact and causation at a distance? Would infants perceive social causality if causation occurs at a distance? Spelke et al. (1996) suggested that 6-month-olds might understand that people can interact without contact. Below we review two studies that test this hypothesis on infants under age 1, with stimuli of simple shapes.

Schlottmann and Surian (1999) showed 9-month-olds launching events with two squares, with one moving toward the other and stopping at a distance before the second square moved. Interestingly, and contrary to Michotte (1963)'s predictions, infants derived an impression of causality despite no contact between the shapes. The causal chain, in this situation, did not break with a spatial gap. It is possible that younger children understand causality within physical constraints, and later around 9 months develop an understanding of social causality. Rochat et al. (2004) tested this by directly comparing inferred causality across ages, in an experiment with animated displays of two chasing discs. While never making contact, the chaser moved at a slow but steady pace toward the chasee which accelerated away, and it was programmed to move constantly closer to the chasee without following its path. Hence, the chaser sought the chasee's heat rather than directly following it. This heat-seeking behavior is an essential attribute of the discs in this experiment, because it renders improbable any direct physical connection between them. Three to 10-month-olds participated in this study, but only infants between 8 and 10 months tended to dishabituate to a role reversal between the chaser and the chasee, which shows their understanding of social causation.

As Rochat et al. (2004) stated, "action at a distance is a trademark of social exchanges", and it is around 8–10 months of age that children make a transition into understanding such mentalistic interactions. Before this age, between 3 (Luo, 2011) and 8 months, infants' thinking about others is limited under a teleological stance. Tomasello (2001) had indeed described a transition in social-cognitive development around 9 months (also coined as the "9 month revolution") when children come to understand others as intentional agents, similar to the perceptual-cognitive change reviewed here. In search for the evolutionary origins of this transition, in the coming section we will review studies of goal-directedness and theory of mind in primates, and will then delve into evidence from neuroscience regarding when and how a mind-understanding mind is developed.

## 5. Attributions of goals and beliefs by monkeys and apes

Humans have tended to consider theory of mind a distinctive human capacity, but numerous discoveries in primates suggest that this notion may be a myth. Bonobos and chimpanzees, who diverged from humans about 7 million years ago, are examples of Great Apes who exhibit the ability to attribute animacy to abstract shapes and an understanding of goal-directed behavior. In an experimental design inspired by Hamlin et al. (2010), Krupenye and Hare (2018) showed bonobos animations of two simple shapes engaged in apparent helping or hindering interactions, with an added cue for animacy (eyes) attached to the shapes. Whereas 3-month-old human infants gaze preferentially to prosocial agents, the bonobos' preference was for the hinderer. This finding conforms with bonobos' behavior in real-world scenarios as they choose dominant individuals over subordinates.

Evidence shows that chimpanzees can also attribute goals to objects and this teleological representation is bounded by the principle of rationality. Uller (2004) measured eye gaze in infant chimpanzees using a task first developed by Gergely et al. (1995) for humans (as illustrated in Figure 2). In this task the chimpanzees were familiarized with an animation of a rectangle traveling along a parabolic path to avoid a barrier and reach a circle. In the test phase without the barrier, the chimpanzees then observed the triangle moving either along the same parabolic trajectory (test 1), or on a straight line (test 2). The chimpanzees looked longer at test 1 that depicted the inefficient, irrational parabolic path, evidence that they recognized the goal and also expected the triangle to move straight toward it.

Old World monkeys diverged from the Great Apes about 20–30 million years ago (Wood et al., 2007; Hayashi et al., 2020). The evidence for whether monkeys can also reason about goal-directed behavior is more mixed. In an experiment with Japanese macaques (*Macaca fuscata*, who belong to the Old World family), monkeys were shown animations with two discs that either moved randomly (Figure 3B) or depicted a runner that moved randomly with a chaser that pursued the same trajectory (chasing, Figure 3A; Atsumi et al., 2017). Similar to human observers who were also included in this study, macaques successfully recognized and selected the chasing events to earn food rewards, which is argued to be evidence that monkeys understand goal-directed behavior. Because the monkeys earned food for their selection, however, others have criticized the study as instead reflecting learned associations between certain low-level movement characteristics and reward (Schafroth et al., 2021).

Acknowledging the issue of learned associations, Schafroth et al. (2021) investigated theory of mind capacities of rhesus monkeys (*Macaca mulatta*, also belonging to the Old World group) in a free-viewing paradigm. This experiment used the same classic Heider-Simmel animations as in human studies, which allows for better interspecies comparisons even though such movies might not be ethologically relevant to monkeys. Whereas humans have longer fixation durations (an indicator of deeper processing) when viewing sequences of interactions best understood using a mentalistic stance, the monkeys fixated longest on animations that could be interpreted from a teleological (goal-directed) stance. Importantly, however, this effect vanished when perceptual variables, including peak motion and motion variability were included as covariates. The authors therefore concluded that there is no evidence that rhesus monkeys have an understanding of goals from simple shapes. They also noted that the monkeys were largely disinterested in the more complex theory of mind animations and glanced around the testing room during those events.

Monkeys' disengagement from Heider-Simmel stimuli might be due to the abstract symbolic nature of these animations. Indeed, in an experiment with rhesus monkeys who observed a human actor reaching for food hidden in one of two containers, the monkeys looked preferentially at the actor's target, evidence that they can make inferences about goals (Wood et al., 2007). Interestingly, this preference was evident only when the action was performed rationally. The monkeys gazed preferentially at the target when actor's hands were occupied holding another object and he reached for the container with his elbow, but not when the actor had empty hands and still (inefficiently) used his elbow. This indicates that the rhesus monkeys were sensitive to the rational nature of the action, consistent with adopting a teleological stance as taken by bonobos (Krupenye and Hare, 2018). Other primates including chimpanzees and tamarins (New World monkeys) were also tested in this study and showed similar rational teleological stances.

New World monkeys diverged from Old World monkeys and Apes about 30–40 million years ago (Wood et al., 2007; Hayashi et al., 2020). The goal-understanding of New World monkeys has been tested with simple animations as well, but with less promising results. Atsumi and Nagasaka (2015), for instance, found squirrel monkeys to be capable of perceiving the chasing of discs, using a similar design as in Atsumi et al. (2017). The same issue of over-training and reward associations, however, also applies here. Another study with marmoset monkeys (*Callithrix jacchus*; Burkart et al., 2012) has also shown that New World monkeys are incapable of attributing goals to moving objects such as a box, but can do so when observing a conspecific, similar to Wood et al. (2007)'s findings with monkeys observing human actors.

There is reason to believe that both Great Apes and Old World monkeys are capable of more sophisticated reasoning about actions than allowed by teleological representations. In a study conducted by Kano et al. (2019) Great Apes (including bonobos, chimpanzees, and orangutans) watched an ape-like actor who hid an object under one of two boxes in the presence of a human-like actor, and then moved the object to the other box when the second actor was away. After the return of the human actor, Great Apes preferentially fixated at the first box, indicating an expectation based on the actor's knowledge and guided by his false belief. A similar paradigm was used in an experiment with Japanese macaques (Hayashi et al., 2020), who also looked longer at the box where they expected the actor to falsely believe to be the location of a hidden object. Furthermore, by disrupting the medial prefrontal cortex (mPFC) of the macaques (by injections of an inhibitory drug) and consequently eliminating the animals' anticipatory looking toward the false-belief location, the authors suggested a causal role for this brain region in mentalizing. This disruption left macaques' other abilities including movement tracking intact.

In sum, a teleological understanding of the world might date back to New World monkeys, about 40 million years ago. More cognitive mental state attributions are however more evolutionary recent, as discovered with false-belief tasks in Old World monkeys who share 30 million-year-old ancestors with Apes. Findings specifically with Heider-Simmel animations were less reliable. It is

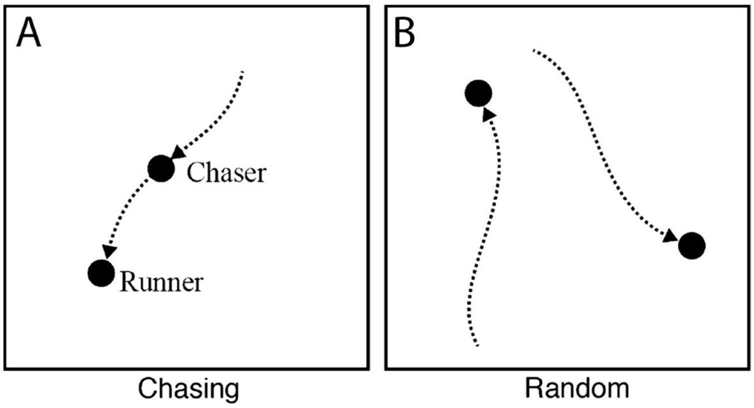


FIGURE 3
Chasing **(A)** and random **(B)** animations showing goal-directed path following and undirected behavior, as shown to macaques and squirrel monkeys in the studies of Atsumi et al. (2017) and Atsumi and Nagasaka (2015). Both groups of monkeys exhibited an understanding of directedness by attributing goals to the chaser. Adapted with permission.

possible that only Great Apes are capable of engaging with symbolic representations of social behavior, and that is also restricted to attributions of goals.

## 6. The neural underpinnings of social cognition

Numerous studies have linked brain areas involved in social cognition to perceived animacy from animations of simple shapes. Using fMRI, Gobbini et al. (2007) investigated neural responses of human adults to animations involving rigid social interactions that conveyed goal-directed action, and to false belief stories. Interestingly, and consistent with previously reviewed behavioral reports, two distinct systems were evoked by goal-directed animations and mentalistic stories. These systems were widely distributed, but notably involved the posterior superior temporal sulcus (pSTS) for representations of goals, and the temporo-parietal junction (TPJ) for mental state attributions, areas known as part of the neural system for theory of mind. Both the pSTS and the TPJ were also found together in the PET scans of individuals in another study who watched mentalistic Heider-Simmel like animations vs. simple action animations that conveyed no social meaning (Castelli et al., 2000). The authors also reported the involvement of the medial prefrontal cortex (mPFC), a midline structure associated with introspective thought, when viewing animations depicting mentalistic attributions. Martin and Weisberg (2003) found further evidence that social interactions between shapes engage the social cognitive brain network. Using long narrative vignettes (21 s) of simple geometric shapes that depicted either social interactions or mechanical relations, the researchers identified distributed patterns of neural activity bilaterally on the STS and within ventral parts of the mPFC (vmPFC), the latter finding proposed to be the results of the narrative eliciting emotional attributions.

These neural findings in adults have identified a set of brain systems that are widely accepted as the so-called "social brain". How does this social brain system develop in infants when the cognitive processes that support perceived animacy and social cognition are coming online? It is very challenging to engage young children in task-related experiments, and maybe near impossible for infants and neonates, especially when conducting neurophysiological measures. While engaging children with specific tasks remains unlikely, resting-state task-free paradigms are starting to elucidate the development of brain networks. The most common of these paradigms include imaging during natural sleep which has been widely utilized for younger children, i.e., newborns (Fransson et al., 2009), children under 3 years (Howell et al., 2019), and 2–4-year-olds (Redcay et al., 2007), and during wake with passive watching of movies of their choice for older children, as conducted by Howell et al. (2019) on children above 3 years, and by Emerson et al. (2015) on 6-year-olds. Although the latter approach can still shed light on brain development, it comes with the issue of engaging children with visual and auditory stimuli, which has potential to shift cortical networks into task-driven states rather than being structured by intrinsic connectivity (Biswal et al., 1995).

Because of our particular interest in the first two years of life in this review, our focus will be on naturally sleeping children, in which spontaneous, low-frequency neural activity results in the emergence of intrinsic functional networks of the brain, known as resting-state networks (RSNs). Of the different modalities used to investigate the correlation between brain regions, resting-state functional MRI (rs-fMRI) has shown more promise as it comes with higher spatial resolution, although the importance of integration with high temporal resolution techniques, including EEG and MEG, is worth noting (Grayson and Fair, 2017). We review resting-state networks—specifically the default mode network—in adults next, as a framework for the section that follows on the development of resting state.

### 6.1. Resting-state and social cognitive networks in adults

In the resting state, adult brain networks organize into a relatively small number of consistent states, which include the default-mode network (DMN), a vastly distributed network consisting of regions including but not limited to the inferior

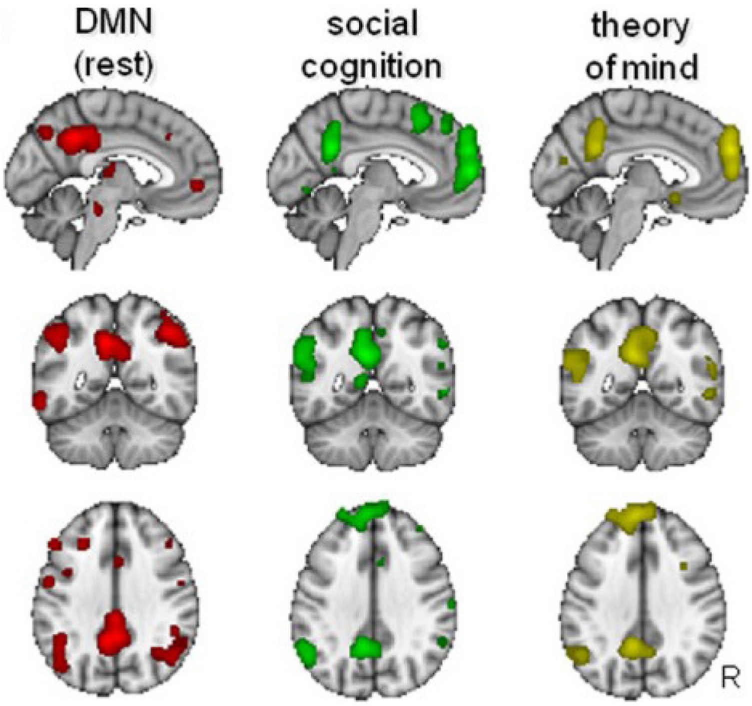


FIGURE 4
Overlap between the default-mode network (DMN), social cognition, and theory of mind. Similarities exist at the posterior cingulate cortex (PCC), the medial prefrontal cortex (mPFC), and the temporo-parietal junction (TPJ). Adapted from Mars et al. (2012) (CC-BY-NC 3.0).

parietal cortex (the inferior parietal lobule, IPL; and the temporo-parietal junction, TPJ), the posterior cingulate cortex (PCC), and the mPFC. This network is the one most commonly identified in the absence of external stimuli (Buckner et al., 2008) during which individuals engage in stimulus-independent or spontaneous thought that may consist of dreaming, mind-wandering or creative thinking (Christoff et al., 2016). The cognitive processes associated with the DMN are commonly linked to internally directed thought, which has been shown to include memory retrieval, planning for the future, and reasoning about others (Harrison et al., 2008), all key cognitive functions for developing mental models of situational context that facilitates navigating social interactions (Yeshurun et al., 2021). Evidence shows that at the core of such functions lies an understanding of the self, which can send projections to these processes or act as a reference (Buckner and Carroll, 2007; Buckner et al., 2008).

Figure 4 illustrates the similarities between the DMN, the social cognitive network, and the system involved in theory of mind, with clear overlap in the parietal, posterior medial and medial frontal regions. The IPL is involved both in the DMN and when humans think about themselves vs. others (Vogeley and Fink, 2003; Schilbach et al., 2008; Mars et al., 2012). Posterior medial parts of the DMN which include the PCC and precuneus are also involved in social cognitive processes such as mentalizing (Saxe and Powell, 2006) and social interactions (Schilbach et al., 2006). These complex systems also have subdivisions within. For example, the TPJ can be split into a posterior and anterior region, which are known to play key roles, respectively, in mentalizing and orientation of attention (Patel et al., 2019). The DMN most strongly overlaps with the posterior TPJ (Mars et al., 2012), which has been associated specifically with the attribution of intentions (Atique et al., 2011).

The mentalizing role of the TPJ engages a widely distributed network that includes the mPFC (Mason et al., 2008; Burnett and Blakemore, 2009; Atique et al., 2011; Baumgartner et al., 2012; Hervé et al., 2012). This network is implicated in mentalizing particularly when reasoning about hidden beliefs that are internal to the mind (Lieberman, 2007). The mPFC has been divided into three subdivisions at its ventral (vmPFC), anterior (amPFC), and dorsal (dmPFC) sides, each with distinct functional specialization and associated network (see Figure 5). Whereas the dmPFC is most activated when selectively reasoning about others (Dd et al., 2012; Denny et al., 2012), the vmPFC involves self-relevant representations and the amPFC is engaged during tasks that require drawing distinctions between self and others (D'Argembeau et al., 2005; Heatherton et al., 2006; Andrews-Hanna et al., 2010). The transition from understanding the self in ventral parts of the mPFC, to representation of others in the dmPFC is demonstrated by Li et al. (2014) as well. Together, the TPJ and mPFC have been linked to the consolidation of recently learned social information, as demonstrated by increased connectivity when measured after exposure and during rest (Meyer et al., 2019).

It is important to note that assignment of reward to social signals has also been found within the social network (Frith, 2007). The TPJ, for instance, shows increased functional connectivity with reward processing regions during empathy (Janowski et al., 2013) and charitable donations (Hare et al., 2010), which can be described as social motivation being associated with social reward (Grimm et al., 2021); more generally, the role of the TPJ might be to

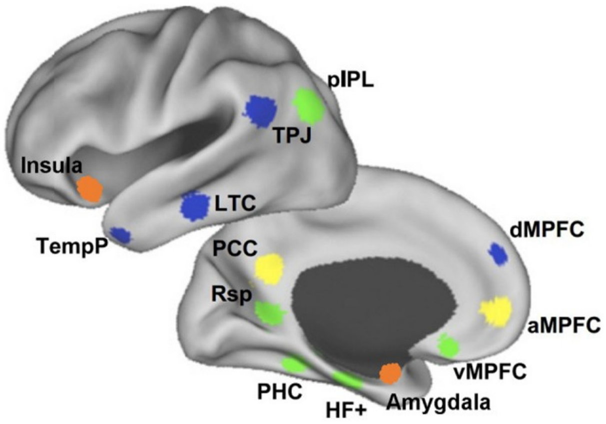


FIGURE 5
Subdivisions of the DMN. The vmPFC (green), through connections with the MTL and the IPL, represents self-relevant thought. The amPFC (yellow) connects to the PCC to draw distinctions between the self and others. The dmPFC (blue) is involved in reasoning about others together with the TPJ. Adapted from Li et al. (2014) (CC-BY 3.0).

compare predictions about the environment with actual outcomes (Abrahamse and Silvetti, 2016). Social reward is mainly processed at the mesolimbic system, which includes the ventral tegmental area (VTA), the PFC, and striatum (Meshi et al., 2013; Serafini et al., 2020), although these regions are involved in various non-social processes as well. The vmPFC is also associated with reward and punishment, and through connections with the amygdala and insula (Carmichael and Price, 1995; Akitsuki and Decety, 2009; Otti et al., 2010), involves in perceiving emotions in social contexts. Social reward, punishment, and motivation have been found in the dmPFC as well (Fehr and Camerer, 2007; Kohls et al., 2013). It is proposed that the vmPFC is involved in self-referenced reward (Dang et al., 2019), while the dmPFC that is linked to cognitive tasks, processes reward information related to others (Apps et al., 2012; Lockwood et al., 2015).

It is worth mentioning that the activation of the social cognitive network is also influenced by cueing and attention, specifically with Heider-Simmel style animations. In an fMRI study, Tavares et al. (2008) showed significant boosts in the social brain network when selective attention was paid to social meaning vs. to spatial properties of the movies. Participants were cued either by the word "behavioral" or by "spatial" before observing animations that showed two circles (i.e., agents) moving through constraints. In the spatial condition, participants were asked to attend to motion features such as speed or trajectory patterns. When cued with "behavioral", however, they were instructed to identify the type of interaction between the circles. Cueing can therefore enhance attributions of mental states toward movies with simple shapes. The pSTS, in particular, has however been shown to respond to interactivity cues irrespective of task, suggesting its automatic involvement in detecting animacy (Schultz et al., 2005).

We will review resting-state development in the coming section, and will also further discuss the roles of the TPJ and mPFC in theory of mind. First, though, it is worth noting that within the DMN various hubs have been identified, which form the basis for the developmental trajectory of the system as a whole. Both Mars et al. (2012) and Buckner et al. (2008) have identified the PCC-Rsp and the mPFC as DMN hubs when evaluated in the resting-state, with the potential for hub properties in the parietal regions of the DMN based on task-based studies of social cognition (Yang et al., 2015; Patel et al., 2019). The PCC, in particular, plays an important role in DMN development (Gao et al., 2009), as will be discussed below. This region not only acts as a key hub within the DMN, it is also involved in attributions of mental states. As Lombardo et al. (2010) have shown, the PCC functionally connects to the TPJ and the mPFC when mentalizing about the self and others, and also responds to self-relevant emotional events (Vogt et al., 2006).

## 6.2. Resting-state in children

As discussed earlier, resting-state fMRI has been a common approach for studying the functional networks of the young brain. In a study of cortical network activity in the first two years of life, Gao et al. (2015b) identified nine functional networks, which divide into topologically adult-like primary networks and widely distributed higher-order networks that are incomplete in younger children. The former includes an early visual and a sensorimotor network, while the latter consists of multiple networks, including a DMN, that become more consolidated through childhood and adolescence (Fair et al., 2008; Mak et al., 2017). All higher-order networks appear in forms that are rudimentary as compared to their adult version.

Investigations on preterm infants at term-equivalent age have also shown the existence of five resting-state networks, as illustrated in Figure 6. The primary visual, auditory, and somato-motor networks resemble the adult counterparts and will only undergo fine developments later in the first 2 years of life (Lin et al., 2008; Gao et al., 2015b). The other networks take more time to develop into mature forms, although not directly into an adult equivalent. In particular, the medial and lateral parietal networks (shown as network D in Figure 6) are regarded by Fransson et al. (2007) as a proto-DMN, which includes the posterior parts of a well-formed DMN. More higher-order early networks have also been detected in preterm infants, including an executive control network (Doria et al., 2010).

Regions that form the proto-DMN, which are mainly bilateral, later develop into brain regions linked to mentalizing, consistent with behavioral findings of goal attribution in infants as young as 3 months (Luo, 2011). Connections within this network, as well as other infant networks, will drastically grow during the first compared to the second year, while between-network segregation also occurs (Gao et al., 2015b). Pruett et al. (2015) specifically demonstrated significant DMN development in the second half of the first year, by showing the network's involvement (together with the dorsal attention network) in identifying 6 vs. 12 month-olds based on their patterns of functional connectivity in the resting-state.

It takes over 10 years for the DMN to find its complete mature form (Hoff et al., 2013; Fan et al., 2021), although this time course will be delayed in autistic children who show

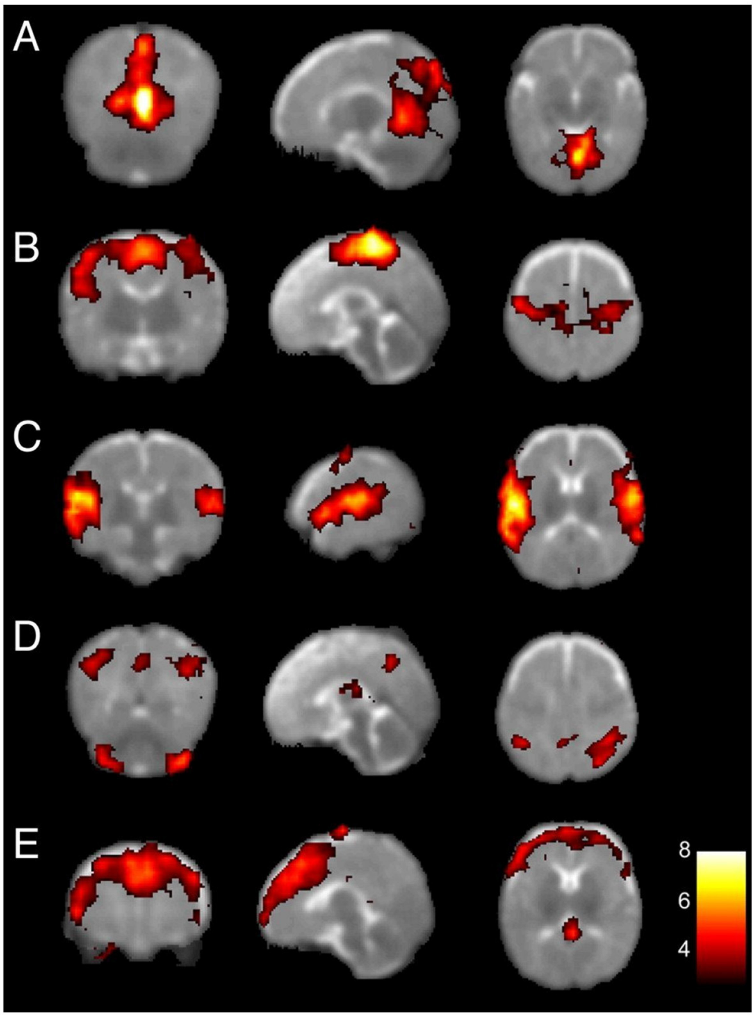


FIGURE 6
Resting-state networks at birth. Primary visual **(A)**, bilateral sensorimotor **(B)**, bilateral auditory **(C)**, proto-DMN, consisting of the lateral parts of the cerebellum, the posterior mid-parietal areas including the precuneus, and the posterior lateral parietal cortex **(D)**, and prefrontal **(E)** networks. Adapted from Fransson et al. (2007). Copyright (2007) National Academy of Sciences.

greater modularization driven by reduced between-subnetwork connectivity (Bathelt and Geurts, 2021). As a milestone in DMN development, connections within the PCC strengthen around the age of 2 to turn it into a hub that functions to link the posterior and anterior regions of the DMN (Gao et al., 2009). This posterior-anterior growth is of chief importance here because of its possible relevance to teleological-mentalistic representations. Other developmental patterns, however, have been identified, in inferior to superior and medial to lateral directions (Gao et al., 2015a).

## 6.3. Social predisposition or learned competence?

If a social tendency, such as the sensitivity to self-propelled motion (Di Giorgio et al., 2017), is evident in newborns even only after a couple of days, it is important to ask whether such tendency is indeed innate or has been influenced by learned mechanisms. As highly altricial species, human babies cannot be deprived of early learning and might never be suitable for investigations of inborn biases. Precocial animals such as domestic chicks, however, can be kept in complete darkness after hatching until tested for predispositions, making them feasible models for nature vs. nurture research.

Similar to human newborns, newly hatched chicks with no prior visual experience are sensitive to face-like configurations (Rosa-Salva et al., 2010; Rosa Salva et al., 2011), biological motion (Rugani et al., 2015; Miura et al., 2020), and to simple shapes that show animacy through self-propelled motion (Mascalzoni et al., 2010), speed change (Rosa-Salva et al., 2016; Versace et al., 2019; Lorenzi et al., 2021), orientation to motion direction (Clara et al., 2009; Rosa-Salva et al., 2018), or gradual trajectory changes (Rosa-Salva et al., 2016). Rosa Salva et al. (2011) tested the preference toward faces in visually naïve 2-day-old chicks (*Gallus gallus*) who never saw the experimenter's face or the face of another chick. They found that chicks preferred human faces to frequency- and color-matched control scrambled images, just as newborn human infants do. The use of human faces for chicks is proposed to demonstrate an innate non-species specific face preference (Morton and Johnson, 1991) that gives way to species, breed and identity selective preferences through environmental exposure during sensitive periods of development (Rosa-Salva et al., 2021).

Newborn chicks also have a spontaneous preference for cues that signal animacy Rosa-Salva et al. (2016). Naïve domestic chicks were placed in a runway apparatus that displayed one shape moving at constant speed at one end, and a speed-changing shape at the other end. The latter accelerated at one third of its path and decelerated to its initial speed at two thirds of the trajectory. Chicks preferred and approached the second shape, showing a predisposition for the animacy cue of speed change. Interestingly, similar to the early transient non-species face preference in chicks, this predisposition exists for only 24 h after hatching, and fades two days later (Rosa-Salva et al., 2021), although it can be restored by administering a hormone associated with the opening of critical windows in imprinting, at least in female chicks (Lorenzi et al., 2021). It is also worth noting the importance of exposure to environmental cues for animacy, as occluding the speed changes will suppress the preference (Rosa-Salva et al., 2016). Predispositions for animacy can be diminished with embryonic injections of Valproic Acid (VPA) as well, which models the behavioral deficits observed in autism spectrum disorder (ASD). Sgadò et al. (2018) showed that VPA exposure impairs newly hatched chicks' predisposition for hen-like objects, while it leaves their subsequent learning intact, evident from their normal imprinting behavior toward a familiar simple shape.

Subcortical—specifically limbic—structures may play an important role in the early detection of animacy and in the imprinting behavior of vertebrates, which function prior to postnatal learning. The nucleus taeniae and arcopallium (amygdala homologues) together with the septal nuclei and the preoptic area (POA) of the hypothalamus of visually naïve chicks are linked to viewing live conspecifics (Mayer et al., 2017a,b) and hen-like objects (Mayer et al., 2019). The POA along with the septum have also been linked to viewing speed changes associated with animacy Lorenzi et al. (2017). The function of limbic structures in promoting perceived animacy may serve to support the imprinting process between newborn chicks and their early social partners given the role of these structures in emotional valence (O'Connell and Hofmann, 2011).

## 7. Conclusions

Like adults, human newborns show preferences toward animacy cues, and are able to connect animate objects to their goals as early as the age of 3 months. They then are capable of attributing beliefs to others toward the end of the first year and understand subjective minds. Here we suggest that the development from a proto-DMN into a maturing DMN between birth and the age of 2 supports the transition from perceiving goals to making attributions about mental states (Pruett et al., 2015). This is consistent with the dramatic development of social cognitive functions that emerge around 9 months of age, specifically with the behavioral findings of false beliefs before the age of two (Onishi and Baillargeon, 2005; Surian et al., 2007; Buttelmann et al., 2009) and even before the first birthday (Rochat et al., 2004; Hamlin et al., 2013). It is important to note that a fully mature DMN is not yet emergent when early theory of mind has appeared. Mind-understanding begins functioning before the first birthday alongside a primitive DMN that includes the TPJ, which represents externally-focused processes. Over time the maturation of the DMN includes the formation of a PCC hub and connectivity to the mPFC, consistent with the emergence of reasoning about internal states of others (Lieberman, 2007). Further investigation is however needed to find direct connections between the early DMN and social cognition.

Here we also reviewed findings of goal and belief attribution in monkeys and apes to discuss whether the teleological-mentalistic sequence found in humans has evolutionary origins. A teleological understanding of the world indeed emerges before mentalizing in evolutionary terms, as evident from findings of goal attribution in New World monkeys, and from findings of mentalizing in Old World monkeys who emerge from a more recent branch in evolution. With Heider-Simmel style animations, which represent social interactions through symbolic abstractions, only Great Apes show an engagement which is limited under a teleological stance. Recognition of animacy cues (e.g., speed change), more generally, has been found in visually naïve chicks as well, suggesting that a wide range of vertebrates are predisposed to animacy.

## 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 supported by the National Science Foundation grant BCS-1658560 to EG.

## 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.

## Publisher's note

# References

Abrahamse, E. L., and Silvetti, M. (2016). Commentary: the role of the parietal cortex in the representation of task–reward associations. *Front. Hum. Neurosci.* 10, 192. doi: 10.3389/fnhum.2016.00192

Akitsuki, Y., and Decety, J. (2009). Social context and perceived agency affects empathy for pain: an event-related fMRI investigation. *Neuroimage* 47, 722–734. doi: 10.1016/j.neuroimage.2009.04.091

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron* 65, 550–562. doi: 10.1016/j.neuron.2010.02.005

Apps, M. A. J., Balsters, J. H., and Ramnani, N. (2012). The anterior cingulate cortex: monitoring the outcomes of others' decisions. *Soc. Neurosci.* 7, 424–435. doi: 10.1080/17470919.2011.638799

Atique, B., Erb, M., Gharabaghi, A., Grodd, W., and Anders, S. (2011). Task-specific activity and connectivity within the mentalizing network during emotion and intention mentalizing. *Neuroimage* 55, 1899–1911. doi: 10.1016/j.neuroimage.2010.12.036

Atsumi, T., Koda, H., and Masataka, N. (2017). Goal attribution to inanimate moving objects by Japanese macaques (*Macaca fuscata*). *Sci. Rep.* 7, 40033. doi: 10.1038/srep40033

Atsumi, T., and Nagasaka, Y. (2015). Perception of chasing in squirrel monkeys (*Saimiri sciureus*). *Anim. Cogn.* 18, 1243–1253. doi: 10.1007/s10071-015-0893-x

Baillargeon, R. (2008). Innate ideas revisited for a principle of persistence in infants' physical reasoning. *Perspect. Psychol. Sci.* 3, 2–13. doi: 10.1111/j.1745-6916.2008.00056.x

Baker, C. L., Jara-Ettinger, J., Saxe, R., and Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.* 1, 0064. doi: 10.1038/s41562-017-0064

Bathelt, J., and Geurts, H. M. (2021). Difference in default mode network subsystems in autism across childhood and adolescence. *Autism* 25, 556–565. doi: 10.1177/1362361320969258

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., and Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behav. Dev.* 23, 223–229. doi: 10.1016/S0163-6383(01)00037-6

Baumgartner, T., Götte, L., Gügler, R., and Fehr, E. (2012). The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum. Brain Mapp.* 33, 1452–1469. doi: 10.1002/hbm.21298

Berry, D. S., and Springer, K. (1993). Structure, motion, and preschoolers' perceptions of social causality. *Ecol. Psychol.* 5, 273–283. doi: 10.1207/s15326969eco0504_1

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409

Brown, M. I., Speer, A. B., Tenbrink, A. P., and Chabris, C. F. (2022). Using game-like animations of geometric shapes to simulate social interactions: An evaluation of group score differences. *Int. J. Select. Assess.* 30, 167–181. doi: 10.1111/ijsa.12375

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004

Buiatti, M., Di Giorgio, E., Piazza, M., Polloni, C., Menna, G., Taddei, F., et al. (2019). Cortical route for facelike pattern processing in human newborns. *Proc. Nat. Acad. Sci. U. S. A.* 116, 4625–4630. doi: 10.1073/pnas.1812419116

Burkart, J., Kupferberg, A., Glasauer, S., and van Schaik, C. (2012). Even simple forms of social learning rely on intention attribution in marmoset monkeys (*Callithrix jacchus*). *J. Comp. Psychol.* 126, 129–138. doi: 10.1037/a0026025

Burnett, S., and Blakemore, S.-J. (2009). Functional connectivity during a social emotion task in adolescents and in adults. *Eur. J. Neurosci.* 29, 1294–1301. doi: 10.1111/j.1460-9568.2009.06674.x

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342. doi: 10.1016/j.cognition.2009.05.006

Carey, S. (1985). *Conceptual Change in Childhood*. The MIT Press. Available online at: https://mitpress.mit.edu/9780262530736/conceptual-change-in-childhood/

Carmichael, S. T., and Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* 363, 615–641. doi: 10.1002/cne.903630408

Castelli, F., Happé, F., Frith, U., and Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12, 314–325. doi: 10.1006/nimg.2000.0612

Chang, D. H. F., and Troje, N. F. (2008). Perception of animacy and direction from local biological motion signals. *J. Vis.* 8, 3. doi: 10.1167/8.5.3

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., and Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: a dynamic framework. *Nat. Rev. Neurosci.* 17, 718–731. doi: 10.1038/nrn.2016.113

Clara, E., Regolin, L., Vallortigara, G., and Rogers, L. J. (2009). Chicks prefer to peck at insect-like elongated stimuli moving in a direction orthogonal to their longer axis. *Anim. Cogn.* 12, 755–765. doi: 10.1007/s10071-009-0235-y

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107, 705–717. doi: 10.1016/j.cognition.2007.08.001

Csibra, G., Gergely, G., Bıró, S., Koós, O., and Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition* 72, 237–267. doi: 10.1016/S0010-0277(99)00039-6

Dang, T. P., Mattan, B. D., Kubota, J. T., and Cloutier, J. (2019). The ventromedial prefrontal cortex is particularly responsive to social evaluations requiring the use of person-knowledge. *Sci. Rep.* 9, 5054. doi: 10.1038/s41598-019-41544-z

D'Argembeau, A., Collette, F., Van der Linden, M., Laureys, S., Del Fiore, G., Degueldre, C., et al. (2005). Self-referential reflective activity and its relationship with rest: a PET study. *Neuroimage* 25, 616–624. doi: 10.1016/j.neuroimage.2004.11.048

Dd, W., Jv, H., and Tf, H. (2012). The representation of self and person knowledge in the medial prefrontal cortex. *Cogn. Sci.* 3. doi: 10.1002/wcs.1183

Denny, B. T., Kober, H., Wager, T. D., and Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* 24, 1742–1752. doi: 10.1162/jocn_a_00233

Di Giorgio, E., Frasnelli, E., Rosa Salva, O., Luisa Scattoni, M., Puopolo, M., Tosoni, D., et al. (2016). Difference in visual social predispositions between newborns at low- and high-risk for autism. *Sci. Rep.* 6, 26395. doi: 10.1038/srep26395

Di Giorgio, E., Lunghi, M., Simion, F., and Vallortigara, G. (2017). Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Dev. Sci.* 20, e12394. doi: 10.1111/desc.12394

Di Giorgio, E., Lunghi, M., Vallortigara, G., and Simion, F. (2021). Newborns' sensitivity to speed changes as a building block for animacy perception. *Sci. Rep.* 11, 542. doi: 10.1038/s41598-020-79451-3

Doria, V., Beckmann, C. F., Arichi, T., Merchant, N., Groppo, M., Turkheimer, F. E., et al. (2010). Emergence of resting state networks in the preterm human brain. *Proc. Nat. Acad. Sci. U. S. A.* 107, 20015–20020. doi: 10.1073/pnas.1007921107

Emerson, R. W., Short, S. J., Lin, W., Gilmore, J. H., and Gao, W. (2015). Network-level connectivity dynamics of movie watching in 6-year-old children. *Front. Hum. Neurosci.* 9, 631. doi: 10.3389/fnhum.2015.00631

Fair, D. A., Cohen, A. L., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Barch, D. M., et al. (2008). The maturing architecture of the brain's default network. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4028–4032. doi: 10.1073/pnas.0800376105

Fan, F., Liao, X., Lei, T., Zhao, T., Xia, M., Men, W., et al. (2021). Development of the default-mode network during childhood and adolescence: a longitudinal resting-state fMRI study. *Neuroimage* 226, 117581. doi: 10.1016/j.neuroimage.2020.117581

Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U. S. A.* 99, 9602–9605. doi: 10.1073/pnas.152159999

Fehr, E., and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427. doi: 10.1016/j.tics.2007.09.002

Flavell, J. H. (1988). *The Development of Children's Knowledge About the Mind: From Cognitive Connections to Mental Representations. Developing Theories of Mind.* New York, NY: Cambridge University Press), 244–267.

Fransson, P., Skiöld, B., Engström, M., Hallberg, B., Mosskin, M., Åden, U., et al. (2009). Spontaneous brain activity in the newborn brain during natural sleep—An fMRI study in infants born at full term. *Pediatr. Res.* 66, 301–305. doi: 10.1203/PDR.0b013e3181b1bd84

Fransson, P., Skiöld, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., et al. (2007). Resting-state networks in the infant brain. *Proc. Natl. Acad. Sci. U. S. A.* 104, 15531–15536. doi: 10.1073/pnas.0704380104

Frith, C. D. (2007). The social brain? *Philos. Trans. R. Soc. Lond. B,. Biol. Sci.* 362, 671–678. doi: 10.1098/rstb.2006.2003

Gao, T., Baker, C. L., Tang, N., Xu, H., and Tenenbaum, J. B. (2019). The cognitive architecture of perceived animacy: intention, attention, and memory. *Cogn. Sci.* 43, e12775. doi: 10.1111/cogs.12775

Gao, T., McCarthy, G., and Scholl, B. J. (2010). The wolfpack effect: perception of animacy irresistibly influences interactive behavior. *Psychol. Sci.* 21, 1845–1853. doi: 10.1177/0956797610388814

Gao, W., Alcauter, S., Elton, A., Hernandez-Castillo, C. R., Smith, J. K., Ramirez, J., et al. (2015a). Functional network development during the first year:

relative sequence and socioeconomic correlations. *Cereb. Cortex* 25, 2919–2928. doi: 10.1093/cercor/bhu088

Gao, W., Alcauter, S., Smith, J. K., Gilmore, J., and Lin, W. (2015b). Development of human brain cortical network architecture during infancy. *Brain Struct. Funct.* 220, 1173–1186. doi: 10.1007/s00429-014-0710-3

Gao, W., Zhu, H., Giovanello, K. S., Smith, J. K., Shen, D., Gilmore, J. H., et al. (2009). Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc. Nat. Acad. Sci. U. S. A.* 106, 6790–6795. doi: 10.1073/pnas.0811221106

Gergely, G., and Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends Cogn. Sci.* 7, 287–292. doi: 10.1016/S1364-6613(03)00128-1

Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193. doi: 10.1016/0010-0277(95)00661-H

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., and Haxby, J. V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814. doi: 10.1162/jocn.2007.19.11.1803

Grayson, D. S., and Fair, D. A. (2017). Development of large-scale functional networks from birth to adulthood: a guide to neuroimaging literature. *Neuroimage* 160, 15–31. doi: 10.1016/j.neuroimage.2017.01.079

Grimm, C., Balsters, J. H., and Zerbi, V. (2021). Shedding light on social reward circuitry: (un)common blueprints in humans and rodents. *Neuroscientist* 27, 159–183. doi: 10.1177/1073858420923552

Hamlin, J. K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559. doi: 10.1038/nature06288

Hamlin, J. K., Wynn, K., and Bloom, P. (2010). 3-month-olds show a negativity bias in their social evaluations. *Dev. Sci.* 13, 923–929. doi: 10.1111/j.1467-7687.2010.00951.x

Hamlin, K. J., Ullman, T., Tenenbaum, J., Goodman, N., and Baker, C. (2013). The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Dev. Sci.* 16, 209–226. doi: 10.1111/desc.12017

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., and Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30, 583–590. doi: 10.1523/JNEUROSCI.4089-09.2010

Harrison, B. J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., et al. (2008). Consistency and functional specialization in the default mode brain network. *Proc. Nat. Acad. Sci. U. SS. A.* 105, 9781–9786. doi: 10.1073/pnas.0711791105

Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., et al. (2020). Macaques exhibit implicit gaze bias anticipating others' false-belief-driven actions via medial prefrontal cortex. *Cell Rep.* 30, 4433–4444.e5. doi: 10.1016/j.celrep.2020.03.013

Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., and Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Soc. Cogn. Affect. Neurosci.* 1, 18–25. doi: 10.1093/scan/nsl001

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259. doi: 10.2307/1416950

Hervé, P.-Y., Razafimandimby, A., Vigneau, M., Mazoyer, B., and Tzourio-Mazoyer, N. (2012). Disentangling the brain networks supporting affective speech comprehension. *Neuroimage* 61, 1255–1267. doi: 10.1016/j.neuroimage.2012.03.073

Hespos, S. J., and vanMarle, K. (2012). Physics for infants: characterizing the origins of knowledge about objects, substances, and number. *WIREs Cogn.* Sci3, 19–27. doi: 10.1002/wcs.157

Hoff, G. E. A.-J., Van den Heuvel, M. P., Benders, M. J. N. L., Kersbergen, K. J., and De Vries, L. S. (2013). On development of functional brain connectivity in the young brain. *Front. Hum. Neurosci.* 7, 650. doi: 10.3389/fnhum.2013.00650

Howell, B. R., Styner, M. A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., et al. (2019). The UNC/UMN Baby Connectome Project (BCP): an overview of the study design and protocol development. *Neuroimage* 185, 891–905. doi: 10.1016/j.neuroimage.2018.03.049

Janowski, V., Camerer, C., and Rangel, A. (2013). Empathic choice involves vmPFC value signals that are modulated by social processing implemented in IPL. *Soc. Cogn. Affect. Neurosci.* 8, 201–208. doi: 10.1093/scan/nsr086

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. 47. doi: 10.2307/1914185

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. Nat. Acad. Sci. U. S. A.* 116, 20904–20909. doi: 10.1073/pnas.1910095116

Kim, J., Allison, S. T., Eylon, D., Goethals, G. R., Markus, M. J., Hindle, S. M., et al. (2008). Rooting for (and then abandoning) the underdog. *J. Appl. Soc. Psychol.* 38, 2550–2573. doi: 10.1111/j.1559-1816.2008.00403.x

Kohls, G., Schulte-Rüther, M., Nehrkorn, B., Müller, K., Fink, G. R., Kamp-Becker, I., et al. (2013). Reward system dysfunction in autism spectrum disorders. *Soc. Cogn. Affect. Neurosci.* 8, 565–572. doi: 10.1093/scan/nss033

Krupenye, C., and Hare, B. (2018). Bonobos prefer individuals that hinder others over those that help. *Curr. Biol.* 28, 280–286.e5. doi: 10.1016/j.cub.2017.11.061

Legerstee, M., Pomerleau, A., Malcuit, G., and Feider, H. (1987). The development of infants' responses to people and a doll: implications for research in communication. *Infant Behav. Dev.* 10, 81–95. doi: 10.1016/0163-6383(87)90008-7

Li, W., Mai, X., and Liu, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Front. Hum. Neurosci.* 8, 74. doi: 10.3389/fnhum.2014.00074

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289. doi: 10.1146/annurev.psych.58.110405.085654

Lin, W., Zhu, Q., Gao, W., Chen, Y., Toh, C.-H., Styner, M., et al. (2008). Functional connectivity MR imaging reveals cortical functional connectivity in the developing brain. *Am. J. Neuroradiol.* 29, 1883–1889. doi: 10.3174/ajnr.A1256

Lockwood, P. L., Apps, M. A. J., Roiser, J. P., and Viding, E. (2015). Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *J. Neurosci.* 35, 13720–13727. doi: 10.1523/JNEUROSCI.1703-15.2015

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Wheelwright, S. J., Sadek, S. A., Suckling, J., et al. (2010). Shared neural circuits for mentalizing about the self and others. *J. Cogn. Neurosci.* 22, 1623–1635. doi: 10.1162/jocn.2009.21287

Lorenzi, E., Lemaire, B. S., Versace, E., Matsushima, T., and Vallortigara, G. (2021). Resurgence of an inborn attraction for animate objects via thyroid hormone T3. *Front. Behav. Neurosci.* 15, 675994. doi: 10.3389/fnbeh.2021.675994

Lorenzi, E., Mayer, U., Rosa-Salva, O., and Vallortigara, G. (2017). Dynamic features of animate motion activate septal and preoptic areas in visually naïve chicks (*Gallus gallus*). *Neuroscience* 354, 54–68. doi: 10.1016/j.neuroscience.2017.04.022

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Dev. Sci.* 14, 453–460. doi: 10.1111/j.1467-7687.2010.00995.x

Mak, L. E., Minuzzi, L., MacQueen, G., Hall, G., Kennedy, S. H., and Milev, R. (2017). The default mode network in healthy individuals: a systematic review and meta-analysis. *Brain Connect.* 7, 25–33. doi: 10.1089/brain.2016.0438

Mars, R. B., Neubert, F.-X., Noonan, M. P., Sallet, J., Toni, I., and Rushworth, M. F. S. (2012). On the relationship between the "default mode network" and the "social brain". *Front. Hum. Neurosci.* 6, 189. doi: 10.3389/fnhum.2012.00189

Martin, A., and Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cogn. Neuropsychol.* 20, 575–587. doi: 10.1080/02643290342000005

Mascalzoni, E., Regolin, L., and Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *Proc. Nat. Acad. Sci. U. S. A.* 107, 4483–4485. doi: 10.1073/pnas.0908792107

Mason, R. A., Williams, D. L., Kana, R. K., Minshew, N., and Just, M. A. (2008). Theory of mind disruption and recruitment of the right hemisphere during narrative comprehension in autism. *Neuropsychologia* 46, 269–280. doi: 10.1016/j.neuropsychologia.2007.07.018

Mayer, U., Rosa-Salva, O., Loveland, J. L., and Vallortigara, G. (2019). Selective response of the nucleus taeniae of the amygdala to a naturalistic social stimulus in visually naïve domestic chicks. *Sci. Rep.* 9, 9849. doi: 10.1038/s41598-019-46322-5

Mayer, U., Rosa-Salva, O., Morbioli, F., and Vallortigara, G. (2017a). The motion of a living conspecific activates septal and preoptic areas in naïve domestic chicks (*Gallus gallus*). *Eur. J. Neurosci.* 45, 423–432. doi: 10.1111/ejn.13484

Mayer, U., Rosa-Salva, O., and Vallortigara, G. (2017b). First exposure to an alive conspecific activates septal and amygdaloid nuclei in visually-naïve domestic chicks (*Gallus gallus*). *Behav. Brain Res.* 317, 71–81. doi: 10.1016/j.bbr.2016.09.031

Meshi, D., Morawetz, C., and Heekeren, H. (2013). Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Front. Hum. Neurosci.* 7, 439. doi: 10.3389/fnhum.2013.00439

Meyer, M. L., Davachi, L., Ochsner, K. N., and Lieberman, M. D. (2019). Evidence that default network connectivity during rest consolidates social information. *Cereb. Cortex* 29, 1910–1920. doi: 10.1093/cercor/bhy071

Michotte, A. (1963). *The Perception of Causality*. Basic Books.

Miura, M., Nishi, D., and Matsushima, T. (2020). Combined predisposed preferences for colour and biological motion make robust development of social attachment through imprinting. *Anim. Cogn.* 23, 169–188. doi: 10.1007/s10071-019-01327-5

Morton, J., and Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychol. Rev.* 98, 164–181. doi: 10.1037/0033-295X.98.2.164

O'Connell, L. A., and Hofmann, H. A. (2011). The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. *J. Comp. Neurol.* 519, 3599–3639. doi: 10.1002/cne.22735

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255–258. doi: 10.1126/science.1107621

Otti, A., Guendel, H., Läer, L., Wohlschlaeger, A. M., Lane, R. D., Decety, J., et al. (2010). I know the pain you feel—how the human brain's default mode predicts our resonance to another's suffering. *Neuroscience* 169, 143–148. doi: 10.1016/j.neuroscience.2010.04.072

Patel, G. H., Sestieri, C., and Corbetta, M. (2019). The evolution of the temporoparietal junction and posterior superior temporal sulcus. *Cortex* 118, 38–50. doi: 10.1016/j.cortex.2019.01.026

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

Pruett, J. R., Kandala, S., Hoertel, S., Snyder, A. Z., Elison, J. T., Nishino, T., et al. (2015). Accurate age classification of 6 and 12 month-old infants based on resting-state functional connectivity magnetic resonance imaging data. *Dev. Cogn. Neurosci.* 12, 123–133. doi: 10.1016/j.dcn.2015.01.003

Rasmussen, C. E., and Jiang, Y. V. (2019). Judging social interaction in the Heider and Simmel movie. *Q. J. Exp. Psychol.* 72, 2350–2361. doi: 10.1177/1747021819838764

Ratajska, A., Brown, M. I., and Chabris, C. F. (2020). Attributing social meaning to animated shapes: a new experimental study of apparent behavior. *Am. J. Psychol.* 133, 295–312. doi: 10.5406/amerjpsyc.133.3.0295

Redcay, E., Kennedy, D. P., and Courchesne, E. (2007). fMRI during natural sleep as a method to study brain function during early childhood. *Neuroimage* 38, 696–707. doi: 10.1016/j.neuroimage.2007.08.005

Rochat, P., Striano, T., and Morgan, R. (2004). Who is doing what to whom? Young infants' developing sense of social causality in animated displays. *Perception* 33, 355–369. doi: 10.1068/p3389

Roemmele, M., Morgens, S.-M., Gordon, A. S., and Morency, L.-P. (2016). "Recognizing human actions in the motion trajectories of shapes," in *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, CA: ACM), 271–281.

Rosa Salva, O., Farroni, T., Regolin, L., Vallortigara, G., and Johnson, M. H. (2011). The evolution of social orienting: evidence from chicks (*Gallus gallus*) and human newborns. *PLoS ONE* 6, e18802. doi: 10.1371/journal.pone.0018802

Rosa-Salva, O., Grassi, M., Lorenzi, E., Regolin, L., and Vallortigara, G. (2016). Spontaneous preference for visual cues of animacy in naïve domestic chicks: the case of speed changes. *Cognition* 157, 49–60. doi: 10.1016/j.cognition.2016.08.014

Rosa-Salva, O., Hernik, M., Broseghini, A., and Vallortigara, G. (2018). Visually-naïve chicks prefer agents that move as if constrained by a bilateral body-plan. *Cognition* 173, 106–114. doi: 10.1016/j.cognition.2018.01.004

Rosa-Salva, O., Mayer, U., Versace, E., Hébert, M., Lemaire, B. S., and Vallortigara, G. (2021). Sensitive periods for social development: interactions between predisposed and learned mechanisms. *Cognition* 213, 104552. doi: 10.1016/j.cognition.2020.104552

Rosa-Salva, O., Regolin, L., and Vallortigara, G. (2010). Faces are special for newly hatched chicks: evidence for inborn domain-specific mechanisms underlying spontaneous preferences for face-like stimuli. *Dev. Sci.* 13, 565–577. doi: 10.1111/j.1467-7687.2009.00914.x

Rugani, R., Rosa Salva, O., Regolin, L., and Vallortigara, G. (2015). Brain asymmetry modulates perception of biological motion in newborn chicks (*Gallus gallus*). *Behav. Brain Res.* 290, 1–7. doi: 10.1016/j.bbr.2015.04.032

Saxe, R., Carey, S., and Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55, 87–124. doi: 10.1146/annurev.psych.55.090902.142044

Saxe, R., and Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17, 692–699. doi: 10.1111/j.1467-9280.2006.01768.x

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., and Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends Cogn. Sci.* 19, 65–72. doi: 10.1016/j.tics.2014.11.007

Schafroth, J. L., Basile, B. M., Martin, A., and Murray, E. A. (2021). No evidence that monkeys attribute mental states to animated shapes in the Heider–Simmel videos. *Sci. Rep.* 11, 3050. doi: 10.1038/s41598-021-82702-6

Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R., and Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Conscious. Cogn.* 17, 457–467. doi: 10.1016/j.concog.2008.03.013

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., et al. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia* 44, 718–730. doi: 10.1016/j.neuropsychologia.2005.07.017

Schlottmann, A., and Surian, L. (1999). Do 9-month-olds perceive causation-at-a-distance? *Perception* 28, 1105–1113. doi: 10.1068/p281105

Scholl, B. J., and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309. doi: 10.1016/S1364-6613(00)01506-0

Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M., and Frith, C. D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* 45, 625–635. doi: 10.1016/j.neuron.2004.12.052

Schultz, J., and Frith, C. D. (2022). Animacy and the prediction of behaviour. *Neurosci. Biobehav. Rev.* 140, 104766. doi: 10.1016/j.neubiorev.2022.104766

Serafini, R. A., Pryce, K. D., and Zachariou, V. (2020). The mesolimbic dopamine system in chronic pain and associated affective comorbidities. *Biol. Psychiatry* 87, 64–73. doi: 10.1016/j.biopsych.2019.10.018

Sgadò, P., Rosa-Salva, O., Versace, E., and Vallortigara, G. (2018). Embryonic exposure to valproic acid impairs social predispositions of newly-hatched chicks. *Sci. Rep.* 8, 5919. doi: 10.1038/s41598-018-24202-8

Shu, T., Peng, Y., Fan, L., Lu, H., and Zhu, S.-C. (2018). Perception of human interaction based on motion trajectories: from aerial videos to decontextualized animations. *Top. Cogn. Sci.* 10, 225–241. doi: 10.1111/tops.12313

Simion, F., Regolin, L., and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proc. Natl. Acad. Sci. U. S. A.* 105, 809–813. doi: 10.1073/pnas.0707021105

Spelke, E. S., Breinlinger, K., Macomber, J., and Jacobson, K. (1992). Origins of knowledge. *Psychol. Rev.* 99, 605–632. doi: 10.1037/0033-295X.99.4.605

Spelke, E. S., Phillips, A., and Woodward, A. L. (1996). "Infants" knowledge of object motion and human action," in *Causal Cognition*, eds D. Sperber, D. Premack, and A. J. Premack (Oxford University Press), 44–78.

Stewart, J. A. (1982). *Perception of Animacy* (dissertations thesis). University of Pennsylvania.

Surian, L., Caldi, S., and Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychol. Sci.* 18, 580–586. doi: 10.1111/j.1467-9280.2007.01943.x

Tavares, P., Lawrence, A. D., and Barnard, P. J. (2008). Paying attention to social meaning: an fMRI study. *Cereb. Cortex* 18, 1876–1885. doi: 10.1093/cercor/bhm212

Tomasello, M. (2001). *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.

Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–951. doi: 10.1068/p3101

Uller, C. (2004). Disposition to recognize goals in infant chimpanzees. *Anim. Cogn.* 7, 154–161. doi: 10.1007/s10071-003-0204-9

Versace, E., Ragusa, M., and Vallortigara, G. (2019). A transient time window for early predispositions in newborn chicks. *Sci. Rep.* 9, 18767. doi: 10.1038/s41598-019-55255-y

Visch, V. T., and Tan, E. S. (2009). Categorizing moving objects into film genres: the effect of animacy attribution, emotional response, and the deviation from non-fiction. *Cognition* 110, 265–272. doi: 10.1016/j.cognition.2008.10.018

Vogeley, K., and Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7, 38–42. doi: 10.1016/S1364-6613(02)00003-7

Vogt, B. A., Vogt, L., and Laureys, S. (2006). Cytology and functionally correlated circuits of human posterior cingulate areas. *Neuroimage* 29, 452–466. doi: 10.1016/j.neuroimage.2005.07.048

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wood, J. N., Glynn, D. D., Phillips, B. C., and Hauser, M. D. (2007). The perception of rational, goal-directed action in nonhuman primates. *Science* 317, 1402–1405. doi: 10.1126/science.1144663

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34. doi: 10.1016/S0010-0277(98)00058-4

Yang, D. Y.-J., Rosenblau, G., Keifer, C., and Pelphrey, K. A. (2015). An integrative neural model of social perception, action observation, and theory of mind. *Neurosci. Biobehav. Rev.* 51, 263–275. doi: 10.1016/j.neubiorev.2015.01.020

Yeshurun, Y., Nguyen, M., and Hasson, U. (2021). The default mode network: where the idiosyncratic self meets the shared social world. *Nat. Rev. Neurosci.* 22, 181–192. doi: 10.1038/s41583-020-00420-w

frontiers | Frontiers in Psychology

TYPE Hypothesis and Theory
PUBLISHED 09 September 2024
DOI 10.3389/fpsyg.2024.1371538

EDITED BY
Michael J. Serra,
Texas Tech University Health Sciences Center,
United States

REVIEWED BY
Francesco-Alessio Ursini,
Central China Normal University, China
Prakash Mondal,
Indian Institute of Technology Hyderabad,
India

*CORRESPONDENCE
Bruno Laeng
✉ bruno.laeng@psykologi.uio.no

RECEIVED 16 January 2024
ACCEPTED 23 August 2024
PUBLISHED 09 September 2024

CITATION
Lobben M and Laeng B (2024) Zooming in and out of semantics: proximal–distal construal levels and prominence hierarchies. *Front. Psychol.* 15:1371538.
doi: 10.3389/fpsyg.2024.1371538

# Zooming in and out of semantics: proximal–distal construal levels and prominence hierarchies

Marit Lobben and Bruno Laeng*

Department of Psychology, University of Oslo, Oslo, Norway

We argue that the "Prominence Hierarchy" within linguistics can be subsumed under the "Construal Level Theory" within psychology and that a wide spectrum of grammatical phenomena, ranging from case assignment to number, definiteness, verbal agreement, voice, direct/inverse morphology, and syntactic word-order respond to Prominence Hierarchies (PH), or semantic scales. In fact, the field of prominence hierarchies, as expressed through the languages of the world, continues to be riddled with riddles. We identify a set of conundrums: (A) vantage point and animacy, (B) individuation and narrow reference phenomena, (C) fronting mechanisms, (D) abstraction, and (E) cultural variance and flexibility. We here propose an account for the existence of these hierarchies and their pervasive effects on grammar by relying on psychological Construal Level Theory (CLT). We suggest that both PH and CLT structure the external world according to proximity or distance from the "*Me*, *Here* and *Now*" (MHN) perspective. In language, MHN has the effect of structuring grammars; in cognition, it structures our lives, our preferences, and choices.

KEYWORDS

**prominence hierarchies, construal level theory (CLT), split ergativity, differential object marking (DOM), nominal classification, culture-language interpretive matrix, abstraction processes, animacy hierarchy**

## Introduction

We propose a domain-general explanation for linguistic prominence effects by merging two independent research traditions that started independently, about 50 years ago, within psychology (Schmitt, 1972; Vinokur et al., 1975; Trope and Burnstein, 1975) and linguistics (Silverstein, 1976). These fields of research have remained apparently oblivious to another, although—as we shall argue—they do relate to similar focus areas: Decision making within the field of social psychology, and a grammatical hierarchy known as "animacy hierarchies," displayed in various structures of transitive sentences in highly unrelated languages.

We note that human decisions are often made in the moment, by an agent from the perspective of an ego, as well as with perspectives that an event may take place in the future. Also, according to linguistics, actions emanate in the unmarked case from a single individual and have consequences beyond the originator's close sphere and further away in time–space. In both cases, the protagonist foresees a mental timeline. The distances traversed may differ in the two cases, but the trajectories are essentially the same, with an origin near a cognizing individual and an end state at some distance from this point, with wide-reaching effects on cognition. At close range, details, means, emotions, peers, time, and place matter. Immersed in a context, a subject cannot see the situation without bias. Transitory emotions are in focus rather than far-reaching often abstract, goals. From a distance, by contrast, the individual can view the essential similarities and differences.

Social psychologists have acknowledged the above aspects in a formal theory on mental construal levels, based on a wide array of empirical experiments (Soderberg et al., 2015). Linguists too, are on their way to understanding that the *prominence* that these grammatical hierarchies spring out from, has to do with an egocentric viewpoint, and not primarily, as often believed, with animacy and other *manifestations* of the hierarchy (Gardelle and Sorlin, 2018). Bridging the gap between linguistics and cognition is in line with a fundamental stance that whatever takes place in mental space is likely to influence the structure of language, in line with the fundamental tenets of cognitive linguistics (Tsoneva-Mathewson, 2009, p. 346):

> "Cognitive linguistics encompasses a number of broadly compatible theoretical approaches to linguistic meaning and structure that share a common basis: the idea that language is an integral part of cognition, and it reflects the interaction of cultural, psychological, and communicative factors which can only be understood in the context of a realistic view of conceptualization and mental processing."

Our argument runs as follows: We first attempt to explain the interdependence between markedness and prototypicality and how this relationship gives rise to prominence hierarchies. We then describe how cross-linguistic variance of seemingly arbitrary borderlines in category structure, as well as cut-off points in grammatical hierarchies, behave as gradient phenomena, and thus formulate a hypothesis based on Construal Level Theory (CLT) that specifies converging aspects of the two fields. From this point onwards, we go on to formulate a set of unsolved questions in prominence phenomena, subsumed under the label "conundrums." To replenish the identified gaps, we concurrently present empirical support for our analysis and propose how prominence phenomena and cross-linguistic differences in category structure can be explained in terms of "psychological distance" within CLT. Finally, we summarize our findings and discuss briefly how our proposal intersects with previous linguistic analyses, as well as formulating predictions of our proposal as cases for future research.

## Markedness and prototypes

Since prototypical categorization (Rosch, 1978) is based on human experience and cognition rather than objective, mind-independent criteria (Geeraerts, 2016), it becomes ubiquitous in language (Taylor, 2003). Note also that prototypical category membership is not binary but exist on a continuum. The co-existence of peripheral and central members is critical to linguistic systems since it permits a high degree of flexibility for cognitive development. While peripheral members possess less of the features and properties associated with the category, prototypical members constitute a core which exhibit the most characteristic features. Importantly, these prototypes serve as cognitive reference points, or anchors, against which new potential category members are compared (van der Auwera and Gast, 2010). As people make new experiences, their mental representations of categories can shift, and prototypical structures evolve over time by means of overlap in features, metaphorical extension, and influence from technological or social changes. Semantic category structure may therefore also eventually be influenced by cultural mindsets (Aikhenvald, 2000, pp. 347, 421).

Prototypicality and markedness are two sides of the same coin. Markedness is what deviates from the prototypical (Croft, 1990, pp. 124–154) or the state of standing out as nontypical or divergent, as opposed to the regular or common. To define markedness, it then becomes paramount to identify the prototypical elements in language. This is done by observing inequalities in structural, behavioral, and frequency data (Greenberg, 1966), both cross-linguistically and within languages. In linguistic marked–unmarked relations, one term of an opposition is the broader, dominant, or typical one (known as *unmarked*); the other one is *marked* and may involve extra morphology or more complex semantics. For example, in the morphosyntactic category of Number, the singular is normally unmarked whereas the plural is marked, since the notion of plurality arises out of adding semantics and morphology to a less complex item. It is this asymmetry aspect that conceptually links markedness to the multivalued categories in implicational universals and grammatical categories: The essential idea behind markedness in typology is the "asymmetric or unequal grammatical properties of otherwise equal linguistic elements – inflections, […] or even syntactic constructions" (Croft, 1990, p. 64).

Markedness may be cancelled in particular contexts (Battistella, 1996, pp. 37, 144), in which the untypical becomes the typical. This is what is known as *markedness reversal* (Aissen, 1999, pp. 679–680) or local markedness (Croft, 1990, p. 66). Such reversals are affected by extralinguistic factors. For example, relative to plurals, singulars are marked in mass nouns (e.g., *oats*, *salt*) and with items typically occurring in pairs (e.g., *paired body parts*) (Tiersma, 1982). Croft (1990, p. 66) explains this behavior as "objects that naturally occur together or are difficult to individuate," a way of thinking pertinent to the definition we adopt for prototypical transitivity.

## Markedness in sentences

Languages across the world encode verbal arguments differently, either by grouping the subject of intransitive verbs (S) and the agent of transitive verbs (A) together, with nominative case against the object (O), marked with accusative case, resulting in the pattern $S = A \neq O$. Ergative-absolutive languages instead encode S and O in the same way, with absolutive case against A, which is ergative, yielding the pattern is $S = O \neq A$. In a large number of languages, these systems intersect in the so-called split case systems. A widely accepted explanation emerges from the observation that, although A and O in transitive events in principle can be both definite and animate, in actual discourse, A tends to be animate and definite, and O to be inanimate and indefinite (Comrie, 1989, p. 128). In other words, transitive action typically runs from a definite and animate A towards an indefinite, inanimate O. This pattern likely reflects frequencies of how human interactions typically play out outside of language (Comrie, 1986, p. 104), or mirrors what humans think is relevant to report (Payne, 1997, p. 151). Markedness is inversely correlated with prototypes, so that any deviation from this pattern of "natural kind of transitive construction" leads to a more marked construction (Comrie, 1989, p. 128). Skewed frequency of preferred referents for subjects and direct objects (DO) of "who does something to whom or what", creates a cross-linguistic pattern of *unmarkedness*, where it is unmarked (typical) for participants higher up in a PH to inhabit the role of subjects, but marked for participants lower in the PH, as illustrated in Figure 1.

Here, "Prototypical transitivity" contrasts with accounts of transitivity that focus on grammatical features that advance efficient transfer of action ("kinesis") from A to O, where instead O needs to be highly individuated, defined as animate and/or definite (Hopper and Thompson, 1980, p. 256). Given Greenberg's (1966) textual and cross-linguistic frequency criteria for protypes/markedness (Croft, 1990, pp. 71–72, 84–89, 92), Comrie's definition of prototypical transitivity is in fact more in line with general prototype/markedness criteria. In this context, an indefinite and/or inanimate A, as well as a definite and/or animate O, would bring about more marked constructions. The former case is referred to as differential subject marking (DSM) (de Hoop and de Swart, 2008), and the latter is known as differential object marking (DOM) (Bossong, 1991; Aissen, 2003). These alignment shifts imply that some languages known as *split case languages* may shift between a nominative-accusative pattern and an ergative-absolutive pattern, when dealing with untypical subjects and objects to make explicit their syntactic functions. Nominative-accusative systems often do not overtly case-mark the subject, since agency is in focus, whereas ergative systems, being oriented towards the patient role, leave absolutive participants zero-marked, motivated by the fact that intransitive subjects are often a bit patient-like (e.g., *melt*). Languages may still differentially case mark agent-like intransitive subjects that are more volitional and in control (e.g., *jump*). It is this alternative systemic focus on agents vs. patients, respectively, that makes each system apt for split case marking at opposite ends of the action chain. Note that languages that always mark their direct objects with a certain case (e.g., accusative case), or never mark them, do not fall under the scope of DOM (Aissen, 2003), and are not in need of an explanation of how PH works.

Other syntactic constructions affected by PH include direct/inverse morphology and voice (active vs. passive) (Aissen, 1999). The primary function of passive voice is to shift the focus away from the agent performing the action to the patient receiving it by raising the patient/object in the corresponding active sentence to subject position, e.g., to maintain cohesion in discourse by focusing on a consistent topic. Direct/inverse morphological systems mark the relative position of arguments on a prominence hierarchy. Direct morphology applies when the subject referent outranks the object on the hierarchy, whereas inverse morphology applies if the speaker wishes to change this ranking order. This operation can be expressed by affixes on clausal participants and/or verbs telling who is "proximal" and who is "obviative" among two 3rd-person participants. The referent considered less important within the construction or in discourse is marked obviative, e.g., the possessed (compared to the possessor), or the inanimate (compared to the animate). As the marking depends on hierarchical status rather than just grammatical role, these systems do not fit neatly into traditional alignment categories like nominative-accusative or ergative-absolutive. Direct/inverse systems are relatively

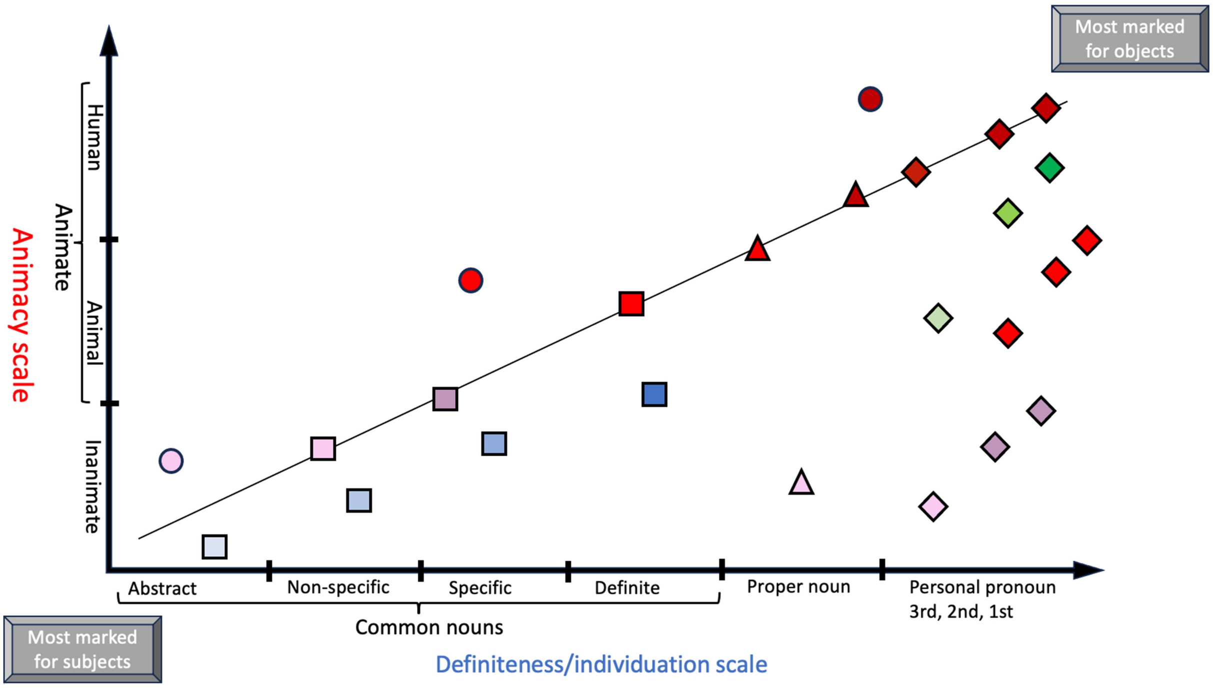


FIGURE 1
A diagram showing relative markedness of subject and object on the dimensions of animacy and definiteness, separately and combined. The placement of symbol on the plot line signifies the assumed likelihood that DOM will occur, given that the two dimensions mutually strengthens DOM likelihood (Aissen, 2003). (Note that the positioning of symbols is approximate and do not reflect data analysis). Key to symbols: Circle = animacy; square = common noun; triangle = proper noun; diamond = personal pronoun. Color key: Dark red = human; crimson red = animal; pink = inanimate; dark blue = definite noun; lighter blue = specific noun; even lighter blue = non-specific/indefinite noun; lightest blue = abstract noun; dark green = 1st person pronoun; lighter green = 2nd person pronoun; lightest green = 3rd person pronoun. Combinations of color and shape symbol shows two-dimensional DOM, i.e., the value is determined by both animacy and definiteness scales (e.g., crimson red square = animal common noun).

rare cross-linguistically, and some linguists have proposed to subsume some types of direct/inverse systems under related phenomena like topic fronting, e.g., Jacques and Antonov (2014).

## Gradience

When assigning case to a clausal participant, split case languages consider not just syntactic functions like subject and object, but also the inherent properties of the referent. These properties are a combination of linguistic and extralinguistic factors. The linguistic features include definiteness/specificity and whether the referent is rendered as a noun or pronoun; the extralinguistic aspects consider if the referent is human, animate, or inanimate, or in some languages, an abstract concept. These properties form clusters of features occupying opposite ends of gradient scales, as proposed in Croft (2003, p. 127) in the Extended Animacy Hierarchy, which really consisted of three hierarchies conflated into one, see 1a-c). A definiteness scale is proposed by Aissen (2003, p. 444) and Croft (1988, pp. 163–164), see 1d).

1) a) Person: 1$^{st}$, 2$^{nd}$ > 3$^{rd}$
   b) Referentiality: pronoun > proper name > common name
   c) Animacy proper: human > animal > inanimate
   d) Definiteness scale: personal pronoun > proper noun > definite NP > indefinite NP

Generally, individuals towards the high (i.e., leftmost) end of prominence hierarchies tend to be definite while participants towards the low end of the prominence hierarchy are indefinite, with those at the far left always being definite (Dixon, 1994, p. 91). Further, a referent may be identified, but not explicitly specified, in which case they are specific, but not definite. A clustering of definiteness appears with the higher values of animacy (Croft, 1990, p. 127): A human is more often definite than an animal, and animate individuals are more often definite than inanimate objects. Pronouns likewise rank high since they per definition have definite reference (Dixon, 1994, p. 91). Nouns, by contrast, may be either definite or indefinite, and animate or inanimate, and therefore appear lower in the hierarchy. Thus, a challenging aspect with the hierarchy underlying split case languages is the lack of a clear, unambiguous boundary along these scales and instead the presence of a continuous spectrum of meanings. This also presents the second challenge, *viz.* that PH are determined by multivariate feature clusters, accounting for tendencies rather than absolute behaviors. Furthermore, PH may respond to a single parameter (one-dimensional DOM), or several parameters combined (two-, or multidimensional DOM), see Figure 2.

A second challenge is that split case languages case-mark direct objects in response to these dimensions at varying cut-off points along a fixed value scale. DOM may be optional, obligatory, or

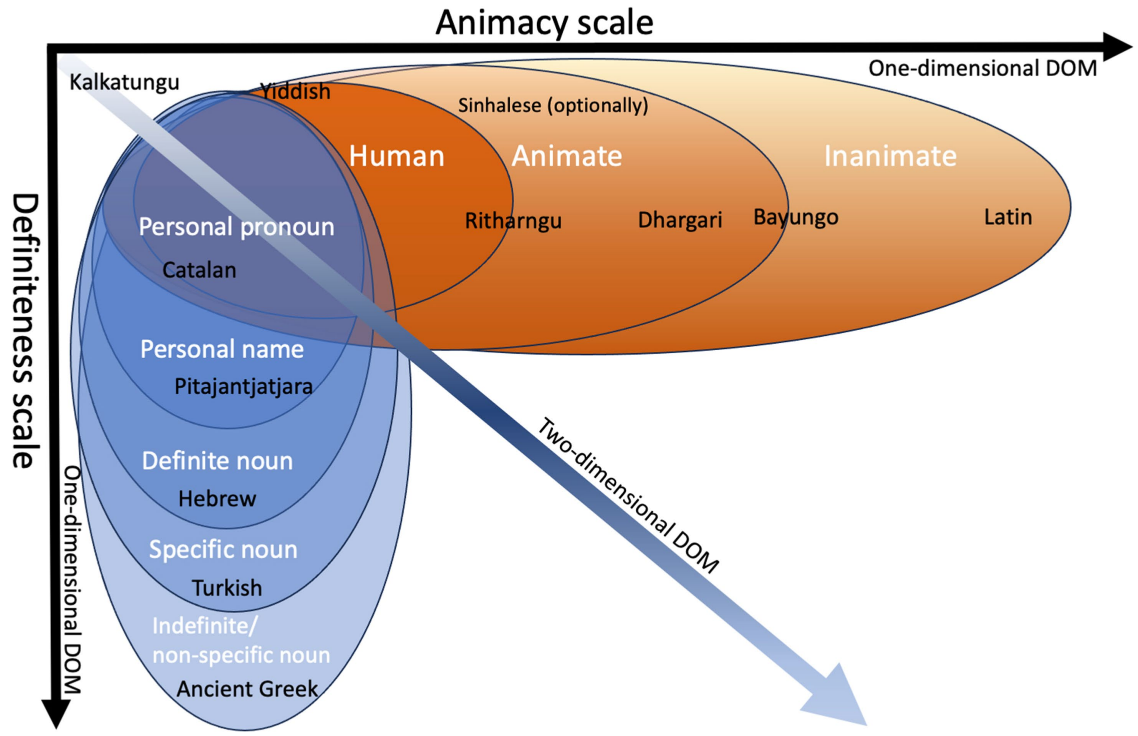


FIGURE 2
Cut-off points for a set of languages responding to DOM phenomena along one of the dimensions, animacy and definiteness. Kalkatungu never case-marks direct objects for either dimension; Yiddish case-marks some human objects only; Singhalese case-marks all animate objects, but only optionally; Ritharngu marks all human and some animate objects; Dhargari marks all animate objects; Bayungo marks all animate and some inanimate objects; Latin case-marks all direct objects. Along the definiteness scale, Catalan case-marks only personal pronoun objects; Pitajantjatjara marks only pronoun and proper name objects; Hebrew marks only pronoun, proper noun (personal name) objects; Turkish case-marks all objects except non-specific nouns (see Aissen, 2003), and finally, Ancient Greek case-marks all direct objects (but with a different case for non-specific nouns; Mardale and Karatsareas, 2020).

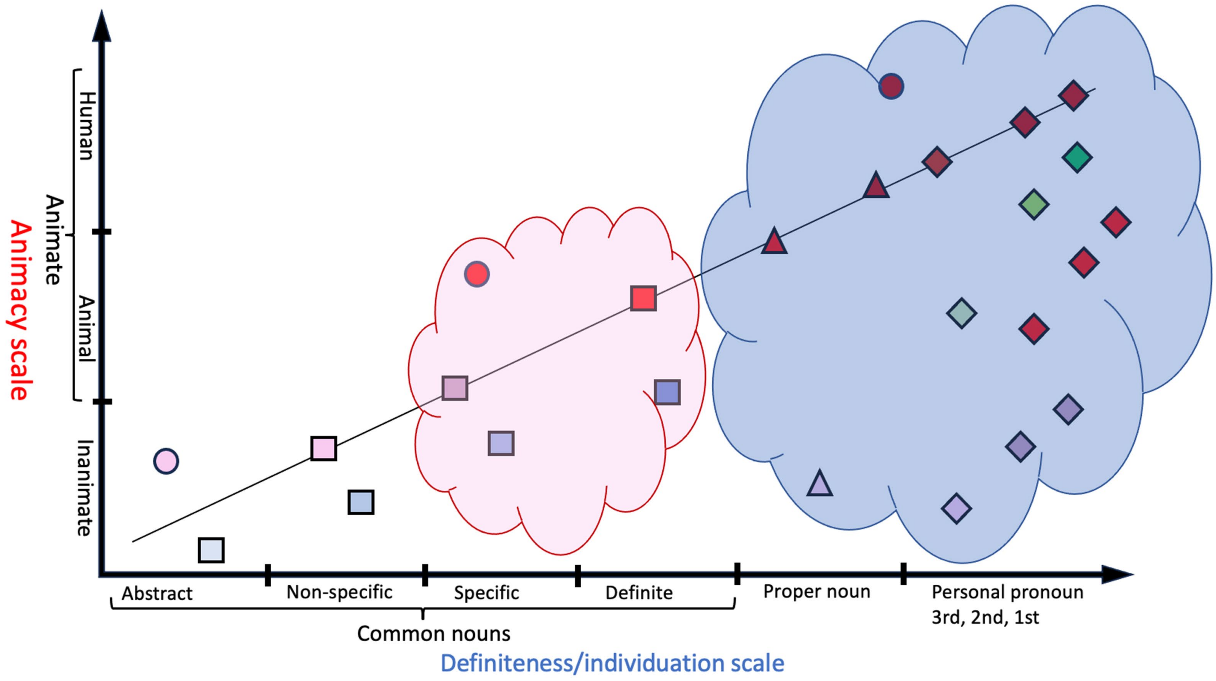


FIGURE 3
An example of DOM in Romanian of how cut-off points for case may be contiguously organized for obligatory (blue cloud), optional (pink cloud) and non-use (no cloud) (Farkas, 1978). Key to symbols: see Figure 1.

excluded, along different portions of a prominence hierarchy. For example, in Sinhalese, case-marking is optional, but only animate-referring objects may be case-marked (Gair, 1970). A second type is Hebrew, in which case-marking is obligatory, but limited to definite objects (Givon, 1978). Finally, a third type of which Romanian is an example, obligatorily marks the case for one group of objects (e.g., at the top of the hierarchy), optionally marks case for an in-between group, and bans case-marking for a third group (e.g., at the bottom of the hierarchy; Farkas, 1978). Notably, while optionality and cut-off points vary from language to language, the hierarchies are consistent in the sense that the order of categories are largely the same across languages (Aissen, 2003, pp. 436–437). Moreover, the parts of the scale where case-marking is mandatory, optional, or absent in each language, are contiguous along the scales (Figure 3).

Grammatical categories display various kinds of organization relative to how categories of species are carved out in biology. Surprisingly, extensions beyond and retractions within the human species and animal taxa in nominal classifier systems, as well as in PH, have been observed in multiple languages. Category structure can be partially culturally determined and cultural knowledge can have an impact on grammar by imposing constraints on morphosyntax (Aikhenvald, 2000, p. 319). Langacker (1994, p. 39) points out that "a specific cultural practice or belief motivates the otherwise unexpected membership of some entity in a conceptually grounded category of grammatical significance." Instead of looking at this as mismatches relative to biological taxa, it is profitable to analyze category membership and cross-linguistic variance in borderlines in terms of what is deemed standard, common, or frequent within a culture, and in principle no different from how prototypes and markedness is defined elsewhere in language.

## Scope

The topic of prominence carves out some limitations to focus areas. First, ranking participants in prominence within a clause can only be done with a minimum of two participants. The original definition of "animacy hierarchy" was made with reference to agent, patient, and transitive propositions (Aissen, 1999, p. 674). Second, a transitive trajectory playing out in space–time requires verbs depicting temporally measurable events. This corresponds to Vendler's (1957) verb class known as "accomplishments," defined as dynamic and durative events with an inherent endpoint. This delimitation can be justified from the viewpoint of prominence as well, although indirectly. Split ergativity, a central facet of prominence hierarchies, correlates with verbs of action, while other types of verbs may follow a different pattern (Tasaku, 1981). In some split ergative languages, the split aligns with semantic properties of verbs, such as the degree of control or volitionally associated with the action (Bohnemeyer, 2004), which again correlates with completed actions. Thus, although the interface between prominence hierarchies and Vendler's verb classes is not a concluded matter, these fields seem to overlap in the sense that they both intersect with the aspectual properties of verbs.

Third, a focus on syntagmatic relationships within sentences is warranted since this is where comparison between clausal participants plays out. A clarification of prominence and markedness with reference to syntagmatic vs. paradigmatic relations might be helpful. The values within a grammatical category that stands in a markedness relation must be logically independent of each other. However, this is not the only requirement and in order to qualify as a markedness pattern, the values in question must be paradigmatic alternatives (Croft, 1990, p. 69). In other words, markedness is restricted to "a relation between features which are mutually exclusive" (Greenberg, 1966, p. 57). For example, an animate direct object stands in a paradigmatic relationship to an inanimate DO, since one and the same sentence cannot have both an animate and an inanimate DO at the same time; they must appear in different sentences. Among the two, the animate is the unusual one and receives overt case marking, e.g., accusative case, the inanimate is unmarked. While being paradigmatic alternatives, these values must also exist at a higher level of abstraction in grammar, in this case the category of "direct object." This means that the markedness and prominence dimensions are orthogonal to each other, but also that paradigmatic opposition is irrelevant to the relative prominence between verbal arguments within a sentence (see Figure 4).

Our analysis will make use of the concept "psychological distance" to explain the above-mentioned challenges, in which temporal distance is one dimension affecting levels of construal. Here, we make explicit that exchanging the present tense with either past or future tense (i.e., more temporally distant from Ego) does not have an effect on the level of construal for the clausal participants. This is because the temporal dimension is measured along the transitive trajectory, as syntagmatic relations, and not along the paradigmatic axis where the various tenses alternate between sentences. This is logical since it is within a particular sentence that participants compete for prominence. Evidently, the axis along which prominence is measured and the temporal dimension *need to converge*, otherwise these two dimensions would not merge as the same psychological distance.

## Preliminary conclusion

Despite prominence hierarchies are ubiquitous in language, it has been notoriously difficult to understand what exactly they are about. It has been challenging to find a common denominator for the whole of the "Animacy hierarchy", as Helmbrecht et al. (2018) point out, and it is not surprising that this hierarchy has been interpreted in different ways in the literature. In fact, a variety of names has been assigned to it since it was first discovered: "Lexical hierarchy" (Silverstein, 1976); "Nominal hierarchy" (Dixon, 1979, 1994, p. 85); "Animacy hierarchy" (Comrie, 1989); "Empathy hierarchy" (Kuno and Kaburaki, 1977; DeLancey, 1981); "Hierarchy of reference" (Zwicky, 1977); and "Prominence hierarchy" (Aissen, 1999; Lockwood and Macaulay, 2012). We therefore propose to cast a broader net by looking at these phenomena within cognitive psychology, specifically within the framework of the Construal Level Theory.

## Construal level theory

In psychology, it has been common to model a variety of mental representations as taking place within a mental or virtual "space" (Shepard, 1962). From the perceptual representation of colors (Krantz, 1975; Smith and Pokorny, 1996; Logvinenko, 2015) to those of human faces (Valentine et al., 2016; Leopold et al., 2001) or of emotional states (Russell, 1980, 2003; Posner et al., 2005), the stimuli corresponding to values or "points" within a multidimensional space, often defined by only three orthogonal polarities (as for color space) or even just two (as for the emotional circumplex). In some of these models, there can be a central point that corresponds to an average or neutral point, where values along the dimensions cancel each other out (as the central achromatic grey in color space; or the sexless, unemotional, and ethnically-hybrid prototypical face). Distance between positions in such spaces indicates how features or entities differ or the degree in

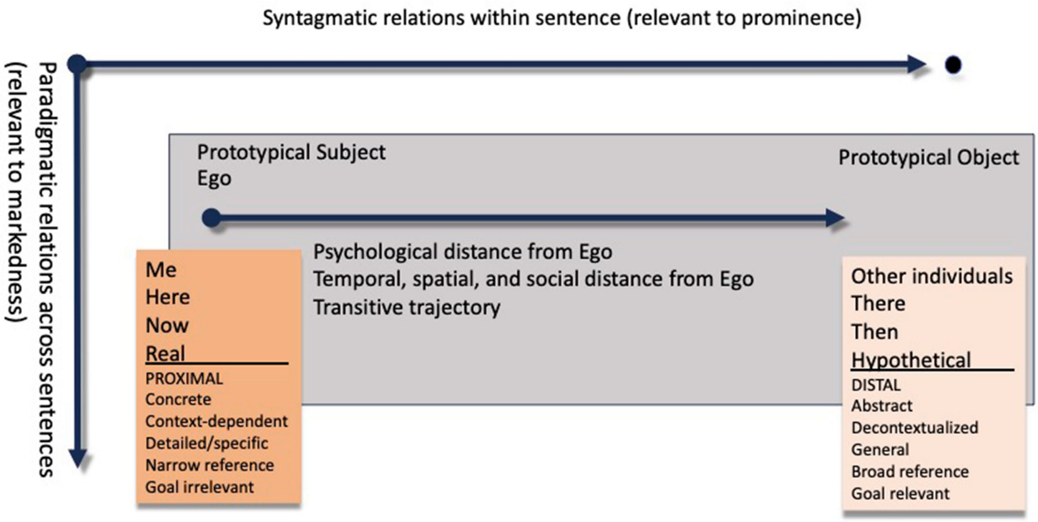


FIGURE 4
Syntagmatic vs. paradigmatic relations and prominence within a transitive sentence within CLT. Temporospatial distance from ego refers to a verbal action towards the direct object with an inherent endpoint (accomplishment).

which they are proximal/distal from each other or whether they are opposites (e.g., as with opponent colors).

However, recent modelling within psychology, where social dimensions are fundamental, posit psychological spaces where the central, or zero, locations are always in relation to self and one's current spacetime coordinates; that is, Ego and the "here-and-now". Increasingly dissimilar features from the Ego or the moment's spacetime coordinates are construed as increasingly mentally distant events or things, as well as increasingly abstracted away from the present and egocentric viewpoint. One such account is known as the "Construal Level Theory" (Trope et al., 2007; Trope and Liberman, 2010, 2012), which would seem most relevant from a linguistic perspective. This account derived from the simple observation that "distance" and level of abstraction appear to be related in people's minds. That is, the model makes explicit that, while distance can happen in different domains (spatial, temporal, social) as well as in terms of hypotheticality, all these dimensions are associated and affect one another, thus subsumed under the unified concept of *psychological distance*. Such a mental distance also implies the subjective experience or feeling that something is close or far away, similar, or dissimilar, in time, space, and social distance from self.

Specifically, the central or zero psychological distance corresponds to the most concrete and most similar situation or object to a Self or Ego situated in the *specious present*, the time duration wherein one's perceptions are considered to be in the present (Kelly, 1882; James, 1893). As thoughts or perceptions are further removed from the Self and the Here-and-Now, or from material reality, mental construals become increasingly distant and, importantly, increasingly abstract. They zoom out from the specific, detailed, subordinate, contextual aspects of a situation or class of objects and extract only the essential, often goal-relevant, features. Thus, mental construal seems a useful cognitive tool that allows human minds to mentally maneuver events and objects in past or future contexts, rather than only living in the present, and to refer to hypothetical and counterfactual events rather than the confines of the real and material situation. Traversing along psychological distance enables people to cross the distances that separate the Self from the Others, the Now from the Then, and Here from There.

Construing psychologically distant or distal events allows the capacity for "mental travel" (Soderberg et al., 2015) and, together with the process of abstraction from a single, concrete, event, may constitute the key mechanism for going beyond the immediate or proximal experience, like when reminiscing, speculating, and making predictions. Importantly, according to CLT (Liberman and Trope, 2008), psychological distance influences the way we represent the world, how we categorize it and communicate about it with others. With respect to social relationships, given that psychological distance is defined in CLT as the extent of divergence from the direct experience of me (Liberman and Trope, 2014), this results in a hierarchical scaling of Self versus proximal and then increasingly distal others, along socially perceived similarity/dissimilarity with other individuals (e.g., in culture, attitudes, appearance), familiarity versus unfamiliarity, as well as a hierarchical scaling of the others in relation to the ingroup (e.g., kin, clan) versus the outgroup classes. According to CLT, the mind understands psychological distance abstractly and literally, in social as well as in physical terms, operating along continuous scales, often without a fixed maximal point of distance. As psychological distance increases away from the proximal to the distal, peoples' representations of objects and events become ever more abstract.

CLT was forged from the results of an impressive number of empirical studies. For example, researchers found not only that proximal concepts triggered concreteness and individuation, but also that this individuation is linked to the first-person perspective. By using an implicit association test, first person pronouns like "ours, ourselves, at our place, for us," and "we" were associated with words that denoted exemplars (*beet, poodle, belt*), while third person pronouns "they, theirs, for them, at their place" were instead associated with categories (*vegetables, animals*, and *clothes*) (Bar-Anan et al., 2006, experiment 3B). Subjected to the same kind of test, participants non-consciously associated socially proximal concepts like "friends, parents, buddies," and "siblings" with concrete words, on the one hand, and socially distal concepts like "enemies," "strangers," "opponents," and "anonymous person" with abstract concepts, on the other (Bar-Anan et al., 2006). Thus, there seems to be a bidirectional, interdependent relationship between level of construal and social distance. Participants who explained individuals' behavior using global dispositional qualities also tended to perceive them as more socially distant, compared to participants who explained the very same behavior in terms of concrete, situational factors (Stephan, 2006: see Bar-Anan et al., 2006).

Pointing to the origin of the concreteness factor, proximal level construals are context-dependent while distal construals are decontextualized. In one study, people's construal of distant-future activities stated the goals of the activities, whereas the construal of near-future activities stated the means and/or the spacetime coordinated for achieving these goals (Liberman and Trope, 1998). The reality people experience when immersed in a specific context are typically more detailed, as they are tied up with practicalities and the "how" aspects of the activities. These subordinate-level construals preserve a stimulus in minute detail, while emphasizing its unique features rather than focusing on a situation's similarity to other stimuli. By contrast, decisions based on distal level construals along the temporal dimension are supervised by desirability concerns—the why's—while downplaying feasibility, potential contextual constraints, and the means necessary for enactment. Because a distant object is decontextualized, only the gist of available information, the superordinate time-stable core features of the event, are mentally represented and considered, which often results in planning fallacies (Trope and Liberman, 2012).

In this research, participants use broader categories in distal than proximal level construals across dimensions. In the social dimension, people tend to describe outgroups using more abstract qualities compared to ingroups, (Fiedler et al., 1993; Werkman et al., 1999), and their properties as more structured and predictable (Linville et al., 1996). Outgroups are perceived as less differentiated into subgroups (Brewer and Lui, 1984; Linville, 1982; Park et al., 1992), and also as more homogenous than ingroups (Jones et al., 1981; Park and Judd, 1990; Park and Rothbart, 1982). As in the social dimension, participants tested in the temporal dimension used broader categories to classify objects for distant-future than for near-future situations, which were instead organized in narrower categories of concrete objects more unstructured and incidental (Liberman et al., 2002, Study 1). Likewise, in the probability dimension, improbable events, being removed from direct experience, seem more distal and instigate participants to categorize objects in fewer, broader groups than those imagining probable events (Trope et al., 2007).

Fewer dimensions underlie people's judgment of temporally distal than proximal events. In an event-rating task, distant-future preferences could be statistically accounted for within two- to four-factor solutions, whereas the near-future preferences always required one factor more to account the same amount of variance (Liberman et al., 2002, Study 2). Thus, near future preferences proved to be more complex, harder to reduce to general underlying dimensions, and were determined by a larger set of distinct factors than the corresponding distant-future preferences, which represented preferences in simpler structures.

In line with this simplification, distant-future events seem to represent more prototypical cases than near-future events. Participants were asked to write down and valence the events they expected to experience, during a good or bad day, in either the near future (tomorrow) or the distant future (a day a year ahead) (Liberman et al., 2002, Study 3). The near future events were described as more diverse than in the distant-future day, and prototypical and more extreme experiences were expected in the distant than in the near future. The researchers concluded that distant-future experiences were more schematic, since intracategory homogeneity was greater, and there was greater intercategory divergence (Trope and Liberman, 2003). In sum, as temporal distance increases, future events are represented more parsimoniously and with greater abstraction.

Several experiments suggest the four dimensions are cognitively interrelated and integrated. For example, abstraction increases with distance and decreases with proximity when the temporal and social dimension are conjoined. Personality ratings were more similar across social roles when participants were thinking of themselves on a day a year later than on the following day. The distant self was thought of in a more integrated, structured manner, and the near self was more contextualized and fluid (Donahue et al., 1993). Another study (Nussbaum et al., 2003) exploited the tendency people have to identify behaviors from underlying, dispositional traits (attribution theory; Trope and Burnstein, 1975). With increased temporal distance, participants were more likely to attribute behavior to personality traits rather than to situational demands. Reasoning that traits are tokens of generalized representations, they found that more abstract, higher-level, construals were used to predict and explain distant-future behaviors. Moreover, the finding that first-person pronouns were associated with exemplars and third-person pronouns with categories translated to the temporal dimension as well, since the same exemplar/category pronominal stimuli were used in an experiment on associations with "near time" (*a second, a minute, now, immediately, soon*) vs. "distant time" (*a year, a decade, later, last year, long ago*) (Bar-Anan et al., 2006, experiment 1B).

Thus, whether investigated separately or in combination, the dimensions produce similar results and participants organize items in broader and more general categories when perceived or imagined at greater distances compared to in proximal level construals. Table 1 summarizes CLT results and predictions.

Hence, we hypothesize that prominence phenomena (hierarchies and categories) fall into the same kind of grid as here described within CLT. Next, we will look at some of the challenges that former analyses have created in presenting a unified understanding of these phenomena.

## Unsolved conundrums of prominence hierarchies and category structure

In this section we organize previously researched aspects of prominence into five overarching topics: A. vantage point and animacy, B. individuation and narrow reference phenomena, C. fronting mechanisms, D. abstraction, and E. linguistic aspects of cultural variance and flexibility. Each of these umbrellas subsume a variety of manifestations in grammars which we outline and exemplify under each point. Notably, several conundrums arise because previous understandings of PH and linguistic categorization are conceptually incoherent on a superordinate level or explain only a portion of PH behavior.

## Vantage point and animacy

### The first conundrum: why are speech-act participants exclusive in prominence hierarchies?

A prominence hierarchy is modified by a person > non-person contrast along the gradient 1st > 2nd > 3rd person pronouns ("I" > "you" > "he/she/it" in the singular). Person hierarchies manifest themselves most famously in direction-marking systems, which mark transitive actions that comply with the prominence hierarchy (i.e., the agent never ranks higher than the patient), with a direct verbal affix, and actions that contradict the prominence hierarchy as "inverse" (Croft, 1990, pp. 136–137). These markers are neutral with regard to the syntactic roles S, A, and O (the addition of inverse affixes do not change the syntactic functions relative to the direct construction), and do not appear on intransitive verbs (Jacques and Antonov, 2014). Direction marking languages leave the direct configuration unmarked, or mark both, as in the Tibeto-Burman language Jyarong, where the direct suffix *-a* occurs with speech-act participants (1st and 2nd person, or SAP) acting on 3rd person participants (SAP →3rd), and the inverse *-uk* suffix occurs with 3rd → SAP.

The most common pattern cross-linguistically distinguishes SAP from all other noun phrases, including 3rd person pronouns

**TABLE 1** Construal level characteristics at proximal vs. distal distances of events and objects (Soderberg et al., 2015; Trope and Liberman, 2003).

| Proximal-level construals | Distal-level construals |
|---|---|
| Concrete | Abstract |
| Complex | Simple |
| Exemplars or narrow categories | Broad categories |
| Individualized information | Dispositional information |
| Details | Gestalts |
| Subordinate | Superordinate |
| Secondary features | Primary features |
| Specific behaviors | Broad traits |
| Situational information | Aggregate information |
| Situation-specific demands | Goal relevant |
| Contextualized | Decontextualized |
| Feasibility concerns | Desirability concerns |

(DeLancey, 1981, p. 628). Number, DOM, and direct/inverse systems all give priority to the 1st and 2nd person pronouns, or the individuals engaged in communication, one as the speaker and the other as the listener. In no known language are 3rd person pronouns grammatically ranked over SAP. Languages generally rank 1st and 2nd equally or rank 1st > 2nd (but the reverse order, 2nd > 1st, exists in Algonquian languages). Still, a noteworthy universal constraint applies to the internal ordering of SAP: the single option not attested in any language so far is inverse marked 1st → 2nd while 2nd → 1st is marked as direct, which warrants primacy to 1st person pronouns over 2nd (Jacques and Antonov, 2014).

As examples of joint vs. gradient SAP ranking types, consider these two Tibeto-Burman languages. In Kham grammar, agents are conceived of as unmarked (2a-c), but 3rd person is not and needs to be marked for agency with ergative case (2d). Participants raised from a lower PH-position to a role reserved for the higher-ranking SAP, need special morphological marking to be accredited this status. Conversely, 1st and 2nd person need to be agreement marked as DO in object position (2a, 2b), but a 3rd person DO does not (2c).

2) a) nga: nən-lay nga-poh-ni-ke.
I you.OBJ 1st.Ag-hit-2nd.Pat-PERF
'I hit you.'

b) nən nga-lay nə-poh-na-ke.
You 1st.OBJ 2nd.Ag-hit-1st.Pat-PERF
'You hit me.'

c) nən no-lay nə-poh-ke.
You he.OBJ 2nd.Ag-hit-PERF
'You hit him.'

d) no-e nən-lay poh-na-ke-o.
He-**ERG** you.OBJ hit-**2nd.Pat**-PERF-**3rd.Ag**
'He hit you.'

In contrast, for the Nocte language, the hierarchy 1st > 2nd > 3rd person is predicted. Agreement is marked on the verb, always prioritizing SAP over 3rd person: agreement with 3rd person occurs only when no SAP are present. The two SAP are nevertheless internally ranked with the PH headed by 1st person, since both 2nd → 1st (3b), and 3rd → 1st (3c) requires the inverse suffix *-h*:

3) a) nga-ma nang hetho-e
I-ERG you teach-1PL
'I will teach you.'

b) nang-ma nga hetho-**h**-ang
You-ERG I teach-**INV**-1st
'You will teach me.'

c) Nga-ma ate hetho-ang
I-ERG he teach-1st
'I will teach him.'

d) Ate-ma nga-nang hetho-**h**-ang
He-ERG I-ACC teach-**INV**-1st
'He will teach me.'

Since DeLancey (1981, p. 644), direct/inverse systems have often been described within an Empathy hierarchy, arguing that discourse participants represent attentional "natural starting points," while also motivated by the proximal/obviative distinction in 3rd person in these languages (Jacques and Antonov, 2014, p. 304), see 4):

4) SAP > 3rd person pronoun > human > animate > inanimate

A typical conversation between single individuals in face-to-face interaction takes place with some asymmetry towards the 1st person as the initiator: the speaker has the primary role of delivering information and guiding the conversation, and decides what to share, controlling the flow of information. This perspective is qualitatively different from other perspectives (Tressoldi et al., 2017).

The primacy of 1st person pronouns in PH also finds a parallel in CLT. Across a wide array of experiments, participants used their own vantage point as basis for conceptualizations about events, objects, and actions with abstractions occurring in increasing psychological distances from Self (Trope and Liberman, 2010). Psychological distance can imply temporal (Trope and Liberman, 2003; Wakslak et al., 2008), or spatial distance (Henderson et al., 2006), unlikeliness of occurrence (Wakslak et al., 2006), and distance along the social dimension, either in terms of an actor being dissimilar or emotionally distant to the perceiver (Liviatan et al., 2008). All of these dimensions matter in the context of SAP.

How a conversation plays out is acutely context-dependent, hinging on the interlocutor's responses and cues in the physical environment (Dohen et al., 2010). The event is factual and dynamic in the sense that it manifests itself in real time with turntaking participants. Experimental evidence suggests that as one moves away from MHN, this changes. Along the temporal dimension, perceivers put less weight on situation-specific states when predicting others' behaviors in near future than distant future events (Nussbaum et al., 2003). The use of social media dissociates interconnectedness of time, space, addressee specificity, and context-dependency in communication. Joshi et al. (2016) found that addressing psychologically distant versus psychologically close audiences has direct effects on communication. Taking an expansive vs. contractive relational scope of addressees in non-face-to-face dialogs alters language use towards high-level, decontextualized messages that are situationally stable and applicable across contexts.

In addition, the nexus of SAP in grammars can be explained with reference to experiments on face-to-face conversations. The listener's role in a conversation goes beyond that of a "hearer"; it demands a collaborator's role who is "a full partner in creating the dialogue," by the use of facial displays, collaborative, interactive gaze patterns, gestures, and brief vocalizations, even when not taking up the speaking turn (Bavelas and Gerwing, 2011). Face-to-face speech communication is a multimodal process and involves not only linguistic but also psychological, affective, and social interaction (Dohen et al., 2010). Speech-act participants relate to each other as individuals, mind to mind. The interlocutors engage mentally in physical acts as well as in mental and emotional interplay, read each other's facial cues and actively strive to understand the other person's intentions and thoughts through "grounding" processes (Bavelas and Gerwing, 2011). Conversation is a two-way process: distracted addressees not able to collaborate in conversation impair speakers' storytelling (Bavelas and Gerwing, 2011). Overhearers differ from addressees by not partaking collaboratively and they do not contribute to mutual understandings. These experimental data support the cross-linguistic roles of SAP in PH.

Social distance affects first-person vs. third person perspective taking (Trope and Liberman, 2010). Research shows that third-party viewpoints are less context-sensitive than first-person viewpoints: Personal memories of behaviors recalled from a third-person perspective produced dispositional descriptions rather than situational terms, whereas the opposite was true for first-person perspectives (Frank and Gilovich, 1989; Nigro and Neisser, 1983). Research within perspective-dependent recall also revealed that perceivers tend to remember more global, dispositional qualities in recalling events from a third person perspective than from a first-person perspective. Imagining performing an activity from a first-person perspective were more vividly depicted than when participants imagined the same activity from a third-person perspective, which instead brought about more abstract and less detailed reports of activities (Libby and Eibach, 2002, Study 4). In terms of CLT, this means that a third-person perspective imposes more psychological distance and higher-level construals than first-person perspective (Trope and Liberman, 2010, p. 448).

In sum, the existing research supports the idea that distance from 1st person towards 2nd and 3rd person pronouns in PH converges with psychological distance emanating from first-person perspective in CLT, and empirical research on face-to-face communication pairs up with the SAP nexus found in many grammars.

### The second conundrum: why does linguistic animacy differ from biological animacy?

In PH, even finer-grained scales rank animals lower than humans, although human animacy is no different from that of animals, biologically speaking. Despite this conspicuous mismatch, it is still more common among linguists to use the term “the animacy hierarchy” than any other term (e.g., Gardelle et al., 2024; Haude, 2024), tacitly accepting a notion of “linguistic animacy” but not stating what that is. Yamamoto (2006, p. 31) argues that “the General Animacy Scale” is based on a kind of hierarchy of animacy *per se* with the assumed natural taxonomy regarding the hierarchy of “living things,” but also that “hierarchical classification of animate beings and inanimate objects is the product of our subjective view of these entities.”

Following this lead, we propose to integrate a linguistic theory of prominence into a wider theory of cognitive construal to offer an alternative, multi-disciplinary view. Recently, a person-centric bias in cognitive representation was robustly documented as the primary way people understand the external world (Kalkstein et al., 2020). Participants were tested for their tendencies towards perceptual mappings of “object-level identities” or towards “relational mappings,” after they observed pairs of identical scenes with minimal manipulations. It was found that the presence (vs. absence) of a person in a scene lead participants to engage in relational mappings, suggesting that the distinct qualities of objects are secondary to how those objects relate to and interact with the person in the scene. Moreover, they were more likely to construe animals in terms of their relationship to humans than to construe humans with reference to their relationship to animals, suggesting person-centric cognition. Hence, people constitute cognitive anchors as to how a scene should be mentally represented and the meaning of objects as well as animals within a scene, becomes defined by their relationship to that central person, therefore giving primacy to people over animals and inanimate objects (Kalkstein et al., 2020, p. 2).

### The third conundrum: why are direct objects more general, non-specific, and abstract than subjects?

Discussions on prominence hierarchies constantly revolve around the clustering of agency and animacy, definiteness/specificity, and the topicality of clausal participants towards the high end of PH (Croft, 1990, p. 127). This clustering has been neither explained, nor have the causes for the absence of these features in direct objects been sufficiently contemplated. Agency is an interactive notion rooted in social landscapes involving “relational reasoning” that classifies them by the role they occupy in relationship to other objects within an event or scene instead of categorizing entities by their perceptual features (Gentner, 1983; Goldwater and Markman, 2011; Kalkstein et al., 2020). The agent and patient roles in canonical active sentences are well known but their distinct asymmetry has not been explicitly addressed as a potential underlying cause of PH or the protypicality of subjects and objects in transitive sentences.

The notion of agency finds a parallel in CLT-research, where it has been shown that powerful agency generates greater distance to and greater abstraction of target categories. An experiment (Smith and Trope, 2006) found that elevated power increases the psychological distance one feels from others, demonstrating a direct relationship between power activation and abstraction. Participants first completed a writing task that activated the experience of either high or low power; subsequently they completed a categorization task to measure inclusiveness of atypical exemplars. Results showed that high-power primed participants were more inclusive in their categorization than low-powered primed participants, demonstrating that power priming leads to more abstract thinking and thus greater breadth of categorization (Trope et al., 2007). The exact same tendency we see in prototypical transitivity: agents are measured by their potency and correlates with broader and less specific direct objects. Thus, in our present proposal, the distance between agent and patient is represented in terms of a mental timeline from the agent, simultaneously representing temporal (in terms of the time it takes to perform the action), the social distance felt by at powerful agent, and a spatial distance between the agent and the patient prior to action. The remaining clustering features relate to individuation and topicality, to which we turn in the next couple of sections.

## Individuation and narrow reference

Languages have a number of techniques to single out exemplars or limited sets rather than generic groups or categories. One way is to let the addressee know that one has a specific item in mind without disclosing which one (specificity), another technique shares this knowledge with the addressee or assumes it to be common knowledge (definiteness). Another way actively points out the referenced item among a set of potential items (demonstratives), some add a unique identifier to contestant candidates (epithets), and some are themselves unique identifiers (names). Finally, some use situated reference points as indirect deixis for identification (kinship terms). These devices have in common that they seek to draw attention to an exemplar, the ultimate narrow reference, rather than a group of items, and common nouns typically refer to classes of items without inherently referencing single items. Nevertheless, some of these common nouns are considered more “animate” than others and are conceived of as individualized, when occurring in a group of entities (generic plural) or as a narrow set of entities (dual, trial, and paucal plurals). In contrast, for the remaining nouns, distinguishing similar entities is irrelevant (unspecified for plurality). These individualized and narrow referenced items all appear towards the higher end of PH, and we believe this is in demand of an explanation that goes beyond animacy.

## The fourth conundrum: why is individuation more prevalent in higher than lower ranked PH categories?

Individuation is more pertinent with participants higher up in the PH than those in lower rank. Number distinction is one way of expressing individuation. To specify plurality is marked relative to specifying singular status by an additional morpheme (Croft, 1990, p. 89), but often required in high-ranking nominals. Thus, for some nouns within certain languages, the opposition between singular, plural, and sometimes dual, is significant, while for other nouns, it is considered irrelevant. This phenomenon is known as "split plurality" and affects "any of the mechanisms used to mark plurality," including agreement between verb-argument or noun-modifier, direct marking of a noun or noun phrase (Smith-Stark, 1974). Pronouns and nouns referring to animate individuals, including humans, may have a number distinction that is not found with common nouns that designate inanimate things. For example, Tiwi grammar distinguishes between singular and plural of humans, e.g., *wuɹalaka* "young girl" and *wuɹalakawi* "young girls", but not of ants, e.g., *waliwalini* "ant/ants" (Osborne, 1974, p. 52), and in Kharia, separate forms represent cats in the singular and plural, *biloi* "one cat" and *biloiki* "cats", but only one form represent stone/stones: *soreŋ* (Biligiri, 1965, p. 36).

When languages have the possibility of inflecting for number, they display the same kind of graded differences in cut-off points as we observed with animacy and definiteness (see Figure 2). They adhere to a hierarchy where within a certain language, 1st and 2nd persons can be distinguished in number (e.g., be agreement marked for number on the verb), but the same grammatical process does not apply to referents lower in the hierarchy such as the 3rd person. For example, in Georgian (Vogt, 1938), only the 1st and 2nd person pronouns are agreement-marked on the verb. Again, cut-off points are language-specific, and, in Kwakiutl, a reduplicating-distributive plural applies mandatorily only to the 1st person pronouns. In 2nd and 3rd person, by contrast, plural is optional (Boas, 1911, p. 444). Pronoun systems may preferentially mark plurality only in SAP, so that Guaraní distinguishes for plurality and inclusive/exclusive in the 1st and 2nd persons, but a common form haʔé represents singular/plural 3rd person referents (Gregores and Suárez, 1967, p. 141). Accordingly, Croft (1990, pp. 111–112) proposed a hierarchy for the markedness for number relations as in 5):

5) 1st and 2nd person pronoun > 3rd person pronoun > proper names > human common noun > nonhuman animate common noun > inanimate common noun

At the same time, animate categories are left unspecified for number when they appear in object position and seen in relation to its usefulness for humans as game/food (e.g., the Norwegian sentence in 6).

6) Han har skutt mye elg.
He has shot much elk.SG
'He has shot a lot of moose.'

Holding this example against the findings of Kalkstein et al. (2020) that entities in a scene are defined against the presence of humans, it becomes evident that individuation is a feature of prominence, not of whether a clausal participant is animate or not.

In sum, categories higher up in PH tend to be individualized, while categories lower in the PH, are not.

## The fifth conundrum: how can kinship terms, epithets, and personal names refer to humans while ranking above the human category?

Split plurality may adhere to narrower borderlines than animacy, sometimes to humanness, but often to the even narrower class of "kin." In Kpelle, only the subgroup of nouns referencing humans that consists of kinship terms are pluralized (Welmers, 1969, p. 82). Plurality may favor close relationship terms, e.g., *brother-in-law*, *wife* in Tlingit, while inanimates optionally take collective plurals (Swanton, 1968, p. 169). Based on split plurality examples from a variety of languages, Smith-Stark proposed the hierarchy of features controlling split plurality of nouns as in 7):

7) Speaker > addressee > kin > rational > human > animate > inanimate

The positioning of kinship terms between the 2nd and 3rd person pronouns categories, e.g., in Gumbayŋgir, makes sense within proximal level construals. Kin are in a direct personal relationship to the ego, whereas 3rd person referents need not be. In kin term split plurality languages too, the proximity of interpersonal relationships with individual attachments trigger individuation strategies, other nouns do not.

Kinship terms, titles, epithets, and names occupy intermediate sections of PH, but always rank above human nouns. Split ergativity systems may rank kin terms and proper names equally and positioned between personal pronouns and human nouns. In Gumbayŋgir, kinship terms but no other nouns must be marked with accusative case in object position while also being mandatorily marked with ergative case when in subject position (Silverstein, 1976). This behavior applies to titles and epithets as well (Silverstein, 1976: "section names"). An epithet functions as an identification device, separating an individual from similar individuals, attributing specific characteristics to a person, e.g., *Richard the Lion-Hearted*. Its function is reminiscent of definite particles in that it singles out an individual for unique identification. Whenever personal names are ranked, they constitute a class between personal names and other nouns (Helmbrecht et al., 2018), see 8):

8) 1st/2nd person > 3rd person > personal name > human common noun > animate common noun > inanimate common noun

Although cross-linguistic evidence remains scant for the existence of proper names as an independent category in this context, there seems to be an intuitive appeal to names being high in PH. In Romanian, the DOM marker *pe* required for human and specific nouns is also always used with pronouns and proper names (Onea and Mardale, 2020, p. 357). In Arabana, a tripartite split ergative/accusative case-marking system ranks personal names second: SG pronouns > personal names > common nouns. Other tripartite systems collapse personal names with adjacent PH categories: common nouns (Diyari), 3rd person/human nouns with demonstratives but not common nouns (Yidin), or personal pronouns with human/animate common nouns (Manipuri, Dhankute Tamang, and Thakali; Helmbrecht et al., 2018).

While names pattern with categories towards the high end of PH, they also associate with demonstrative and personal pronouns (in other languages besides Yidin). Norwegian has two sets of grammaticalized 3rd person personal pronouns with deictic and specificity semantics preceding personal names (Johannessen, 2008). One set are the psychological proximal demonstratives (PPD) labelled "preproprial articles" in Johannessen (2008, p. 170). These introduce personal names/kin terms of close family relations, e.g., *mother*, *father*. PPD agree with the semantic gender of the referent, e.g., *a/hu/ho Gerd* "she Gerd", *a/hu/ho mor* "she mother", *n/han far* "he father", and are used of people that are personally known to SAP. The other set constitutes demonstrative markers which are identical to the set of 3rd person singular personal pronouns, e.g., *hu/hun Berit* "she Berit" (feminine), and *han Ola* "he Ola" (masculine). These demonstratives, labelled "psychological distal demonstratives" (PDD), are used with specific reference to human or human-like nouns (e.g., pets). PPD and PDD are phonologically, semantically, pragmatically, and syntactically distinct, with a higher degree of grammaticalization in PPD and usage reflecting the notion of psychological proximity vs. distance: while the proximal set is used with first names and family relations, the distal set is used with all kinds of nouns denoting humans. In Johannessen's analysis, PDD signals social distance from speaker and/or addressee towards the referent, connoting the speaker's negative attitudes for this individual and excluding the possibility of personal acquaintance, while simultaneously indicating that the referent is a specific individual. The use of PDD is categorically conditioned by its specificity; Johannessen (2008, p. 167) states that "The PDD must refer to something specific, never hypothetical or non-existing." Specificity distinguishes them from regular definite articles, while also interacting with definiteness in interesting ways. This description is in line with the rank of names and demonstratives in the intermediate section of PH: not as proximal as SAP, but still more proximal than the low right end of PH.

A related case is the German usage of definite articles with personal names (*die Gisela,* feminine; *der Jonas*, masculine; Patterson and Schumacher, 2021). A case from noun classes corroborates that personal relationships and individuation are closely connected concepts: while the Setswana noun class 1/2 is reserved for only and most personal nouns, class 1a/2a comprises all personal names, kinship terms and personified animals (Tsonope, 1988, p. 40). The former class represents the Ego, the latter the Ego's immediate environment.

The reason why kinship terms and uniquely identifying labels like names, epithets, titles, and psychological demonstrative pronouns rank above human nouns when they indeed reference humans, cannot be explained by pointing to "animacy." Nor has it been explained why these terms associate closely with definiteness, specificity, and demonstratives.

## The sixth conundrum: if animacy underlies prominence, how can definiteness and specificity be independent drivers of prominence hierarchies?

Surprisingly, whereas PH in some languages respond to animacy alone, other languages are sensitive exclusively to definiteness or specificity (see Figure 2). This means that not just animacy and humanness, but definiteness and specificity as well can independently drive DOM effects. In person-sensitive verbal agreement languages, definiteness plays an essential part; e.g., in Hungarian, agreement with subject is marked on the verb, but objects are only agreement marked if they are definite (Kiss, 2017). If PH is truly based on animacy, aspects like definiteness and specificity are left unexplained. We believe the answer to this could be concreteness and detail associated with proximal construal as it is documented in CLT. Grammatical behavior itself bears witness of this; individuation is a more prevalent feature in pronouns than in nouns and clusters with definiteness and number. For example, number distinctions can be present on pronouns, including 3rd person, that are not present with common nouns, as in Mandarin: Pronouns for 3rd person are differentiated for plurality, e.g., 3rd singular *tā* "he/she/it" vs. 3rd plural, *tāmen* "they", but not for common nouns, e.g., *shū* "book/books" (Croft, 1990, p. 111; Li and Thompson, 1981, p. 13).

We believe the answer to conundrums 4–6 is that these are instance of individualization and narrow reference, which matches the predictions and experimental findings of CLT that, in proximal construal, people visualize concrete exemplars and construct narrower categories with fewer members than in distal level construal (see Table 1).

# Fronting mechanisms

Various linguistic devices share a formal marking with entities high in PH but has not been linked to these hierarchies in linguistics. Fronting mechanisms refer to various syntactic processes that move an element from its typical position to the beginning of a sentence or clause for various cognitive and communicative purposes, such as emphasis and focus, topic marking, and creating contrast or cohesion in discourse. A shared function of fronting may be to make certain elements more cognitively accessible or easier to process by placing them in a prominent position, potentially reducing cognitive load and facilitating attention.

## The seventh conundrum: why does one "empathize with" inanimate discourse topics but not with inanimate direct objects?

The "Discourse Topic Empathy Hierarchy" was proposed in Kuno and Kaburaki (1977) and Kuno (1976, p. 267): "It is easier for the speaker to empathize with an object (e.g., a person) which he has been talking about than with an object that he has introduced anew into the discourse: Discourse-Anaphoric > Discourse-Nonanaphoric." Similarly, in attempting to bridge topic-worthiness and agentivity, Payne (1997, p. 151) invokes the concept of empathy (see also Kuno, 1976) to explain why certain topics are favored in human conversation, see 9):

9) "Humans tend to select as topics entities with whom they empathize, first of all themselves, then the person they are speaking to, then other human beings, and finally the inanimate world. Therefore, morphosyntactic expressions whose function is to refer to topical entities indirectly tend to refer to entities that speakers empathize with."

In direct/inverse systems, the speaker's assumed degree of linguistic empathy (Kuno, 1976, 1987; Oshima, 2007a, 2007b) follows a set of constraints: (1) Speech act: the speaker cannot empathize with someone else more than with himself, (2) Topic: the speaker "empathizes" more easily with the discourse topic than non-topic matters, (3) Surface structure: It is easier for the speaker to "empathize with" the referent of the grammatical subject than with referents of other NPs in the sentence, (4) Descriptors: Given descriptor *x* (e.g., Peter) and another descriptor *f(x)* (e.g., Peter's sister), it is easier to empathize with *x* than with *f(x)*, which is indirectly accessed via the first descriptor. To these generalizations is added a caveat that a single sentence cannot contain logical conflicts in empathy relationships and empathy relationships within a sentence must be consistent (Oshima, 2007a).

The logic of the empathy hierarchy is that proximal participants are considered central to the story and thus higher on an empathy hierarchy than the obviative marked participant within the same sentence (Oshima, 2007a; Kuno and Kaburaki, 1977). However, taking a closer look at these principles, one observes that, on the one hand, inanimate objects are lowest ranking given that one cannot empathize with unliving things, and yet on the other hand, when a non-living entity becomes the topic, it outranks a living, also human, subject, in that respect. Linguistic empathy was recently reiterated as the speaker's attitude towards and identification with the referents therein (Kann et al., 2023), a concept resembling the social dimension of psychological distance, but which is unintuitive of fronted inanimate objects, and lacking the potentials to be extended into a full model that can explain all aspects of PH, in particular why languages use the same grammatical devices for discourse topicality of inanimate objects as for emotion-based proximity.

Several languages express both types by a unique identifier, e.g., direct/inverse morphology, particles, or case, suggesting the underlying mechanism is basically the same, *viz.* to hold in mind (or draw close) an entity for proximal construal, evident from the following examples. First, Navajo adheres to a strict agency hierarchy (Zúñiga, 2014; Witherspoon, 1977; Drexel, 2014). Deviating from the hierarchy requires the use of the inverse marker *bi-* on the verb when two participants are equal on the PH. While *bi-* in Navajo is a device that marks reversed agency order (DeLancey, 1981), if attached to the postpositional phrase instead of to the verb, it functions as a topicalizing device (Perkins, 1978). Second, the *wa* particle in Japanese is used as a "pragmatic case marker" for participants in the clause that possess the highest degree of inherent topicality, but also for the attention-evoking function of contrast (Maruyama, 2003), and as a topic-marker and a marker for something that is known from previous conversation (Drexel, 2014; Kittilä et al., 2011; Kuno, 1973). Third, while the primary function of ergative case is to mark subjects in transitive sentences (Dixon, 1979), and secondarily to signal agency of inanimate or lower ranked items in DOM languages, the ergative also serve other functions that align with fronting mechanisms in optional ergative systems (Fauconnier, 2011). In the Gooniyandi and Warwa languages, an unexpected discourse topic that is drawn to attention receives ergative marking (McGregor, 1992, 1998, 2006). In Foré, ergative case signals contrastive focus (Donohue and Donohue, 1998, p. 85), and in Umpithamu (Verstraete, 2010) and Waskia (Ross and Paol, 1978, pp. 36–39), ergative simply represents focus. Besides, in Gooniyandi, ergative in commands signifies that a 2nd person agent is accorded special focus or prominence because of surprising or unusual involvement in a process. By contrast, absence of ergative marking on such agents signifies that the addressee agent is obscured or not individuated, because in an avoidance relationship to the speaker (McGregor, 1992). These cases demonstrate that ergative case is not just about agency, but that instead it generally signals a proximal-level construal.

Instead of interpreting such devices as signals of animacy or topicality in the traditional sense, they can be viewed as attention shifters from a default egocentric viewpoint to a reference point—a topic—intently chosen by the speaker in the here-and-now moment. In the CLT framework this would amount to the same thing, since a topic is what is being talked about. In other words, topicalization is to bring an entity into the speech-act participants' immediate sphere of attention. Topicality infers that the referent of a noun phrase is identifiable by the hearer, creating a sense of "aboutness" (Reinhart, 1981, p. 5) and implicitly shared knowledge (Gundel, 1985, p. 92): "the topic of a speech act will normally be some entity that is already familiar to both speaker and addressee." Topics are time-sensitive, "what is of current interest or concern" (Strawson, 1964, p. 104); consequently, a topical referent is something that at a certain moment in time presides SAP's attention and as a result, is attributed salience, or cognitive prominence. Kaiser (2009, p. 335) links salience to attention, stating that anaphoric pronouns correlate with the most salient entities in discourse, in the sense that these referents are at the center of attention.

An alternative to postulating a caveat to grammatical patterns only applicable to linguistics is to invoke a known cognitive constraint, the scarcity of human attention (Simon, 1971), or the fact that attention can only be fully directed to one target at a time. Scarcity of attention interacts with linguistic information structure in important ways. Both provide a narrow and selective focus (Walker, 2002). Just as attention is selective due to its limited capacity, topics in linguistic structure serve to focus the listener to highlighted information in discourse, helping to direct limited attentional resources efficiently and playing a crucial role in organizing discourse structure (Walker, 2002). In discourse, shifts in topics can be alikened to the reallocation of attentional focus when switching tasks; the linguistic cues signal topic changes to help redirect the listener's limited attentional resources. Thus, attention necessitates prioritization of information, just as topics in linguistic structure indicate what information is deemed most relevant in a given context. Finally, topic constructions align with the constraints of working memory, presenting information in chunks that can be readily processed and maintained in an active state (Oberauer, 2019; Schmidt and Schmidt, 2001).

Topicality is thus a derivative of attention and how this is implemented in grammar finds parallels in human biology. Attention constantly relocates with moving cognizers within their peripersonal space or the individual's action radius that is neurally encoded in the human brain's "body schema", linked to hands, head, and trunk (Di Pellegrino and Làdavas, 2015). It can be reallocated to items by having human eyes and minds directed at it at will. In linguistics, the first type corresponds to inherent "topic-worthy," and the latter to "context-imparted" topicality (Payne, 1997). In the unmarked case, an entity higher in PH is preferred as the topic as a function of inherent topicality, but this can be reversed if a low-ranking referent is highlighted as a topic for communicative purposes, requiring linguistic cues that are overtly marked. The answers to an apparent enigmatic multifunctionality of inherent vs. context-imparted topicality lies in the level of construal of topics. Within a CLT model of PH, fronting mechanisms are understood in conjunction with the roles of SAP. Topics are marked as proximal in grammars because they inhabit the minds of speech-act participants and the mental space between them in a here-and-now moment. What part of a linguistic construction is higher on PH, is determined either by situatedness and context when drawn mentally close by joint attention. Also, determined by the omnipresence of conscious cognizers aware of their own animate capacities as potential initiators of action chains, but never at the same time due to limitations of human attention. Note also that the SAP pronouns "I" and "you" shift their reference with who speaks and is who spoken to, placing them firmly within context-dependency.

We suggest that linguistic empathy, topicalization in direct/inverse systems, agency and animacy rankings are all aspects of proximal level construal, the egocentric viewpoint in the here and now. By including in the egocentric reference point not just of the self but also the "here and now", as proposed in CLT, some formerly misconceived phenomena of prominence hierarchies can make sense; since topics are by definition something that a speaker holds in his or her attention, at the moment of utterance. Empathy emanates from the self *and* results from attention, the act of applying the mind to something in the present moment, and that can potentially evoke feelings towards what is presently attended to, as a secondary function of the immediate attention.

### The eighth conundrum: why can oblique DOM participants be passivized when their homophone obliques cannot?

Split ergative languages employ ergative for nontypical subjects and accusative for nontypical objects. In so-called oblique DOM languages, however, direct objects with specifications characteristic of higher prominence such as being human, animate, definite, or specific, are instead marked with *oblique* prepositions or cases: dative, genitive, or locative. Elsewhere in the language these oblique cases are used in their primary functions of beneficiary, part/owner of, or place/goal.

In discourse, passive voice can help maintain cohesion by keeping the focus on a consistent topic by highlighting the patient role and suppressing the agent. In nominative-accusative languages, the accusative marked patient object in an active sentence, typically low in animacy and prominence, is raised to subject position, attaining all the typical characteristics of subjects. In oblique DOM languages, however, only direct objects can be fronted to fill the subject position in passive constructions, while other arguments marked with these cases within a certain language, cannot (Irimia, 2023). For example, Spanish marks a human direct object with a dative preposition *a*, while inanimate direct objects are unmarked. While the dative DOM can convert to a subject in a passive sentence, the same operation for an ordinary indirect object, with the exact same dative case marking, is ungrammatical, see 10) and 11); examples from Irimia (2023).

10) Veo **a** la mujer/(*a)
See.1SG DAT=DOM DEF.F.SG woman/DAT=DOM

la casa.
DEF.F.SG. house
'I see the woman/the house.'

La mujer/la casa fue vista
DEF.F.SG woman/DEF.F.SG house was seen.F.SG
'The woman/the house was seen.'

11) Le doy el libro
CL.3SG.DAT give.1SG DEF.M.SG book

**a** la mujer.
DAT DEF.F.SG woman
'I give the book to the woman.'

*La mujer fue dada/dado el libro.
DEF.F.SG woman was given.F.SG./M.SG DEF.M.SG book
'The woman was given the book.'

The prevailing understanding for this difference in grammatical behavior is that "Oblique DOM is not an oblique syntactically" (Irimia, 2023). However, to analyze this purely in syntactic terms masks the choice of an oblique case for DOM, as well as the connection to other usages of that case within the language. Instead, semantic extension leading to polysemy of these oblique cases to DOM may reveal that case assignment was not made at random. For example, the DOM marker *pe* in Romanian likely evolved from embedded topics and were polysemous with the goal/locative/topical usages *for*, *on*, *concerning*, and *about* in Old Romanian (Onea and Mardale, 2020, p. 362). These usages are lost in modern Romanian and instead developed the values "human" and "definite" required for DOM. The authors point out that "different languages tend to exhibit the same or similar patterns." Alternatively, to focus on the recipient role in, e.g., *Am cumpărat flori* pe *mama*. "I bought flowers for my mother" would highlight how Romanian overlaps with Spanish DOM.

Thus, in response to this conundrum, we suggest that oblique DOM, as part of cross-linguistic PH phenomena, cannot be explained by syntax alone. Instead, it should be viewed as one of several fronting mechanisms that exploit the semantics of case to associate the characteristics of proximal construal.

## Abstraction

### The ninth conundrum: why are abstract concepts low in prominence?

In many languages, a broader PH is observed that includes abstract concepts at the lower end of the scale. An expanded hierarchy might look like in 12) (Langacker, 1991a, p. 307):

12) Speaker > hearer > human > animal > plants > physical object > abstract entity

For example, abstract concepts like *hunger*, *health*, or *happiness* are the lowest in prominence in Navajo (Drexel, 2014, p. 9). Blackfoot also makes an interesting case where nouns are classified as animate vs. inanimate/abstract (Ritter, 2014). An intriguing case is Old Romanian where the oblique DOM marker *pe/pre* mandatorily marks proper nouns, which are highly individuated nouns that call for proximal construal, but with the only two recorded exceptions being names of *Hristos* "Christ" and *Dumnezeu* "God" (Onea and Mardale, 2020, p. 359); a likely explanation being that deities were conceived of abstract concepts.

While proximal construal associates with individuation, abstraction associates with group thinking. Bantu noun classes tend to classify abstract concepts with collective nouns, e.g., Setswana

(Tsonope, 1988, p. 40). In many nominal classifier systems, insects are classified as inanimate: Indonesian (*kaki*), Lahu (*mà*), Vietnamese (*cai*), or as abstract: Fulfulde (*ngu*) (Lobben et al., 2020, supplement). The gradience here suggests that absence of individuation better explains these collocations than animacy.

To sculpt abstract concepts into a grid made for concrete objects might not be very informative of their constitution or explain why they are the lowest category in PHs. Curiously, very little research has gone into possible reasons for abstract concepts being low in prominence. Generally, the animacy account has been assumed but never substantiated with abstract concepts. For example, Navajo abstract concepts are labelled "incorporeal inanimates", contrasting with "corporeal inanimates" (Witherspoon, 1977; Perkins, 1978).

In a model of PH within CLT, there is no need for *ad hoc* assumptions of inanimacy in abstract concepts. Instead, abstraction inhabits the far end of hierarchies as a function of psychological distance. Soderberg et al. (2015) review the theoretical rationale for expecting a link between psychological distance and abstraction and provide multiple experiments testing this link. The effect of distance was significant, produced medium-sized effects on construal-level, and was similar across different types of psychological distance: temporal, spatial, social, and hypothetical, supporting CLT's central prediction that variation along any dimension of psychological distance will influence level of abstraction.

## Cultural flexibility and variance

### The tenth conundrum: why does linguistic categorization not comply with biology?

Animacy-sensitive DOM may respond to the human category while extending beyond, or retracting within, the human category "in ways that are clearly culturally determined" (Aissen, 2003, p. 456). For example, Yiddish marks direct objects differently depending on humanness, but mandatory case marking is restricted to a set of masculine nouns that denote humans culturally defined as "worthy of respect": *grandfather*, *teacher*, or an ethnic group member. Older feminine relatives are only optionally case-marked; e.g., *grandmother*, *mother, aunt*. Splitting of the human group can result in categorizing people as non-human. The Marind of Papua New Guinea assign women to animal classes while men inhabit the human class. While such examples take a narrower view to what should be included within the egocentric sphere, there are also examples of the opposite: Ritharngu, an Australian language, extends case-marked objects beyond the human category to include so-called "higher animals" (e.g., kangaroos, dogs, and emus). Animals considered lower on the hierarchy, like fish and raccoons, are not case-marked (Heath, 1980). Moreover, case-marking may leak across the animate-inanimate boundary; in Bayungo, all animate-referring objects are case-marked, including humans, but in addition the two inanimate nouns of *meat* and *vegetable food* (Austin, 1981). This narrow selection of inanimate, case-marked, DOM nouns is unlikely to be random as these nouns reference biological material for consumption and therefore associate with human bodies.

Divisions into "higher" vs. "lower" animals can affect individuation as well. In Tiwi, higher animals are number-marked like humans, while lower animals and inanimates are unspecified for number (Haspelmath, 2013). Further, in Manam, dual and paucal noun forms are used only for humans and "higher animal," in line with the typological tendency that unmarked grammatical categories display more values than marked ones (Croft, 1990, p. 78). What counts as a higher animal in DOM, however, may come down to whether the animal is domestic or wild, discounting the actual species. Humans are always considered "higher animals," but pigs, dogs, fowl, and goats only when domesticated (Lichtenberk, 1983, pp. 110, 256). The same kind of ranking between domestic and wild animals is found in Navajo (Drexel, 2014; Perkins, 1978). Although it seems to matter how humans interact with animals, taxa matter towards extremes in natural size and individuation feasibility; the Navajo PH ranks insects lower than "small animals" but higher than natural forces > plants/inanimate objects (Lockwood and Macaulay, 2012).

Divisions of biological sex are largely clear-cut but, in Lokono, grammatical gender seems to follow in-group versus out-group distinctions reminiscent of the PH egocentric systems. Masculine gender is applied to all men within the Lokono tribe, unless they are despised, as well as all things that one considers positive, even animals, things, and spirits (if they are thought to be good or are protagonists within a story). Men of other tribes can be referred to with masculine gender if they are friends of the speaker or in a mutual respectful relationship to him. Feminine gender, by contrast, apart from being used of women within the tribe, are used for despicable men, and men from other tribes (Aikhenvald, 2000, p. 262), see Figure 5. This pans out even more clearly in the Brazilian language Jarawara, where a masculine nominal class can be used selectively to revere women. Conversely, in Amharic culture, feminine gender is applied to men as a marker of respect. Evidently, the prototype from which psychological distance is measured is male in the former two cases, and female in the latter. Apparently, the differences in cut-off points along prominence scales are grounded in construals defined by the culturally dominant supplier of premises and social values that speech societies make within each culture.

Psychological distance can explain these mismatches between culture and biology in linguistic categories, e.g., why some human classifiers do not align with biological distinctions of humans like male and female, why external ethnicities are sometimes excluded from human classification, why domestic animals are ranked higher in PH than wild animals within the same species, and why grammars treat pets different from other animals. Just as the semantic scales in PH and noun categorization are gradient, the conception of psychological distance is continuous with no fixed or predefined cut-off points.

Yet, the clearest evidence of psychological distance impetus comes from cognitive neuroscience. In human brains, ventral visual areas are topographically organized along a gradient scale analogue to linguistic animacy scales; the neural population that respond to face perception is located adjacent to that of primates, birds occupy an area in between the primate and the insect areas, which again border on the regions for inanimate entities (Connolly et al., 2012). A similar topography exists in the monkey (*makaka mulatta*) brain; for example, within face responding neurons, distinct clusters activated primate and non-primate (goat/horse/dog) faces, and primate faces activated separate clusters for humans and monkeys (Kiani et al., 2007). In other words, primates may be innately endowed with a neural grid of animacy gradience and, in humans, this gets expressed in most languages' structures. Knowing this, one cannot disregard the parallel between the neuroscientific finding that brain regions responding to dogs' faces were closer to human faces in the human brain's representational space than, for example, the less familiar but genealogically closer monkeys (Connolly et al., 2011). Moreover, the grammatical status of pets and domestic animals overlaps with that of humans. Since monkeys do not keep dogs as pets, their brains may express the "original" topography.

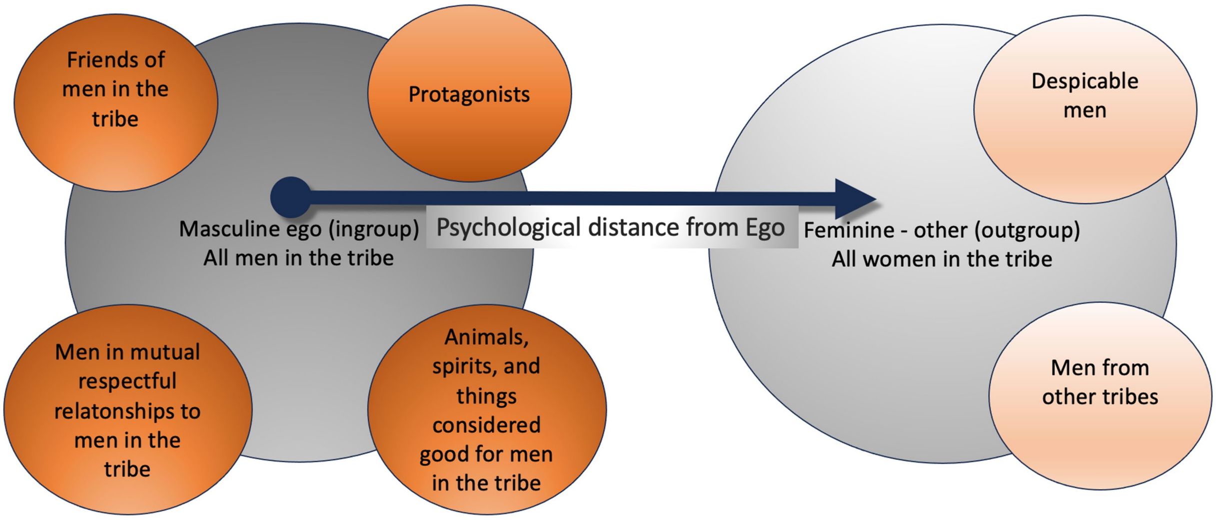


FIGURE 5
Prototypical category structure in Lokono explained in terms of psychological distance. The value in the large circle is the prototype, and the smaller circles represent the semantic extensions, defined from the viewpoint of a culturally dominant ego.

PH and psychological distance are both *malleable* concepts, as evidenced by cross-linguistic variations in grammatical structures. However, it is important to remember that it is the brain's neural network that possesses this malleability or plasticity. Evidently, proximal construal overrules genealogical and phylogenetic information and may produce lasting effects on the brain, and in turn, on grammar. By consequence, if there is evidence that peoples' relationships to their pets can affect brain organization, there are strong reasons to believe that men's culturally determined relationships to women, and ingroups' relationships to outgroups, also affect their brains, and in turn, their grammars. All this makes psychological distance a very real-world phenomenon, embodied in the human brain.

## Merge: where linguistics meets psychology

### Cross-field correspondences in construal levels

As discussed in the preceding, the common ground between linguistic prominence and construal levels resulting from psychological distance are multifold. The most conspicuous overlap occurs with the egocentric viewpoint. In taking an integrative approach to linguistics, all the properties associated with the Ego in its own understanding of itself becomes available to characterize the proximal viewpoint. Agency, animacy, empathy become part of a linguistic expression's prominence as manifestations of the most proximal construal level in any linguistic proposition, whether a 1st or 3rd person pronoun, or an animate noun. At close range, the Ego observes details, individuals, and is at the mercy of the immediate context. Topics are part of this context because they are at the center of current attention in speech-act participants' minds. When immersed in immediate context, people tend to think about exemplars rather than categories of things; hence, items are definite and specific, and at the social level, personal relationships are one-to-one. This results in narrow categories with either just one or a limited number of referents. In prominence hierarchies, these narrower categories manifest as ranked personal pronouns, personal names, titles, demonstrative personal pronouns, and kinship terms. Although kinship terms are words with generalized meanings, when in use they reference one individual at a time. CLT is resourceful to linguistics in throwing new light on fronting mechanisms. The use of common grammatical marking in several apparently disparate grammatical functions such as topicality, focus, and agency suggests conceptual unity. By incorporating linguistic prominence under CLT's notion of proximal level construal, these devices will receive a unitary account. As a consequence, the formerly proposed subhierarchies of PH in 1a-d) above become superfluous and can be replaced by *one* hierarchy. Table 2 summarizes the parallels between PH and CLT at the proximal construal level.

**TABLE 2 Parallel structural characteristics of proximal construal in CLT and PH.**

| Proximal construal level characteristics in psychology | Proximal features in prominence hierarchies in languages |
|---|---|
| Concrete | Concrete |
| First person perspective | 1st person pronoun, may include both SAP |
| Social ingroup, emotionally close/ positive relations (e.g., friends) | Kin, the other SAP, tribal ingroup, dominant gender; domestic animals and pets |
| Exemplars or narrow categories | Singular pronouns, kinship terms, personal names |
| Individualized information, details | Personal names; dual, paucal and plural forms |
| Specific behaviors | Specific/definite referents, including pronouns |
| Contextualized, situation-specific demands, situational information | Context-imparted topics; pronominal referents vary with context (who speaks, textual context) |
| Goal irrelevant | Grammatical subject, intransitive actions |
| Feeling of elevated power | Strength of agency |

The CLT variables are sampled from Soderberg et al. (2015) and Trope and Liberman (2003).

In thought and language, the third person perspective brings about different effects on cognition and grammar than first person perspective. While PH prioritize first over third person, considering them more prototypical agents than third person participants, in thought, first-person perspective causes more detailed recalls about events, while third person perspectives bring about more global, dispositional, less detailed, and abstract qualities of the same events. Social distance towards outgroups is related to negative emotions in psychology as well as in language, the latter conspicuous in nominal classifiers. In psychology, social outgroups are associated with structured, predictable, and abstract features, and more homogenous than ingroups, while in language, outgroups are classified with categories further to the abstract, lower end of PH.

In CLT and PH, as the psychological distance from Ego increases, categories become wider and more general, although heterogenous in real life. The divisions along animal species and heterogeneity of inanimate objects are disregarded, unless drawn to attention and supplied with the proximal construal tokens of individuation and concreteness. As the scope widens, categories become more inclusive but also abstract, in line with the functioning of taxonomic systems: the level of abstraction and complexity of features are inversely correlated; superordinate concepts contain more referents but are captured in terms of fewer features, and vice versa (Rosch and Lloyd, 1978). Thus, the most superordinate concepts are also the most abstract. PH in linguistics differ, however, in including in abstract concepts that are intangible (deities, emotions, and sensations), however, this may be a matter of semantic extension in linguistic systems.

Table 3 sums up the properties of distal level construal in language and psychology.

## The mental timeline in transitive trajectories

Temporal distance in transitive sentences is the time elapsed from when an agent initiates an action to the impact this has on the direct object. We speak here of mental timelines (Arzy et al., 2009; Oliveri et al., 2009; Bonato et al., 2012; Corballis, 2009). In mental representational space, time can be stretched or compressed; e.g., verbs can reference processes or punctual events, which have vast effects on grammar (Vendler, 1957). The idea of traversing temporal distance is not new to linguistics. Cognitive grammar imagines the trajectory in transitive sentences to be a mental timeline (e.g., Langacker, 1987, pp. 402–403), and cognitive grammar linguists analyzed spatial and temporal expressions with reference to mental paths (Langacker, 2017, p. 178), and conceived time (Langacker, 1991b, p. 150). Conceptual archetypes and image schemas describe events as trajectory-landmark relationships. In a transitive sentence, the subject (typically an agent) is conceptualized as the (moving) trajectory, while the object is the (stationary) landmark. The interaction between these two entities is mapped onto the mental timeline. Recently, temporospatial construals of events were corroborated by the theory of cognitive spacetime (Stocker, 2014), supported by linguistic analysis and substantial experimental evidence, e.g., from eye tracking experiments (Demarais and Cohen, 1998). Thus, while more research could be useful specifically with regard to the PH and spatiotemporal dimensions of psychological distance, there is already ample independent support for the idea that transitivity is mentally represented along temporospatial domains.

TABLE 3 Parallel structural characteristics of distal construal in CLT and PH.

| Distal construal level characteristics in psychology | Distal features in prominence hierarchies in languages |
|---|---|
| Abstract | Abstract |
| Third person perspective | 3rd person pronouns |
| Social outgroup; emotionally distant/negative relations (e.g., enemy) | Other tribes and genders; wild animals and game |
| Broad categories; broad traits | Superordinate categories, e.g., 'animals', 'inanimate objects'; lack of individuation, e.g., insects = abstract |
| Superordinate; aggregate and dispositional information, primary features | Superordinate 'animal' and 'inanimate objects'; fewer and only essential semantic features of superordinate categories |
| Decontextualized | Word reference is more independent of situated context (contrast pronouns which have reference variable with context to nouns) |
| Goal relevant | Direct objects are targets in transitive actions |
| Overarching goals | Inanimate objects are goals in transitive actions |

The CLT variables are sampled from Soderberg et al. (2015) and Trope and Liberman (2003).

## Discussion and conclusion

We aimed to provide a deeper understanding of prominence in hierarchies by taking as starting point the prototypical transitive sentence, defined as typical human-to-inanimate object interactions. We then explained markedness phenomena as deviations from prototypes, stressing the fact that prototypicality as well as markedness reversal are commonly motivated by extralinguistic factors. We identified the transitive sentence with a minimum of two participants as the cognitive unit where prominence occurs, taking into consideration that, while prominence operates within syntagmatic relations (i.e., within a sentence), markedness happens within paradigmatic relations (i.e., between sentences). Deviations from the prototypical sentence arise because other clausal participants compete with the subject for prominence; e.g., topics or human/animate direct objects that are deemed psychologically more proximal to a cognizing Ego. We identified the tokens for prominence as individuation and concreteness, measured in psychological distance from the vantage point of this Ego. The same kind of distance explains lack of individuation in prototypical direct objects.

What sets our analysis apart from other prominence analyses is that lower-ranked categories are accounted for, including abstract categories of nouns at the lowest end of PH. Direct objects are prototypically wider and more general, less individuated, and increasingly abstract *because* they are construed at psychologically distal levels. In our analysis, all fronting mechanisms are grammars' devices to change a distal level entity to proximal level construal. We see the signs of this in how some grammars treat inverse constructions, topics and focus alike. While the points above constitute our main findings, we relied heavily on the empirical and theoretical work carried out for decades within CLT that temporal, spatial, social, and factual distance is in fact perceived and conceived of in the human mind as *one* measure of distance. This framework allowed us to treat the transitive space–time trajectory emanating from the agent as forged with social distance. While former linguistic analyses have hypothesized that emotional distance might govern prominence constructions, the unification of all these dimensions of psychological distance was in fact the crucial aspect. It scaffolded the present proposal that transitivity itself imposes a temporospatial distance concurrent with the social distance. A serendipitous finding was that gradience in differentially marked objects matched gradience in semantic category structure in noun categorization. By applying the Ego vs. Other psychological distance, we could identify motivations for biology/linguistic category mismatches, and explain, e.g., that categorizing insects as "inanimate" is likely due to lack of individuation rather than misconceptions that insects are not "living things."

## Prior research

Our proposal differs from former proposals in several respects. Linguists have tried to resolve the puzzle of differential object marking in various ways, including asking "why and how DOM would arise in the first place" (Onea and Mardale, 2020). While recognizing the challenges in finding conclusive answers, we notice that these previously offered solutions remain within linguistic structure and behavior. For example, the idea that DOM arises because "the animacy and referentiality scales are good indicators of prototypical subjects," or "are indexed for their properties on general notions such as *affectedness*, *transitivity* or *animacy*" (Onea and Mardale, 2020, p. 352). In contrast, we have ventured the inclusion of linguistic prominence within general cognition, which appears to be a radically novel way of thinking about prominence. Nevertheless, we made use of previous analyses along this path since many were integral to and compatible with our proposal. In particular, the notions of egocentricity (Gardelle and Sorlin, 2018), viewpoint and attention flow (DeLancey, 1981), the empathy hierarchy (Kuno and Kaburaki, 1977; DeLancey, 1981; Langacker, 1991a; Oshima, 2007a; Matthews, 2007), counting in the analysis of empathy as a radial schema with the self in the center (Yamamoto, 1999), and finally the innumerable accounts on agency and animacy, often seen as inseparable concepts (Yamamoto, 2006, p. 29).

Where our account radically differs from some other proposals is in their practice and belief that languages are in essence autonomous systems that cannot only be described but also explained by mechanisms exclusive to language (e.g., "modularity of mind," see Fodor, 1985 and "parallel architecture," see Jackendoff and Audring, 2019). In our view, instead of being endowed with a special, isolated language faculty, humans possess a language faculty that is deeply intertwined with and building on general cognitive processes, making it an integral part of human cognition. Indeed, "separatists" need to explain the remarkable parallels between prominence phenomena and proximal-distal construal. Much previous meticulous work by linguists (e.g., Aissen, 2003; Irimia, 2023; Starke, 2017; Caha, 2009) may not suffice to fully explain prominence phenomena, as these formalisms may stay at descriptive levels. Thus, our goal is not restricted to *how* languages may be constructed but extends to *why* state of affairs is the way they are.

Our proposal also differs from others in assigning a pivotal role to cultural differences regarding the makeup of linguistic categories. Cultural biases in prominence hierarchies confirms the egocentric perspective. Contrary to Gardelle and Sorlin (2018, p. 134) that "despite an obviously cultural basis, the notion of Animacy Hierarchy appears to be restricted to linguistics; it does not seem to be used, for instance, in sociology, anthropology or philosophy," we have shown instead that this perspective is indeed what influences the egocentric viewpoint in PH. Given the role of empathy and egocentric perspective in the prominence hierarchy, it is not surprising that cultural differences in social organization appear to influence the grammatical systems, culture here being understood as "collective patterns of thought within a speech society" (Enfield, 2002).

The way in which language systems can co-vary with the cultural belief systems of their speakers has been discussed since the appearance of linguistic relativism via the descriptions offered in Enfield (2002). The idea that cognitive systems may underlie linguistic systems in a culture-specific way, was first elaborated by Wierzbicka (1979), who coined the term "ethnosyntax," suggesting that "every language embodies in its very structure a certain world view." Importantly, she went a step further from Sapir-Whorfism by stating that language does more than to code for culture-based semantic cultural content, but code culture-based grammars as well. This approach has been deemed relevant to aspects of the animacy hierarchy (Drexel, 2014). Specifically, ethnosyntax refers to "the study of connections between the cultural knowledge, attitudes, and practices of speakers, and the morphosyntactic resources that employ in speech" (Enfield, 2002, p. 4). Indeed, some of linguists' challenges "that beset grammatical theory derive from trying to analyze native speakers' linguistic knowledge as a self-contained system"; as Keesing (1979) pointed out after years of fieldwork on the Solomon Islands with the Mailaita language.

Although linguists have hinted at cultural explanations for linguistic behaviors described as "cultural flexibility" above, no explanation has been offered that is valid cross-linguistically. For example, Aissen (2003, p. 457) suggests that there are two ways to analyze these cases; either are the categories HUMAN, ANIMATE, and INANIMATE understood differently in particular languages, or there is further language-particular ranking within these ontological categories. Neither of these proposals offer a real explanation beyond the descriptive level. Crucially, none of them provided a comprehensive or *unitary* understanding to all aspects of prominence with reference to universal principles of human cognition, while at the same time allowing for considerable cultural flexibility. Finally, we missed an account that went beyond the mere descriptive level to elaborate on *why* prominence hierarchies look the way they do, how they arise and are maintained in grammars.

## Predictions

Several potential predictions followed from integrating PH and nominal categorization within CLT. One was that as more languages are investigated, prominence and markedness phenomena already described within linguistics will lend themselves to the characteristics of proximal vs. distal level construals, marked with the set of resources for marking prominence existing in the individual language. Perhaps more important was the prediction that prominence in language are concurrent with peoples' perceptions of prominence outside of language, including their cultural experience and social organization. Another bold prediction was that cross-linguistic variation in PH cut-off points respond to cultural beliefs and practices, or even that within individual languages that possess both type of phenomena, semantic classification in nominal classifier systems and PH will concur. All this can be tested further and should set the ground for a whole new paradigm of research.

## Future research

By aligning linguistic structure with the concept of construal levels modified by psychological distance, we hope has opened up a whole new field for potential new research. One possible direction for future research regards the assumed correspondence between culture and social distance on the one hand, and linguistic structure on the other. This is relevant for nominal categorization as well as prominence hierarchies. As for the question of cultural variance in noun classification, we indicated some "unexpected memberships" that contradicted biological facts and instead followed preferences along gender and ethnicity. We suppose that these are cultural or ethnocentric, defined subconsciously by the dominant social group within societies, although anthropological studies will need to confirm this connection.

Another proposal for future research is to explore the variance in cut-off points along the PH with respect to split case systems. One should note that differences could be motivated by culture or by general cognitive processes. Research on diachronic change shows that the evolutionary paths of DOM may proceed as gradual spreading from the left to the right on animacy and referentiality scales (Onea and Mardale, 2020, p. 351). For example, Lichtenberk (1983, p. 133) predicted that DOM in Manam presently is used for humans/higher animals but is "moving toward a stage where *-di* will be the predominant, if not the only, 3pl object marker." If this is indeed the case in many languages, the arbitrary cut-off points represent could reflect general tendencies in historical development. The overall variation in categories involved in prominence hierarchies speaks against this being the only cause, however; more likely the causes could be mixed.

Finally, there is a dire need for more research on how abstract concepts behave within PH, including the description of multiple typically varied languages, and a semantic characterization of such abstract concepts.

## Conclusion

Psychological distance is able to explain categorization in aspects of noun classification as well as in PH. Thus, we have proposed an analysis valid for two independent aspects of "animacy" in human grammars by reference to one and the same mechanism. Psychological distance subsumes all the features of PH: agency, emotion, cognition, animacy, and abstraction, and makes it highly plausible that the special features of prominence hierarchies arose out of how the human mind in general cognizes about proximal and distal events and objects; that is, from a subjective perspective towards the external world. All of these individual *manifestations* of prominence can be understood in terms of the overarching notion of psychological distance.

## 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

ML: Conceptualization, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. BL: Conceptualization, Writing - original draft, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## 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.

## Publisher's note

## References

Aikhenvald, A. Y. (2000). Classifiers: A typology of noun categorization devices. Oxford: Oxford University Press.

Aissen, J. (1999). Agent focus and inverse in Tzotzil. *Language* 75, 451–485. doi: 10.2307/417057

Aissen, J. (2003). Differential object marking: iconicity vs. economy. *Nat. Lang. Linguist. Theory* 21, 435–483. doi: 10.1023/A:1024109008573

Arzy, S., Adi-Japha, E., and Blanke, O. (2009). The mental time line: an analogue of the mental number line in the mapping of life events. *Conscious. Cogn.* 18, 781–785. doi: 10.1016/j.concog.2009.05.007

Austin, P. (1981). A grammar of Diyari, South Australia. Cambridge: Cambridge University Press.

Bar-Anan, Y., Liberman, N., and Trope, Y. (2006). The association between psychological distance and construal level: evidence from an implicit association test. *J. Exp. Psychol. Gen.* 135:609. doi: 10.1037/0096-3445.135.4.609

Battistella, E. L. (1996). The logic of markedness. New York: OUP.

Bavelas, J. B., and Gerwing, J. (2011). The listener as addressee in face-to-face dialogue. *Int. J. Listening* 25, 178–198. doi: 10.1080/10904018.2010.508675

Biligiri, H. S. (1965). Kharia: Phonology, grammar and vocabulary. Poona: Deccan College.

Boas, F. (1911). "Introduction," in *Handbook of American Indian Languages. Bureau of American Ethnology, Bulletin 40*. Washington: Government Print Office (Smithsonian Institution, Bureau of American Ethnology), 1, 1–83.

Bohnemeyer, J. (2004). "Argument and event structure in Yukatek verb classes," in *The semantics of under-represented languages in the Americas (SULA)*. GLSA.

Bonato, M., Zorzi, M., and Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neurosci. Biobehav. Rev.* 36, 2257–2273. doi: 10.1016/j.neubiorev.2012.08.007

Bossong, G. (1991). "Differential object marking in romance and beyond" in New analyses in romance linguistics, 69. eds. D. Wanner and D. A. Kibbee (Amsterdam: John Benjamins Publishing Company), 143–170.

Brewer, M. B., and Lui, L. (1984). Categorization of the elderly by the elderly: effects of perceiver's category membership. *Personal. Soc. Psychol. Bull.* 10, 585–595. doi: 10.1177/0146167284104012

Caha, P. (2009). *The nanosyntax of case*. [PhD dissertation]. Tromsø: University of Tromsø. Available at: https://hdl.handle.net/10037/2203 (Accessed December 14, 2023).

Comrie, B. (1986). "Markedness, grammar, people, and the world" in Markedness. eds. F. R. Eckman, E. A. Moravcsik and J. R. Wirth (New York: Plenum Press), 85–106.

Comrie, B. (1989). Language universals and linguistic typology. *2nd* Edn. Chicago: University of Chicago Press.

Connolly, A. C., Gobbini, M. I., and Haxby, J. V. (2011). "Three virtues of similarity based multivariate pattern analysis: an example from the object vision pathway" in Visual population codes. Toward a common common multivariate framework for cell recording and functional imaging. eds. N. Kriegeskorte and G. Kreiman (Cambridge, MA: The MIT Press), 335–356.

Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y. C., et al. (2012). The representation of biological classes in the human brain. *J. Neurosci.* 32, 2608–2618. doi: 10.1523/JNEUROSCI.5547-11.2012

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

Croft, W. (1988). "Agreement vs. case marking and direct objects" in Agreement in natural language: Approaches, theories, descriptions. eds. M. Barlow and C. Ferguson (Stanford, CA: CSLI), 159–179.

Croft, W. (1990). Typology and universals. Cambridge: Cambridge UP.

Croft, W. (2003). Typology and universals. *2nd* Edn. Cambridge: Cambridge UP.

de Hoop, H., and de Swart, P. (2008). "Cross-linguistic variation in differential subject marking" in Differential subject marking. eds. H. de Hoop and P. de Swart (Dordrecht: Springer Netherlands), 1–16.

DeLancey, S. (1981). An interpretation of split ergativity and related patterns. *Language* 57, 626–657. doi: 10.2307/414343

Demarais, A. M., and Cohen, B. H. (1998). Evidence for image-scanning eye movements during transitive inference. *Biol. Psychol.* 49, 229–247. doi: 10.1016/S0301-0511(98)00042-8

Di Pellegrino, G., and Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia* 66, 126–133. doi: 10.1016/j.neuropsychologia.2014.11.011

Dixon, R. M. W. (1979). Ergativity. *Language* 55, 59–138. doi: 10.2307/412519

Dixon, R. M. W. (1994). Ergativity: Cambridge studies in linguistics, 69. Cambridge: Cambridge University Press.

Dohen, M., Schwartz, J. L., and Bailly, G. (2010). Speech and face-to-face communication–an introduction. *Speech Comm.* 52, 477–480. doi: 10.1016/j.specom.2010.02.016

Donahue, E. M., Robins, R. W., Roberts, B. W., and John, O. P. (1993). The divided self: concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *J. Pers. Soc. Psychol.* 64:834. doi: 10.1037/0022-3514.64.5.834

Donohue, C., and Donohue, M. (1998). Fore case marking. *Lang. Ling. Melanesia* 28, 69–98.

Drexel, I. (2014). Animacy and topicality on a scale of cultural context [Thesis]. Available at: http://hdl.handle.net/10066/12513 (Accessed April 21, 2023).

Enfield, N. J. (Ed.) (2002). Ethnosyntax: Explorations in grammar and culture: Explorations in grammar and culture. Oxford: OUP.

Farkas, D. (1978). "Direct and indirect object reduplication in Romanian," in *Papers from the* fourteenth *regional meeting. Chicago Ling. Soc. Chicago, Ill* (Vol. 14), 88–97.

Fauconnier, S. (2011). Differential agent marking and animacy. *Lingua* 121, 533–547. doi: 10.1016/j.lingua.2010.10.014

Fiedler, K., Semin, G. R., and Finkenauer, C. (1993). The battle of words between gender groups a language-based approach to intergroup processes. *Hum. Commun. Res.* 19, 409–441. doi: 10.1111/j.1468-2958.1993.tb00308.x

Fodor, J. A. (1985). Precis of the modularity of mind. *Behav. Brain Sci.* 8, 1–5. doi: 10.1017/S0140525X0001921X

Frank, M. G., and Gilovich, T. (1989). Effect of memory perspective on retrospective causal attributions. *J. Pers. Soc. Psychol.* 57, 399–403. doi: 10.1037/0022-3514.57.3.399

Gair, J. (1970). Colloquial Sinhalese clause structures. The Hague: De Gruyter Mouton.

Gardelle, L., Mignot, E., and Neveux, J. (2024). Nouns and the Morphosyntax/semantics interface. Cham, Switzerland: Palgrave Macmillan.

Gardelle, L., and Sorlin, S. (2018). Introduction: anthropocentrism, egocentrism and the notion of animacy hierarchy. *Int. J. Lang. Cult.* 5, 133–162. doi: 10.1075/ijolc.00004.gar

Geeraerts, D. (2016). Prospects and problems of prototype theory. *Diacronia* 4, 1–16. doi: 10.17684/i4A53en

Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cogn. Sci.* 7, 155–170. doi: 10.1016/S0364-0213(83)80009-3

Givon, T. (1978). "Definiteness and referentiality" in Universals of human language, vol. 4. ed. J. Greenberg (Stanford, CA: Stanford University Press), 291–330.

Goldwater, M. B., and Markman, A. B. (2011). Categorizing entities by common role. *Psychon. Bull. Rev.* 18, 406–413. doi: 10.3758/s13423-011-0058-0

Greenberg, J. (1966). Language universals, with special reference to feature hierarchies. (Janua Linguarum, series minor 59). The Hague: De Gruyter Mouton.

Gregores, E., and Suárez, J. A. (1967). A description of colloquial Guaraní (vol. 27). Berlin: Walter de Gruyter GmbH and Co KG.

Gundel, J. K. (1985). 'Shared knowledge' and topicality. *J. Pragmat.* 9, 83–107. doi: 10.1016/0378-2166(85)90049-9

Haspelmath, M. (2013). 34 occurrence of nominal plurality. *The world atlas of language structures online*. Available at: https://amor.cms.hu-berlin.de/~h2816i3x/Lehre/2007_VL_Typologie/03_Haspelmath_Occurrence_Plurality.pdf (accessed August 21, 2024)

Haude, K. (2024). Between symmetrical voice and ergativity: inverse and antipassive in Movima. *Int. J. Am. Ling.* 90, 1–36. doi: 10.1086/727522

Heath, J. (1980). *Basic materials in Ritharngu: grammar, texts and dictionary*. Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

Helmbrecht, J., Denk, L., Thanner, S., and Tonetti, I. (2018). "Morphosyntactic coding of proper names and its implications for the animacy hierarchy" in Typological hierarchies in synchrony and diachrony. eds. S. Cristofaro and F. Zúñiga (Amsterdam: John Benjamins Publishing Company), 377–401.

Henderson, M. D., Fujita, K., Trope, Y., and Liberman, N. (2006). Transcending the "here": the effect of spatial distance on social judgment. *J. Pers. Soc. Psychol.* 91:845. doi: 10.1037/0022-3514.91.5.845

Hopper, P. J., and Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language* 56:251. doi: 10.2307/413757

Irimia, M. A. (2023). Oblique DOM in enriched case hierarchies. *Glossa* 8, 1–56. doi: 10.16995/glossa.5751

Jackendoff, R., and Audring, J. (2019). "The parallel architecture" in Current approaches to syntax: A comparative handbook. eds. A. Kertész, E. Moravcsik and C. Rákosi (Berlin, Boston: De Gruyter Mouton), 215–238.

Jacques, G., and Antonov, A. (2014). Direct/inverse systems. *Lang. Ling. Compass* 8, 301–318. doi: 10.1111/lnc3.12079

James, W. (1893). The principles of psychology. New York: H. Holt and Company, 609.

Johannessen, J. B. (2008). The pronominal psychological demonstrative in Scandinavian: its syntax, semantics and pragmatics. *Nordic J. Ling.* 31, 161–192. doi: 10.1017/S0332586508001923

Jones, E. E., Wood, G. C., and Quattrone, G. A. (1981). Perceived variability of personal characteristics in in-groups and out-groups: the role of knowledge and evaluation. *Personal. Soc. Psychol. Bull.* 7, 523–528. doi: 10.1177/014616728173024

Joshi, P. D., Wakslak, C. J., Raj, M., and Trope, Y. (2016). Communicating with distant others: the functional use of abstraction. *Soc. Psychol. Personal. Sci.* 7, 37–44. doi: 10.1177/1948550615590449

Kaiser, E. (2009). "Investigating effects of structural and information-structural factors on pronoun resolution" in Information structure from different perspectives. eds. M. Zimmermann and C. Féry (Oxford: Oxford University Press), 332–353.

Kalkstein, D. A., Hackel, L. M., and Trope, Y. (2020). Person-centered cognition: the presence of people in a visual scene promotes relational reasoning. *J. Exp. Soc. Psychol.* 90:104009. doi: 10.1016/j.jesp.2020.104009

Kann, T., Berman, S., Cohen, M. S., Goldknopf, E., Gülser, M., Erlikhman, G., et al. (2023). Linguistic empathy: behavioral measures, neurophysiological correlates, and correlation with psychological empathy. *Neuropsychologia* 191:108650. doi: 10.1016/j.neuropsychologia.2023.108650

Keesing, R. M. (1979). Linguistic knowledge and cultural knowledge: some doubts and speculations. *Am. Anthropol.* 81, 14–36. doi: 10.1525/aa.1979.81.1.02a00020

Kelly, R. E. (1882). The alternative: A study in psychology. London: Macmillan and Co.

Kiani, R., Esteky, H., Mirpour, K., and Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J. Neurophysiol.* 97, 4296–4309. doi: 10.1152/jn.00024.2007

Kiss, K. É. (2017). The person-case constraint and the inverse agreement constraint are manifestations of the same inverse topicality constraint. *Ling. Rev.* 34, 365–395. doi: 10.1515/tlr-2017-0004

Kittilä, S., Västi, K., and Ylikoski, J. (2011). "Introduction to case, animacy and semantic roles" in *Case, animacy and semantic roles*, *99*. eds. S. Kittilä, J. Ylikoski and K. Västi (Amsterdam: John Benjamins), 1–26. doi: 10.1075/tsl.99.01kit

Krantz, D. H. (1975). Color measurement and color theory: I. Representation theorem for Grassmann structures. *J. Math. Psychol.* 12, 283–303. doi: 10.1016/0022-2496(75)90026-7

Kuno, S. (1973). The structure of the Japanese language. Cambridge, Massachusetts: MIT Press.

Kuno, S. (1976). The speaker's empathy and its effect on syntax: a reexamination of Yaru and Kureru in Japanese. *J. Assoc. Teach. Jpn* 11, 249–271. doi: 10.2307/489466

Kuno, S. (1987). Functional syntax: Anaphora, discourse and empathy. Chicago: University of Chicago Press.

Kuno, S., and Kaburaki, E. (1977). Empathy and syntax. *Ling. Inq.*, 627–672. Available at: https://www.jstor.org/stable/4178011

Langacker, R. W. (1987). Foundations of cognitive grammar, vol. 1, theoretical prerequisites. Stanford: Stanford University Press.

Langacker, R. W. (1991a). Foundations of cognitive grammar, vol. 2, descriptive application. Stanford: Stanford University Press.

Langacker, R. W. (1991b). Concept, image, and symbol: The cognitive basis of grammar. Berlin and New York: De Gruyter Mouton.

Langacker, R. W. (1994). "Culture, cognition, and grammar" in Language contact and language conflict, 71. ed. M. Pütz (Amsterdam: John Benjamins), 25–53.

Langacker, R. (2017). Ten lectures on the elaboration of cognitive grammar, vol. 18. Leiden and Boston: Brill.

Leopold, D. A., O'Toole, A. J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat. Neurosci.* 4, 89–94. doi: 10.1038/82947

Li, C. N., and Thompson, S. A. (1981). Mandarin Chinese: A functional reference grammar. Berkeley and Los Angeles: University of California Press.

Libby, L. K., and Eibach, R. P. (2002). Looking back in time: self-concept change affects visual perspective in autobiographical memory. *J. Pers. Soc. Psychol.* 82:167. doi: 10.1037/0022-3514.82.2.167

Liberman, N., Sagristano, M. D., and Trope, Y. (2002). The effect of temporal distance on level of mental construal. *J. Exp. Soc. Psychol.* 38, 523–534. doi: 10.1016/S0022-1031(02)00535-8

Liberman, N., and Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: a test of temporal construal theory. *J. Pers. Soc. Psychol.* 75, 5–18. doi: 10.1037/0022-3514.75.1.5

Liberman, N., and Trope, Y. (2008). The psychology of transcending the here and now. *Science* 322, 1201–1205. doi: 10.1126/science.1161958

Liberman, N., and Trope, Y. (2014). Traversing psychological distance. *Trends Cogn. Sci.* 18, 364–369. doi: 10.1016/j.tics.2014.03.001

Lichtenberk, F. (1983). A grammar of Manam. *Ocean. Linguist. Spec. Publ.* 18:i-647. Available at: http://www.jstor.org/stable/20006696

Linville, P. W. (1982). The complexity–extremity effect and age-based stereotyping. *J. Pers. Soc. Psychol.* 42, 193–211. doi: 10.1037/0022-3514.42.2.193

Linville, P. W., Fischer, G. W., and Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: the outgroup covariation effect. *J. Pers. Soc. Psychol.* 70, 421–436. doi: 10.1037/0022-3514.70.3.421

Liviatan, I., Trope, Y., and Liberman, N. (2008). Interpersonal similarity as a social distance dimension: implications for perception of others' actions. *J. Exp. Soc. Psychol.* 44, 1256–1269. doi: 10.1016/j.jesp.2008.04.007

Lobben, M., Bochynska, A., Tanggaard, S., and Laeng, B. (2020). Classifiers in non-European languages and semantic impairments in western neurological patients have a common cognitive structure. *Lingua* 245:102929. doi: 10.1016/j.lingua.2020.102929

Lockwood, H. T., and Macaulay, M. (2012). Prominence hierarchies. *Lang. Ling. Compass* 6, 431–446. doi: 10.1002/lnc3.345

Logvinenko, A. D. (2015). The geometric structure of color. *J. Vis.* 15:16. doi: 10.1167/15.1.16

Mardale, A., and Karatsareas, P. (2020). Differential object marking and language contact: An introduction to this special issue. *J. Lang. Contact.* 13, 1–16. doi: 10.1163/19552629-bja10001

Maruyama, A. (2003). Japanese Wa in conversational discourse: a contrast marker. *Stud. Lang.* 27, 245–285. doi: 10.1075/sl.27.2.03mar

Matthews, P. H. (2007). The concise Oxford dictionary of linguistics. *2nd* Edn. Oxford: OUP.

McGregor, W. (1992). The semantics of ergative marking in Gooniyandi. *Linguistics* 30, 275–318. doi: 10.1515/ling.1992.30.2.275

McGregor, W. (1998). "Optional" ergative marking in Gooniyandi revisited: implications to the theory of marking. *Leuvense Bijdragen* 87, 491–534.

McGregor, W. (2006). Focal and optional ergative marking in Warrwa (Kimberley, Western Australia). *Lingua* 116, 393–423. doi: 10.1016/j.lingua.2005.02.002

Nigro, G., and Neisser, U. (1983). Point of view in personal memories. *Cogn. Psychol.* 15, 467–482. doi: 10.1016/0010-0285(83)90016-6

Nussbaum, S., Trope, Y., and Liberman, N. (2003). Creeping dispositionism: the temporal dynamics of behavior prediction. *J. Pers. Soc. Psychol.* 84:485. doi: 10.1037/0022-3514.84.3.485

Oberauer, K. (2019). Working memory and attention–a conceptual analysis and review. *J. Cogn.* 2, 1–23. doi: 10.5334/joc.58

Oliveri, M., Koch, G., Salerno, S., Torriero, S., Gerfo, E. L., and Caltagirone, C. (2009). Representation of time intervals in the right posterior parietal cortex: implications for a mental time line. *NeuroImage* 46, 1173–1179. doi: 10.1016/j.neuroimage.2009.03.042

Onea, E., and Mardale, A. (2020). From topic to object: grammaticalization of differential object marking in Romanian. *Can. J. Ling.* 65, 350–392. doi: 10.1017/cnj.2020.12

Osborne, C. R. (1974). The Tiwi language: Grammar, myths and dictionary of the Tiwi language spoken on Melville and Bathurst Islands, northern Australia. Canberra: The Australian Institute of Aboriginal Studies.

Oshima, D. Y. (2007a). Syntactic direction and obviation as empathy-based phenomena: a typological approach. *Linguistics* 45, 727–763. doi: 10.1515/LING.2007.022

Oshima, D. Y. (2007b). On empathic and logophoric binding. *Res. Lang. Comput.* 5, 19–35. doi: 10.1007/s11168-006-9020-0

Park, B., and Judd, C. M. (1990). Measures and models of perceived group variability. *J. Pers. Soc. Psychol.* 59, 173–191. doi: 10.1037/0022-3514.59.2.173

Park, B., and Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: memory for the subordinate attributes of in-group and out-group members. *J. Pers. Soc. Psychol.* 42, 1051–1068. doi: 10.1037/0022-3514.42.6.1051

Park, B., Ryan, C. S., and Judd, C. M. (1992). Role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *J. Pers. Soc. Psychol.* 63, 553–567. doi: 10.1037/0022-3514.63.4.553

Patterson, C., and Schumacher, P. B. (2021). Interpretation preferences in contexts with three antecedents: examining the role of prominence in German pronouns. *Appl. Psycholinguist.* 42, 1427–1461. doi: 10.1017/S0142716421000291

Payne, T. E. (1997). Describing morphosyntax: A guide for field linguists. Cambridge: Cambridge University Press.

Perkins, E. T. (1978). *The role of word order and scope in the interpretation of Navajo sentences*. [PhD dissertation]. Tucson: The University of Arizona.

Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17, 715–734. doi: 10.1017/S0954579405050340

Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27, 53–94.

Ritter, E. (2014). Featuring animacy. *Nordlyd* 41, 103–124. doi: 10.7557/12.3315

Rosch, E. (1978). "Principles of categorization" in Cognition and categorization. eds. E. Rosch and B. Lloyd (London: Routledge), 27–48.

Rosch, E., and Lloyd, B. (1978). Cognition and categorization. (London: Routledge).

Ross, M., and Paol, J. N. (1978). *A Waskia grammar sketch and vocabulary*. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University. doi: 10.15144/PL-B56.cover

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39:1161. doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110:145. doi: 10.1037/0033-295X.110.1.145

Schmidt, R., and Schmidt, R. (2001). "Attention" in Cognition and second language instruction. ed. P. Robinson (Cambridge: Cambridge University Press), 3–32.

Schmitt, M. H. (1972). Near and far: a re-formulation of the social distance concept. *Sociol. Soc. Res.* 57, 85–97.

Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. I. *Psychometrika* 27, 125–140. doi: 10.1007/BF02289630

Silverstein, M. (1976). "Hierarchy of features and ergativity" in Grammatical categories in Australian languages. Linguistic series 22. ed. R. M. W. Dixon (Canberra: Australian Institute of Aboriginal Studies), 112–171.

Simon, H. A. (1971). Designing organizations for an information-rich world. Baltimore, MD: Johns Hopkins University Press, 37–52.

Smith, V. C., and Pokorny, J. (1996). The design and use of a cone chromaticity space. *Color. Res. Appl.* 21, 375–383. doi: 10.1002/(SICI)1520-6378(199610)21:5<375::AID-COL6>3.0.CO;2-V

Smith, P. K., and Trope, Y. (2006). You focus on the forest when you're in charge of the trees: power priming and abstract information processing. *J. Pers. Soc. Psychol.* 90:578. doi: 10.1037/0022-3514.90.4.578

Smith-Stark, T. C. (1974). "The plurality split" in Papers from the tenth regional meeting, Chicago linguistic society, April 19–21, 1974. eds. M. W. La Galy, R. A. Fox and A. Bruck (Chicago: Chicago Linguistic Society), 657–671.

Soderberg, C. K., Callahan, S. P., Kochersberger, A. O., Amit, E., and Ledgerwood, A. (2015). The effects of psychological distance on abstraction: two meta-analyses. *Psychol. Bull.* 141:525. doi: 10.1037/bul0000005

Starke, M. (2017). Resolving (DAT=ACC)≠GEN. *Glossa* 2, 1–8. doi: 10.5334/gjgl.408

Stephan, E. (2006). Social distance and its relations to level of construal, temporal distance and spatial distance: Tel Aviv University.

Stocker, K. (2014). The theory of cognitive spacetime. *Metaphor. Symb.* 29, 71–93. doi: 10.1080/10926488.2014.889991

Strawson, P. F. (1964). Identifying reference and truth-values. *Theoria* 30, 96–118. doi: 10.1111/j.1755-2567.1964.tb00404.x

Swanton, J. R. (1968). *The Handbook of American Indian Languages*. ed. F. Boas (Washington: US Government Printing Office).

Tasaku, T. (1981). Split case-marking patterns in verb-types and tense/aspect/mood. *Linguistics* 19, 389–438. doi: 10.1515/ling.1981.19.5-6.389

Taylor, J. R. (2003). Linguistic categorization. Prototypes in linguistic theory. Oxford: Oxford University Press.

Tiersma, P. M. (1982). Local and general markedness. *Language* 58, 832–849. doi: 10.2307/413959

Tressoldi, P. E., Facco, E., and Lucangeli, D. (2017). On the primacy and irreducible nature of first-person versus third-person information. *F1000Res.* 6, 1–8. doi: 10.12688/f1000research.10752.1

Trope, Y., and Burnstein, E. (1975). Processing the information contained in another's behavior. *J. Exp. Soc. Psychol.* 11, 439–458. doi: 10.1016/0022-1031(75)90047-5

Trope, Y., and Liberman, N. (2003). Temporal construal. *Psychol. Rev.* 110:403. doi: 10.1037/0033-295X.110.3.403

Trope, Y., and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychol. Rev.* 117:440. doi: 10.1037/a0018963

Trope, Y., and Liberman, N. (2012). "Construal level theory" in Handbook of theories of social psychology, vol. 1, 118–134. Available at: http://digital.casalini.it/9781446250068 (Accessed June 23, 2023).

Trope, Y., Liberman, N., and Wakslak, C. (2007). Construal levels and psychological distance: effects on representation, prediction, evaluation, and behavior. *J. Consum. Psychol.* 17, 83–95. doi: 10.1016/S1057-7408(07)70013-X

Tsoneva-Mathewson, S. T. (2009). "Cognitive linguistics" in Linguistics. eds. V. Muhvic-Dimanovski and L. Socanac (Paris, France: EOLSS Publications).

Tsonope, J. (1988). *The acquisition of Setswana noun class and agreement morphology – With special reference to demonstratives and possessives*. [PhD dissertation]. Buffalo: State University of New York.

Valentine, T., Lewis, M. B., and Hills, P. J. (2016). Face-space: a unifying concept in face recognition research. *Q. J. Exp. Psychol.* 69, 1996–2019. doi: 10.1080/17470218.2014.990392

van der Auwera, J., and Gast, V. (2010). "Categories and prototypes" in The Oxford handbook of linguistic typology. ed. J. J. Song (Oxford: Oxford Academic).

Vendler, Z. (1957). Verbs and times. *Philos. Rev.* 66, 143–160. doi: 10.2307/2182371

Verstraete, J. C. (2010). Animacy and information structure in the system of ergative marking in Umpithamu. *Lingua* 120, 1637–1651. doi: 10.1016/j.lingua.2009.05.011

Vinokur, A., Trope, Y., and Burnstein, E. (1975). A decision-making analysis of persuasive argumentation and the choice-shift effect. *J. Exp. Soc. Psychol.* 11, 127–148. doi: 10.1016/S0022-1031(75)80016-3

Vogt, H. (1938). "Esquisse d 'une grammaire du moderne," in Norsk Tidsskrift for Sprogvidenskap, IX and X.

Wakslak, C. J., Nussbaum, S., Liberman, N., and Trope, Y. (2008). Representations of the self in the near and distant future. *J. Pers. Soc. Psychol.* 95, 757–773. doi: 10.1037/a0012939

Wakslak, C. J., Trope, Y., Liberman, N., and Alony, R. (2006). Seeing the forest when entry is unlikely: probability and the mental representation of events. *J. Exp. Psychol. Gen.* 135:641. doi: 10.1037/0096-3445.135.4.641

Walker, M. A. (2002). Toward a model of the interaction of centering with global discourse structure. *Verbum* 22, 1–23. Available at: https://www.researchgate.net/publication/2478110

Welmers, W. E. (1969). The morphology of Kpelle Nominals. *J. Afr. Lang.* 8, 73–101.

Werkman, W. M., Wigboldus, D. H., and Semin, G. R. (1999). Children's communication of the linguistic intergroup bias and its impact upon cognitive inferences. *Eur. J. Soc. Psychol.* 29, 95–104. doi: 10.1002/(SICI)1099-0992(199902)29:1<95::AID-EJSP898>3.0.CO;2-Z

Wierzbicka, A. (1979). Ethno-syntax and the philosophy of grammar. *Stud. Lang.* 3, 313–383. doi: 10.1075/sl.3.3.03wie

Witherspoon, G. (1977). Language and art in the Navajo universe. Ann Arbor, MI: University of Michigan Press.

Yamamoto, M. (1999). Animacy and reference: A cognitive approach to corpus linguistics. Amsterdam: John Benjamins.

Yamamoto, M. (2006). Agency and impersonality: Their linguistic and cultural manifestations. Amsterdam: John Benjamins.

Zúñiga, F. (2014). Inversion, obviation, and Animacy in native languages of the Americas: elements for a cross-linguistic survey. *Anthropol. Ling.* 56, 334–355. doi: 10.1353/anl.2014.0016

Zwicky, A. (1977). "Hierarchies of person," in *Papers from the thirteenth regional meeting of the Chicago linguistic society* (Chicago: Chicago Linguistic Society), 712–733.

# Frontiers in
# Psychology

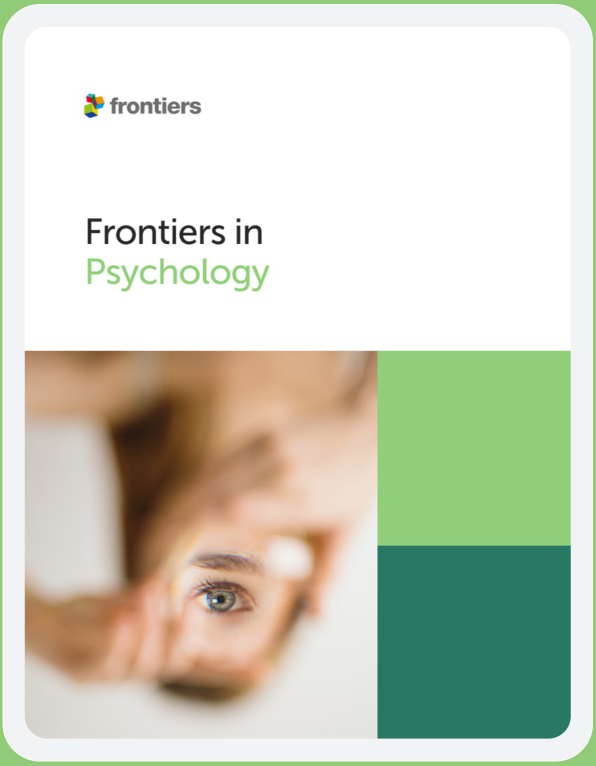
**Paving the way for a greater understanding of human behavior**

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

---

## Discover the latest Research Topics

See more →



frontiers

Frontiers in
Psychology

**Frontiers**

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

**Contact us**

+41 (0)21 510 17 00
frontiersin.org/about/contact