The background of the cover features a red horizontal band across the top. Below and above this band are abstract, textured green patterns composed of overlapping geometric shapes and lines, resembling a complex, layered structure.

CONCEPTUAL CATEGORIES AND THE STRUCTURE OF REALITY: THEORETICAL AND EMPIRICAL APPROACHES

EDITED BY: Paul M. W. Hackett

PUBLISHED IN: Frontiers in Psychology



frontiers

Frontiers Copyright Statement

© Copyright 2007-2019 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-731-1

DOI 10.3389/978-2-88945-731-1

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: researchtopics@frontiersin.org

CONCEPTUAL CATEGORIES AND THE STRUCTURE OF REALITY: THEORETICAL AND EMPIRICAL APPROACHES

Topic Editor:

Paul M. W. Hackett, Visiting Professor, University of Gloucestershire, United Kingdom; Visiting Professor of Health Research Methods, University of Suffolk, United Kingdom; Honorary Fellow, Durham University, United Kingdom; Emerson College, United States

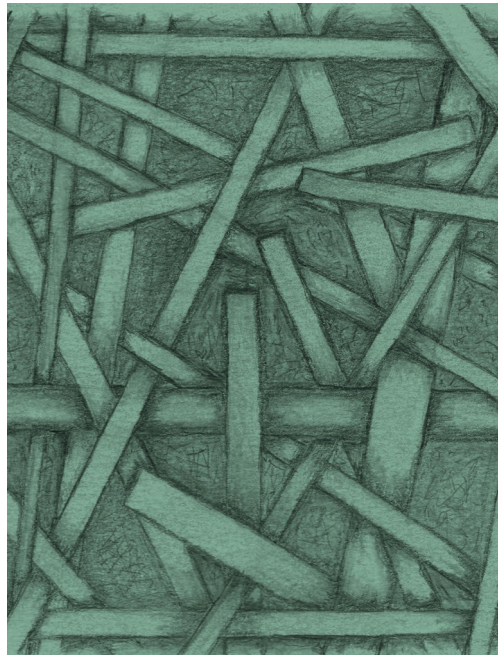


Image: 71015, graphite on paper, Paul Hackett

In this eBook, *Conceptual Categories and the Structure of Reality*, the title very well describes the book's content. Within the book's pages a selection of academics from a variety of human behaviour, human/social science and humanities disciplines write about their research all of which can be typified by their consideration of how categories are used to structure understanding of phenomena. These authors have considered how reality may be understood through notions such as categorial and structural ontologies, part-whole relationships (mereology), the qualitative, quantitative and philosophical use of the facet theory approach to research, mapping sentences and declarative mapping sentence, hermeneutics, concepts and constructs, similarities and differences. The resulting collection presents the foregoing conceptual and empirical approaches to knowledge development in general (chapter 1&3 Hackett); Phillips and Wislons' review of compositional syntax in bird

calls (chapter 2); neurobehavioral decision systems (chapter 4 Foxall); representations of human psychological processes (chapter 5 Juan-Miguel López-Gil; Rosa Gil; Roberto García); free associations mirroring and its relation to self- and world-related concepts (chapter 6 Martin Kuška; Radek Trnka; Aleš Antonín Kuběna; Jiří Růžička); local knowledge and going beyond the data (chapter 7 Steven Phillips); categorical etiologies of speech sound disorders (chapter 8 Kelly Farquharson); similarity of visual appearance (chapter 9 Nao Nakatsuji; Hisayasu Ihara; Takeharu Seno; Hiroshi Ito); and a consideration of the seminal writing of David Oderberg's on the categorial classification of reality (chapter 10 Hackett).

Citation: Hackett, P. M. W., ed. (2019). *Conceptual Categories and the Structure of Reality: Theoretical and Empirical Approaches*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-731-1

Table of Contents

- 05 Editorial: Conceptual Categories and the Structure of Reality: Theoretical and Empirical Approaches**
Paul M. W. Hackett
- 07 Commentary: Experimental Evidence for Compositional Syntax in Bird Calls**
Steven Phillips and William H. Wilson
- 11 Facet Theory and the Mapping Sentence as Hermeneutically Consistent Structured Meta-Ontology and Structured Meta-Mereology**
Paul M. W. Hackett
- 17 Metacognitive Control of Categorical Neurobehavioral Decision Systems**
Gordon R. Foxall
- 35 Web Ontologies to Categorical Structure Reality: Representations of Human Emotional, Cognitive, and Motivational Processes**
Juan-Miguel López-Gil, Rosa Gil and Roberto García
- 47 Free Associations Mirroring Self- and World-Related Concepts: Implications for Personal Construct Theory, Psycholinguistics and Philosophical Psychology**
Martin Kuška, Radek Trnka, Aleš A. Kuběna and Jiří Růžicka
- 60 Going Beyond the Data as the Patching (Sheaving) of Local Knowledge**
Steven Phillips
- 73 Language or Motor: Reviewing Categorical Etiologies of Speech Sound Disorders**
Kelly Farquharson
- 76 Visualizing Similarity of Appearance by Arrangement of Cards**
Nao Nakatsuji, Hisayasu Ihara, Takeharu Seno and Hiroshi Ito
- 83 Book Review: "Classifying Reality," by David S. Oderberg (ed.) (2013)**
Paul M. W. Hackett



Editorial: Conceptual Categories and the Structure of Reality: Theoretical and Empirical Approaches

Paul M. W. Hackett *

School of Communication, Emerson College, Boston, MA, USA

Keywords: categories, ontology, mereology

Editorial on the Research Topic

Conceptual Categories and the Structure of Reality: Theoretical and Empirical Approaches

The process or activity of forming mental categories is both fundamental and necessary to the existence of all living creatures from bi-polar categories such as safe-dangerous, wet-dry, edible-not edible, to higher and more complex forms of categorial discrimination (see, for example, Greggor and Hackett, 2017). The worlds within which sentient creatures live are made sense of through categories. The identification of categories, and the development of categorial understandings, are also important within research.

This special edition of Frontiers in Psychology stands in support of this last claim with papers by researchers from a diverse selection of disciplinary backgrounds. For instance, there are contributions that employ, a range of psychological perspectives; philosophical, personal construct, cognitive, perceptual, neuropsychological, etc. (for recent research in these areas see: Mahon and Caramazza, 2009; Khalidi, 2015). The articles also reflect a spread of approaches from theoretical to applied and demonstrate the ways in which categorization is understood and employed in research from the sciences and social sciences.

Notions of parthood and part-whole interaction are similarly concepts that permeate many academic areas. In this special edition Hackett reviews parthood in his paper, “Facet Theory and the Mapping Sentence As Hermeneutically Consistent Structured Meta-Ontology and Structured Meta-Mereology.” In both of his contributions Hackett proffers a theoretical framework for the design and interpretation of research that uses categories. He places emphasis upon the identification of super-ordinate ontologies that are present within a research domain and the explication of mereological (part to part and part to whole) relationships within these ontologies. This meta-theoretical approach to knowledge development in the social sciences is couched within the rubric of the mapping sentence and facet theory. Hackett argues that by adopting this approach it is possible to achieve hermeneutic consistency with theoretical and empirical validity.

Another contribution that stresses the mereological arrangement of the categories they identify is that by Kuška et al. These authors considered categorization, or the use of categories to generate knowledge, in an applied and particularistic sense, whilst conducting theory-driven research. In their paper “Free Associations Mirroring Self- and World-Related Concepts,” the authors couched their linguistic research within the framework of personal construct psychology, presenting findings from the use of free-association to investigate how reality was construed. They claim their findings indicated that people construe reality through the employment of basic units of meaningful categorization. Their methodology required respondents to offer words that related to the words *world* and *self*. They claim that this procedure accesses, in a relatively direct manner, what they called the basic units of meaningful categorization. They discovered that some categories were expressed as semantic polarities such as nature versus culture. Some of the other verbal category

OPEN ACCESS

Edited and reviewed by:

Markus Werning,
Ruhr University Bochum, Germany

*Correspondence:

Paul M. W. Hackett
paul@hackettspace.com

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 02 February 2017

Accepted: 31 March 2017

Published: 20 April 2017

Citation:

Hackett PMW (2017) Editorial:
Conceptual Categories and the
Structure of Reality: Theoretical and
Empirical Approaches.
Front. Psychol. 8:601.
doi: 10.3389/fpsyg.2017.00601

groupings were related to respondent understanding of the words *world* and *self*-whilst others mediated pathways in this category-based network.

The article, “Language or motor: reviewing categorical etiologies of speech sound disorders” by Farquharson, is also concerned with linguistic categories. However, in this instance the research presented addresses speech disorders. With foundations in her category-based consideration of this applied area, Farquharson stresses the need for better conceptualizations of the mechanisms associated with speech sound disorders. In doing this, she hopes that this will lead to an improvement in the diagnosis and treatment of children with these forms of disorder.

Questions regarding methodological issues associated with identifying categories in psychological research, are posed by Nakatsuji et al. They employed a categorically related form of data analysis by analyzing respondents’ similarity ratings. This is a less complex form of multidimensional scaling (MDS) which analogously explicates psychological space through analysing pair-wise evaluations. In the study reported, participants completed a sort procedure, arranging cards on the basis of their degree of similarity. The authors rigorously compare their approach to traditional MDS and discovered their results obtained closely resembled those obtained using non-metric MDS. However, they argue that their approach was parsimonious, needing approximately one third of the time to complete. They further claim superior sensitivity for their approach. In conclusion, the authors proposed their category deriving method to save time when conducting research that assesses the similarity of appearance.

Foxall’s paper, Metacognitive Control of Categorical Neurobehavioral Decision Systems (CNDS), offers a highly theoretical real-world application of the effects of categories of neural activity. Human decision-making, says Foxall, often involves a person ignoring the future consequences of their decisions and this disregard is dependent upon activity within the limbic and paralimbic regions of the brain relative to activity

in the prefrontal cortex. His model depicts the relationship between categorically distinct neurophysiological, behavioral and cognitive systems. The degree of balance achieved between these categories results in, he claims, normal or addictive behavior. The author discusses these neural elements and proposes a category-based structure to allow understanding of the effects of CNDS on behavior.

In his review of Oderberg’s (2013) book, *Classifying Reality*, Hackett again proffers the mapping sentence as a declarative tool that, in this instance, may enable understanding of the writing about ontologies by other scholars, which allows the development of a categorical structure for experiences contained in an ontology. A similar perspective is taken by López-Gil et al. in their use of web ontologies to categorically structure reality. Their ontology uses the semantic web language OWL (Web Ontology Language) to represent rich and a complex mereologically associated knowledge of the world. In taking an applied outlook their view depicts online students’ emotional, cognitive and motivational state as a web ontology, which interacts with distance or blended educational systems. Their categorical ontology, which they have empirically tested, does not impose a specific way of organizing emotional responses but is able to model reality in association to student affect and motivation.

As can be seen from the above descriptions, the contributions in the special edition represent an eclectic mixture of empirical and theoretical psychological approaches. These have been applied to investigate phenomena that are categorical or have used categories to better understand psychological events. Together, the papers offer insight into contemporary use of category-based knowledge.

AUTHOR CONTRIBUTIONS

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

REFERENCES

- Greggor, L., and Hackett, P. M. W. (2017). “Categorization by the animal mind,” in *Mereologies, Ontologies and Facets: The Categorical Structure of Reality*, ed P. M. W. Hackett (Lanham, MD: Lexington Press).
- Khalidi, M. A. (2015). *Natural Categories and Human Kinds: Classification in the Natural and Social Sciences*. Cambridge: Cambridge University Press.
- Mahon, B. Z., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annu. Rev. Psychol.* 60, 27–51. doi: 10.1146/annurev.psych.60.110707.163532
- Oderberg, D. S. (ed.). (2013). *Classifying Reality*. Chichester: John Wiley and Sons Ltd.
- Conflict of Interest Statement:** 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.

Copyright © 2017 Hackett. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Commentary: Experimental evidence for compositional syntax in bird calls

Steven Phillips^{1*} and William H. Wilson²

¹ Mathematical Neuroinformatics Group, Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan, ² School of Computer Science and Engineering, University of New South Wales, Sydney, NSW, Australia

Keywords: systematicity, compositionality, category theory, syntax, universality

A commentary on

Experimental evidence for compositional syntax in bird calls

by Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2016). *Nat. Commun.* 7:10986. doi: 10.1038/ncomms10986

OPEN ACCESS

Edited by:

Richard Stöckle-Schobel,
Ruhr University Bochum, Germany

Reviewed by:

Ross Gayler,
Formerly affiliated with La Trobe
University, Australia
Michelle Spierings,
Leiden University, Netherlands

*Correspondence:

Steven Phillips
steve@ni.aist.go.jp

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 June 2016

Accepted: 22 July 2016

Published: 03 August 2016

Citation:

Phillips S and Wilson WH (2016)
Commentary: Experimental evidence
for compositional syntax in bird calls.
Front. Psychol. 7:1171.
doi: 10.3389/fpsyg.2016.01171

Suzuki et al. (2016) report a remarkable discovery: the first evidence of a combinatorial syntax and semantics in non-humans; specifically, Japanese great tits. However, remarkable discoveries require remarkable evidence. Their data provide impressive support for a compositional syntax. Yet, evidence for compositionality is not necessarily evidence for one of the hallmarks of human language and thought: *systematicity*—a structural equivalence relation over cognitive capacities (Fodor and Pylyshyn, 1988). Some versions of compositionality support systematicity and some do not (Aizawa, 2003; Phillips and Wilson, 2010). We surmise that the question remains open as to whether the version of compositionality that is evident in the bird calls study does indeed support systematicity. Drawing on a theory of systematicity (Phillips and Wilson, 2010) we derive testable criteria for systematicity in the context of bird calls. These criteria must be met before claims of *human-like* compositional syntax in non-humans could be justified.

Systematicity is a property of (some core aspects of) human language and thought whereby having the capacity to understand certain expressions or situations implies having the capacity to understand certain other, structurally related, expressions/situations (Fodor and Pylyshyn, 1988). The archetypal example of systematicity is where one has the capacity to understand the sentence “John loves Mary” if and only if one has the capacity to understand the sentence “Mary loves John,” assuming one also understands the constituents John, loves, and Mary, where the common structural relation between entities John and Mary is love. Other forms of systematicity follow from the systematic nature of thought, generally. For example, in reasoning, if one is told that John and Mary went to the store, then one can infer that John went to the store—*P* and *Q* implies *P* (see Fodor and Pylyshyn, 1988; McLaughlin, 2009). Hence, systematicity is a central property of human language and thought that warrants investigation in non-humans if the evolutionary story is to be properly told.

The authors’ claim of a compositional syntax and semantics for Japanese great tits aligns with some aspects of the classical (symbol systems) notion of compositionality, which is sometimes called *classical compositionality*. The experiments revealed that great tits extracted different meanings for notes and their syntactic compositions: an ABC note means “scan for danger,” a D note means “approach the caller,” and their syntactic combination ABC-D means “scan for danger then approach the caller,” whereas the (agrammatical) combination D-ABC has no meaning. The classical compositionality account says that the meaning of a complex

utterance is understood from the meanings of the constituent utterances and the correspondence between syntactic relationships among those constituents and their semantic relationships. In the case of great tits, the meaning of bird call ABC-D (scan, then advance) is understood from the meanings of the constituent calls ABC (scan) and D (advance) and the correspondence between their syntactic relationship (ABC is-followed-by D) and their corresponding semantic relationship (scanning is-followed-by advancing). Notice, importantly, that the common structural relation in the bird calls case is not simply temporal order, as evidenced by the counterexample, D-ABC. Accordingly, classical compositionality includes typing (relational role) information that determines the allowable syntactic constructions, which are supposed to be aligned with the corresponding semantic structural relations (see below for further examples). Failure to make this distinction may be seen as one reason why simple associative/statistical models that are based on co-occurrence relations among the constituents of complex capacities (e.g., words in sentences) fail to account for the complexity and subtlety of human language and thought (Fodor and Pylyshyn, 1988; Everaert et al., 2015).

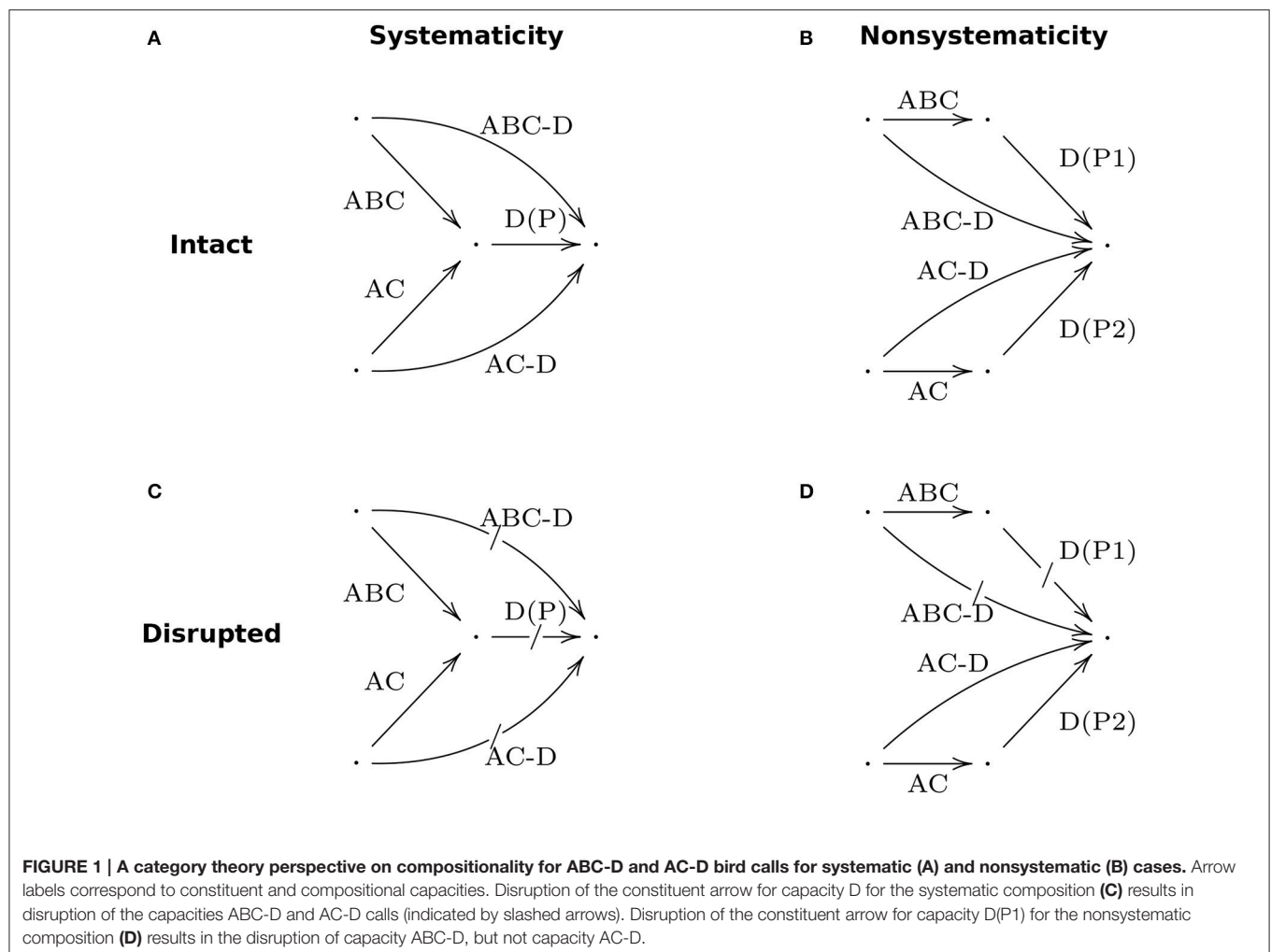
The point of departure from the authors' claim of compositionality and the notion of classical compositionality in humans is in regard to systematicity. A demonstration of systematicity requires evidence that the capacity for two structurally equivalent abilities is *indivisible*. For the *loves* example, such a demonstration involves evidence that the capacity to understand "John loves Mary" is equivalent to the capacity to understand "Mary loves John," because these two capacities share the same syntactic/semantic relation, *loves*. Put simply, there is no situation of having one capacity, but not the other. Clearly, however, this form of (symmetric) structural similarity is meaningless to a great tit in the context of predator deterrence, presumably because the capacity to understand D-ABC, i.e., advance before scanning for predators, has severe consequences for survival. Similar, nonsensical situations arise in human language. For instance, one can say "John fed hay (to the horse)," but it makes no sense to say "Hay fed John." Systematicity need not be confined to symmetric structures. Instances of systematicity based on asymmetric structures also exist, for example, where one has the capacity to understand the sentence "John fed hay" if one has the capacity to understand the sentence "Mary fed hay," assuming that one understands the constituents John, fed, hay and Mary.

An analogous criterion for great tits can also be derived as a test for the systematicity of their "linguistic" ability, using a category theoretic approach to systematicity (Phillips and Wilson, 2010, 2014). Category theory is a branch of mathematics for reasoning about systems of entities and their relationships: a system regarded as a *category* consists of a collection of *objects*, a collection of relations between objects, called *arrows*, *morphisms*, or *maps*, and a composition operation that composes pairs of (compatible) arrows into other arrows. A category theory approach to modeling a cognitive system is to regard sets of cognitive states as objects, cognitive processes that map states to states as arrows, and composition of cognitive processes as the composition operation. In this way, a category

theory explanation for systematicity says that every cognitive capacity in a collection of systematically related capacities is the composition of a *universal* arrow that is common to all capacities in that collection and a unique arrow that is specific to that particular capacity, so one has each and every capacity if and only if one has the universal arrow and a way to compose arrows, which includes the universal arrow with the unique arrows. An advantage of deriving criteria from a category theory perspective on systematicity, as opposed to other approaches, is that it isolates just those properties that are essential (i.e., necessary and sufficient) for systematicity from those properties that are idiosyncratic to the domain at hand. In particular, the characterizations of systematicity introduced above, which were drawn from the original classicists' perspective, presume peripheral (symbolic) capacities that are far beyond those of non-humans, whence it is unclear how systematicity is even testable in non-human cohorts. In contrast, a categorical, "objects and arrows" perspective generalizes the notion of compositionality in a way that affords realistic tests of systematicity in non-humans. A visual/geometric intuition of the formalism that underlies the example, provided next, is given in **Figure 1**.

The authors explain that great tits have a variety of calls associated with different predators, such as AC-D and BC-D. Then an instance of systematicity is when a great tit demonstrates the capacity to understand the ABC-D calls if and only if it demonstrates the capacity to understand the structurally-related AC-D or BC-D calls. Because systematicity is a structural equivalence relation over capacities (McLaughlin, 2009), demonstrating that a bird understands both ABC-D and AC-D calls is only half of the criteria for systematicity in this example. One must also demonstrate that there exists a component process that when absent or disrupted results in the absence or disruption of both ABC-D and AC-D capacities, not the exclusive disruption of just one or the other capacity. Naturally, this criterion for systematicity extends to situations of having more than two structurally related capacities, e.g., ABC-D, AC-D, and BC-D. The case where only one capacity is disrupted, say ABC-D but not AC-D, is evidence against the kind of compositionality possessed by humans. The essential problem to be addressed empirically is that there are two ways to realize capacities ABC-D and AC-D: (1) via a shared component process P that realizes constituent capacity D, and (2) via distinct component processes P1 and P2 that separately realize constituent capacity D for the capacities ABC-D and AC-D, respectively. In the first case, disruption of P implies disruption of both capacities—systematicity. In the second case, disruption of P1 implies disruption of the ABC-D capacity, but not the AC-D capacity, since the component process P2 is intact—no systematicity.

As the authors point out, studies of language-like behavior in non-humans are important to establish the missing link in the evolutionary story of human language. Systematicity is afforded by effective (re)use of cognitive resources, as the categorical perspective highlights. The potential relationship between cognitive resource use (cost) and cognitive capacity to successfully interact with the environment (benefit) leads



naturally to important questions regarding the extent that compositionality is driven by environmental forces vs. genetic good fortune (Hauser et al., 2002), or some combination of these. For instance, one can envisage situations where the lack of variability in the environment places little demand for a systematic compositional syntax—e.g., small variation in predator types, whereby each situation is represented without representing the common structural relations (nonsystematic compositionality). Alternatively, environments filled with different types of predators requiring different types of related responses may drive systematic compositionality. Such situations would suggest that systematic compositionality is driven by environmental forces. However, systematicity in the absence of diverse environmental contingencies suggests fortuitous genetic endowment, in which case the environment plays a lesser role.

Empirical data that dissociate systematic from nonsystematic compositionality in other species is evidence of a branch point in the evolution of human language and thought, a hallmark of which is the sheer diversity of situations that are within the capacity of our cognitive system.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by a Japanese Society for the Promotion of Science Grant-in-aid (26280051).

REFERENCES

Aizawa, K. (2003). *The Systematicity Arguments*. Studies in Mind and Brain. New York, NY: Kluwer Academic.

Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., and Bolhuis, J. J. (2015). Structures, not strings: linguistics as part of the cognitive sciences. *Trends Cogn. Sci.* 19, 729–743. doi: 10.1016/j.tics.2015.09.008

- Fodor, J. A., and Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71. doi: 10.1016/0010-0277(88)90031-5
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569
- McLaughlin, B. P. (2009). Systematicity redux. *Synthese* 170, 251–274. doi: 10.1007/s11229-009-9582-0
- Phillips, S., and Wilson, W. H. (2010). Categorical compositionality: a category theory explanation for the systematicity of human cognition. *PLoS Comput. Biol.* 6:e1000858. doi: 10.1371/journal.pcbi.1000858
- Phillips, S., and Wilson, W. H. (2014). “A category theory explanation for systematicity: universal constructions,” in *The Architecture of Cognition: Rethinking Fodor and Pylyshyn’s Systematicity Challenge*, Chapter 9, eds P. Calvo and J. Symons (Cambridge, MA: MIT Press), 227–249. doi: 10.7551/mitpress/9780262027236.003.0009
- Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nat. Commun.* 7:10986. doi: 10.1038/ncomms10986
- Conflict of Interest Statement:** 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.
- Copyright © 2016 Phillips and Wilson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Facet Theory and the Mapping Sentence As Hermeneutically Consistent Structured Meta-Ontology and Structured Meta-Mereology

Paul M. W. Hackett *

¹ School of Communication, Emerson College, Boston, MA, USA, ² Department of Philosophy, University of Oxford, Oxford, UK

OPEN ACCESS

Edited by:

Richard Stöckle-Schobel,
Ruhr-University Bochum, Germany

Reviewed by:

John G. Holden,
University of Cincinnati, USA
Patrick Mair,
Harvard University, USA

*Correspondence:

Paul M. W. Hackett
paul_hackett@emerson.edu

An earlier version of this paper (Koval and Hackett, 2015) was presented at the 14th International Facet Theory Conference, Fordham University, New York, August 2015.

Specialty section:

This article was submitted to Theoretical and Philosophical Psychology, a section of the journal *Frontiers in Psychology*

Received: 09 January 2016

Accepted: 17 March 2016

Published: 31 March 2016

Citation:

Hackett PMW (2016) Facet Theory and the Mapping Sentence As Hermeneutically Consistent Structured Meta-Ontology and Structured Meta-Mereology. *Front. Psychol.* 7:471. doi: 10.3389/fpsyg.2016.00471

When behavior is interpreted in a reliable manner (i.e., robustly across different situations and times) its explained meaning may be seen to possess hermeneutic consistency. In this essay I present an evaluation of the hermeneutic consistency that I propose may be present when the research tool known as the mapping sentence is used to create generic structural ontologies. I also claim that theoretical and empirical validity is a likely result of employing the mapping sentence in research design and interpretation. These claims are non-contentious within the realm of quantitative psychological and behavioral research. However, I extend the scope of both facet theory based research and claims for its structural utility, reliability and validity to philosophical and qualitative investigations. I assert that the hermeneutic consistency of a structural ontology is a product of a structural representation's ontological components and the mereological relationships between these ontological sub-units: the mapping sentence seminally allows for the depiction of such structure.

Keywords: ontology, mereology, facet theory, mapping sentence, meta-ontology, meta-mereology

INTRODUCTION

When thinking about the world around us it is commonplace and may even seem natural to sub-divide our experiences in attempting to achieve better understanding. The practice of partitioning research content has a long history dating back to at least the time of the ancient, classical philosophers, where such well-known examples include ontologies by Aristotle (1975) and Plato (Harte, 2002). During the subsequent millennia, categorial ontologies have been developed by a wide range of psychologists and philosophers, each of who have concerned themselves with attempting to understand the basic components of human existence (see for example: in psychology, Piaget and Inhelder, 1969, Kelly, 2013; in philosophy, Chisholm, 1996). Given the multitude of ontologies and other componential existential models that exist, the question may be asked as to whether a meta-ontology may be developed that speaks about how ontologies may be understood in structurally theoretical terms. Moreover, questions may also be posed as to the possibilities of developing a meta-mereological structure, which explicates the combined relations of the meta-ontology. During

this essay I provide answers to these questions¹, however, I will initially clarify the precise terms of my exposition.

DEFINING TERMS

In the title of this essay I have employed three phrases that qualify my understanding of the requirements of categorial research investigations: *hermeneutic consistency*, *structured meta-ontology*, and *structured meta-mereology*. These expressions have been carefully selected to emphasize what I believe a qualitative facet theory approach is able to achieve and an initial review of these terms will explicate the nature of the ontology/mereology in which I am interested.

HERMENEUTIC CONSISTENCY

Hermeneutical is an adjective that implicates and focuses ontologies as being interpretative tools. *Hermeneutically consistent* implies that the ontology I offer is reliable in terms of the structure and the interpretation of its content. In the usage of the phrase hermeneutic consistency, hermeneutical refers to a specific interpretive methodology as understood through the writing of Heidegger (1962) and Gadamer (2013). These authors were interested in knowledge and truth and in their work the phrase *hermeneutic consistency* refers to the ability to achieve a coherent explanation of an informational source. Many other philosophers, especially epistemologists, are interested in knowledge and truth and the coherence of explanations about sources of information². However, Heidegger and Gadamer are of particular import as it may be claimed that their influence has spread more widely than some other scholars. For instance, both Heidegger and Gadamer are commonly cited within sociological, psychological and perhaps most importantly to this paper, within research design lecture series and textbooks. Furthermore, the hermeneutical process is of great importance within disciplines that seek interpretation of complex events (as an illustration see: Osborne, 2007; Porter and Robinson, 2011) who provide introductory accounts of hermeneutical processes in reading scripture. In the same way, facet theory based interpretations are also concerned with the interpretative interplay between an event and those experiencing and attempting to understand these occurrences.

¹In this paper I present facet theory and the mapping sentence as a philosophical/qualitative approach to the study of behavior. Criticisms of inadequacy can be made of analytic systems of formal logic (the mapping sentence may be seen as one such system) in their difficulty in differentiating meaning such as those embodied in sarcasm and irony versus sincerity where formally these may be indistinguishable [see for example the work of Gibbs (e.g., Gibbs and Colston, 2012) and the collection of writings by (Gibbs and Colston, 2007)]. However, the mapping sentence would address this difference by the inclusion of elements of sincerity and sarcasm within a content facet of degree of genuineness. Furthermore, there have been many developments, such as self-organizing systems and impredicative declarations which are of interest to the development of system based definition but beyond the scope of this article. The interested reader is guided toward the writing of Turvey and Moreno (2006).

²This ability is studied or thought about in many disciplines and contexts other than philosophical.

ONTOLOGY

Ontology refers to the basic components underlying nature of experience, and structured ontology explicates such understanding within a determinate composition. Ontology has slightly dissimilar meanings when used within the different disciplines that have incorporated ontology into part of their lexicon and way of thinking. For example: in *philosophy*—ontology is a branch of metaphysics concerned with the nature of being; within *logic*—ontology is the set of entities that a given theory assumes beforehand; in *technology*—ontology provides a systematic explanation of existence; within *information and computer sciences*—ontology is the rigorous designation of existent components (sorts, characteristics) and their inter-associations. From these definitions it can be seen that to some extent there are common elements in what ontology is taken to mean. Ontology may therefore appear to refer to being and components of existence, which are perhaps instantiated by a scholar prior to consideration of a content area. Given the differences in the use of the term ontology I wish to escape any possible confusion that may arise by providing a precise definition and understanding of ontology:

Ontology is the study and formal explication of a domain of content in terms of its more fundamental or basic categorial components as these may be understood at this fundamental level and as their meaning may be further revealed through consideration of more sub-ordinate, particular, or evident categorial entities.

I use the term *meta-ontology* to imply that the qualitative ontology I propose constitutes an ontology about ontologies rather than being an ontology of a specific or substantive content area³. My use of this term refers to an ontology of the different, often instrumentalist, ontologies that different disciplines of enquiry adopt to characterize and delimit their frameworks.

Furthermore, the term *structured ontology* and *structured mereology* respectively bring together the concepts of ontology and mereology (or the underlying nature of experience) within a determinate structural template under the definition of ontology I have provided. The next term in my title is mereology.

MEREOLGY

Mereology is concerned with attempts to understand the relationships between, and implications of, part-to-whole and part-to-part associations within a categorial system or ontology. Mereology is defined within metaphysics as: "... any theory of part hood or composition." (Harte, 2002, p7). However, as with the term ontology, mereology is understood in slightly different manners dependent upon the discipline of usage (e.g., philosophy, science, logic, mathematics, semantics). I wish to avoid possibilities of confusion and misinterpretation and I therefore provide my own definition of mereology as follows:

³I am using meta-ontology in the sense of theory that underlies a generic framework for the constitution of interpretative consistency of a research domain, whilst not imposing an external structure to either the content under investigation or to the interpretation itself.

Mereology is the systematic and explicit investigation, analysis and resulting understanding of the relationships within a structured ontology, in terms of the part to part, part to whole, part to context, part to background, and part to observation range, relationships.

A meta-mereology is a mereology that is concerned with the nature of mereologies rather than the content of any particular or specific mereology⁴. Structured meta-mereology implicates an interest in the configuration of mereological relationships. I must provide one final definition that applies to my specifications of both ontology and mereology. On these understandings, ontologies and mereologies exist where and when:

Context and background are essential and inherent components of the existence and realization of the structured ontological/mereological system, where changes in background and context would result in significant differences in the structured ontology/mereology, and where the specification of a different range of observations would significantly alter the content of the structured ontology/mereology and the nature of knowledge embodied within such structure.

So far I have provided a limit to the scope of my essay and in the following sections I offer facet theory and the mapping sentence as a means for achieving a structured ontology/mereology under the constraints of these definitions. I advance my ontology/mereology under the belief that if a researcher understands the components of the behaviors of interest and the interrelationships between these components, a greater appreciation of the total behavior may result.

QUALITATIVE FACET THEORY AND THE MAPPING SENTENCE

Louis Guttman originated facet theory with an implicitly point of view that understand human activities and knowledge about such activities as being formed of discrete components (Guttman, 1947; Levy, 1994). Guttman (1959, p130) defined a facet as "... a set that is a component of a Cartesian product." and in his authoritative text, Canter (1985a, p22) states how a facet is constituted as a "... labeling of a conceptual categorization underlying a group of observation." Facet theory has been defined as, "a strategy for research in psychology and other sciences that study complex behavioral systems. Facet Theory centers on the formalization of research contents and on intrinsic data analysis for the purpose of discovering stable laws and conducting theory-based measurements in those sciences"⁵.

Facet theory has traditionally been based in quantitative research approaches and the statistical analysis (e.g., Borg and

Shye, 1995; Canter, 1985a,b; Shye, 1978; Shye and Amar, 1985; Shye and Elizur, 1994). After having used facet theory in a traditionally quantitative manner, Hackett (2013, 2014) has, over the past few years, developed a qualitative facet theory⁶. During the course of this brief essay I offer a qualitative⁷ facet theory approach as an instantiation of a meta-ontology and meta-mereology. In this paper I evaluate facet theory, and its major instrument the mapping sentence, as a qualitative and philosophical stance toward the understanding of behavior.

The philosophical and theoretical bases of facet theory along with qualitative facet theory approach to research design, data collection and analysis is best understood and reported using the mapping sentence. A mapping sentence is a formal statement of a research domain which includes the respondents, sub-categories of the research content along with the range over which observations will be made, in the structure of a sentence written in normal prose. The mapping sentence is both the major tool of facet theory research design and analysis and also a series of structural/spatial hypotheses. As Canter (1985b) says: "... a piece of facet research is a process of refinement, elaboration and validation of a mapping sentence." (p266): I will be using a mapping sentence in precisely these terms in this paper. Philosophically, the mapping sentence is a structural ontology and in application to any substantive area of research and understanding may also be seen as a mereological device. Related to the notion of the mapping sentence is that a mereology is a compositional identity, where composition is the relation between a whole and its specific parts, in which parts form the whole and where the whole is nothing more than its parts: the whole *is* its parts and parts may only be understood within the whole (see, Cotnoir and Baxter, 2014).

In qualitative facet theory and within a facet theoretical philosophy two central theses arise from the above definitions of ontology and mereology:

When taken together, a specified structured ontology and a mereological account of this structure form what is known as a mapping sentence.

For any specified area of interest, a mapping sentence provides a hermeneutically consistent account of a domain of interest. Thus, facet theory and specifically the mapping sentence is well characterized through the use of the terms structural ontology and mereology with the explicit intent of developing hermeneutically consistent knowledge.

⁴I employ term meta-mereology in the same sense as I use meta-ontology (see footnote 2).

⁵Borg (1978, p65) defined facet theory as: "... a general methodology for investigation in the social sciences: it provides a general framework for the precise definition of an universe of observations, which is directly related to both the specification of the various elements of empirical studies (stimuli, subjects, responses) and to theories about the structure of those observations."

⁶Hackett has carried-out qualitative analyses of specific research domains using mapping sentences as a framework for the conception and design of research projects. He has then progressed by analysing qualitative and conceptual data within a facet theory mind-set to allow theory development. This has required Hackett to use facet theory as a philosophical perspective that he has taken when viewing the subject matter of his research into human behavior (Hackett, 2013, 2014).

⁷In facet theory the term qualitative has been used to mean a qualitatively arranged facet rather than a linear or quantitatively ordered facet: This is not the sense of the word that I will employ. I use qualitative to imply rich observational, non-numerical information. Subsequently I analyse data to establish reliable and valid interpretative hermeneutics.

Person (x) understands the world to exist in terms of kinds that are:

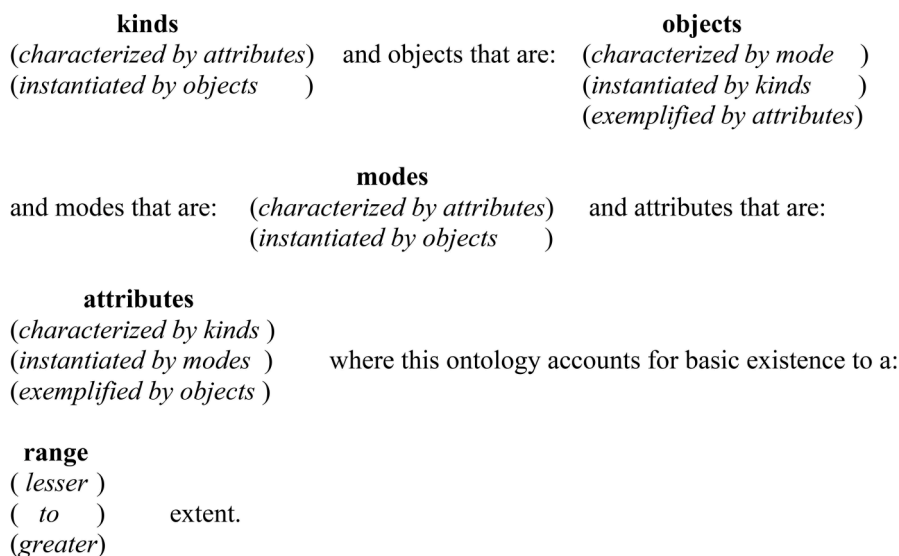


FIGURE 1 | Mapping sentence for Lowe's four-category ontology.

EXAMPLE OF A QUALITATIVE MAPPING SENTENCE

In earlier research I have demonstrated the utility of a non-numerically based facet theory that employs the conceptual rigor that the mapping sentence has provided in my investigation of the mereology of Aristotle's *Categories* (Aristotle, 1975)⁸. This mapping sentence offers an account of *The Categories* that clearly displays Aristotle's ontology and uniquely a potential mereological relationship between the *Categories* parts-to-parts and parts-to-whole and in so doing offers further exploration of Aristotle's ontology. In figure 1 I provide a mapping sentence for a more contemporary ontology by Lowe (2007) in his *four-category ontology*⁹. Lowe's ontology embodies the notion that the world may be understood as comprising three distinct types of objects, two kinds of events, two modes and three attributes. Lowe settled with this structure as he believes that this four-category ontology provides "a uniquely satisfactory metaphysical foundation for the natural sciences" (Lowe, 2007 Page 16).

TAKE IN FIGURE 1. ABOUT HERE

The mapping sentence for Lowe's ontology in figure 1 offers a transparent modeling of Lowe's conceptions of the basics of existence. Uniquely, the mapping sentence demonstrates not only the ontology's structure but also the interplay (or mereological

⁸ Aristotle's categories are: 1, Substance (οὐσία); 2, Quantity (ποσόν); 3, Quality (ποιόν); 4, Relation (πρός); 5, Place (πού); 6, Time (ότε); 7, Being-in-a-position (κεῖθαι); 8, Having (ἔχειν); 9, Action (ποιεῖν); 10, Affection (πάσχειν). From these Categories I developed a mapping sentence for Aristotle's categorial system (Hackett, 2014).

⁹ Lowe's ontology has appeared in several slightly varied iterations all of which Lowe portrayed as an ontological square.

arrangement) of Lowe's ontology. When Lowe's ontology is modeled in a mapping sentence the interplay of elements is stressed and by clearly explicating a possible mereology of elements the active role of the reader is also emphasized. Furthermore, the mapping sentence requires the researcher to consider the nature of the context of the evaluation and background features that may affect content.

A HERMENEUTICALLY CONSISTENT TEMPLATE

In this paper I am claiming that the mapping sentence is at the heart of traditional, philosophical and qualitative explorations employing a facet theory outlook in both exploratory and confirmatory research. The mapping sentence is the basis for investigations, structural hypothesis testing and theory generation and as a stand-alone research approach. Mapping sentences specify research domains allowing the definition of the domain's sub-aspects and sub-aspect interrelationships availing appreciation of the domain's content. To further illustrate a qualitative application of a mapping sentence in **Figure 2** I provide a mapping sentence of the theoretical content of this essay. This qualitative/philosophical mapping sentence demonstrates the hermeneutic consistency of understanding that arises from non-numerical research that is organized through using a mapping sentence.

TAKE IN FIGURE 2 ABOUT HERE

In this mapping sentence the *range* facet delimits the substantive concern of the mapping sentence, which in this case is the extent to which a mapping sentence structured ontology can

The content of this paper, when read by person (x) understands facet theory to

embody a mapping sentence with:

ontology
(*facets*)
([*elements*]) and with the structure
(*background facets*)
(*range*)

between these ontological components being in terms of the: **mereology**
(*part-to-part*)
(*part-to-whole*)

relationships, and judges this to have: **range**
(*more*)
(*to*) hermeneutic consistency in relation
(*less*)

to the ontological domain.

FIGURE 2 | Mapping sentence for the hermeneutic consistency of a mapping sentence.

avail a hermeneutically consistent understanding of a content domain. Returning to the start of the mapping sentence, person (x) is taken to be any individual reading and understanding the mapping sentence. Continuing along the sentence, the combinatorial arrangements of the two content facets are determinants of the values observed in the range. In this sentence: the ontology facet specifies the content of the mapping sentence ontology to be—*facets* (with sub-divisions of facet elements); *background* (which lists background characteristics of the instantiation of the ontology); *range* which specifies the epistemological/characteristics of the observations that constitute the mapping sentence's logic. Thus, the mereology facet characterizes the nature of the relationships that are extant within the mapping sentence ontology as being either *part-to-part* (facet/facet element-to-facet/facet element) or *part-to-whole* (facet/facet element-to-mapping sentence).

CONCLUSIONS

I commenced by proposing that understanding a content domain may result from sub-dividing the domain into relevant categories. I then noted how facet theory has achieved a category-based epistemological exposition of many research areas under a quantitative research rubric. In this paper I

have provided support for claims regarding the potential of qualitative or philosophical research that is undertaken within a facet theory framework. I have claimed utility for the use of a mapping sentence as a purely philosophical outlook when attempting to understand human experience by offering a mapping sentence as a philosophically coherent approach to understanding Lowe's ontology and as a tool to investigate the hermeneutical consistency of research.

It is my contention that the hermeneutic consistency of a structural ontology is a product of a structural representation's ontological components and the mereological relationships between these ontological units: the mapping sentence seminally allows for the depiction of such structure. Finally, I claim facet theory and mapping sentences form a precise though flexible framework for the designing research and writing within philosophical and qualitative psychological research¹⁰.

AUTHOR CONTRIBUTIONS

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

¹⁰I would like to thank the associate editor of Frontiers for suggesting that the facet theory approach may be used in future research to thematically investigate the philosophy of science and especially the structure of scientific theories.

REFERENCES

- Aristotle (1975). *Aristotle's Categories and de Interpretatione*. Oxford: Oxford University Press.
- Borg, I. (1978). "Some basic concepts in facet theory," in *Geometric Representation of Relational Data*, ed J. Lingoes (Ann Arbor, MI: Mathesis Press).
- Borg, I., and Shye, S. (1995). *Facet Theory: Form and Content (Advanced Quantitative Techniques in the Social Sciences)*. Thousand Oaks, CA: Sage Publications, Inc.
- Canter, D. (ed.). (1985a). *Facet Theory: Approaches to Social Research*. New York, NY: Springer Verlag.
- Canter, D. (1985b). "How to be a facet researcher," in *Facet Theory: Approaches to Social Research*, ed D. Canter (New York, NY: Springer Verlag), 265–276.
- Chisholm, R. M. (1996). *A Realistic Theory of Categories: An Essay on Ontology*. Cambridge: Cambridge University Press.
- Cotnoir, A. J., and Baxter, D. L. M. (eds.). (2014). *Composition as Identity*. Oxford: Oxford University Press.
- Gadamer, H. G. (2013). *Truth and Method*. New York, NY: Bloomsbury Academic.

- Gibbs, R. W., and Colston, H. L. (eds.). (2007). *Irony in Language and Thought: A Cognitive Science Reader*. New York, NY: Routledge.
- Gibbs, R. W., and Colston, H. L. (2012). *Interpreting Figurative Meaning*. Cambridge: Cambridge University Press.
- Guttman, L. (1947). "Scale and intensity analysis for attitude, opinion and achievement in new methods in applied psychology," in *Proceedings of the Maryland Conference on Military Contributions to Methodology in Applied Psychology held at the University of Maryland, November 27–28, 1945, under the auspices of the Military Division of the American Psychological Association*, ed G. A. Kelly (College Park, MD: University of Maryland).
- Guttman, L. (1959). "Introduction to facet design and analysis," in *Proceedings of the Fifteenth International Congress of Psychology, Brussels–1957*. Amsterdam: North Holland, 130–132.
- Hackett, P. M. W. (2013). *Fine Art and Perceptual Neuroscience: Field of Vision and the Painted Grid*. New York, NY: Psychology Press.
- Hackett, P. M. W. (2014). *Facet Theory and the Mapping Sentence: Evolving Philosophy, Use and Application*. Basingstoke: Palgrave.
- Harte, V. (2002). *Plato on Parts and Wholes: The Metaphysics of Structure*. Oxford: Oxford University Press.
- Heidegger, M. (1962). *Being and Time*. New York, NY: Harper Perennial Modern Classics.
- Kelly, G. A. (2013). *A Theory of Personality: The Psychology of Personal Constructs*. New York, NY: W. W. Norton and Company.
- Koval, E., and Hackett, P. M.W. (2015). "Hermeneutic consistency, structured ontology and mereology as embodied," in *Facet Theory and the Mapping Sentence, paper presented at the fourteenth International Facet Theory Conference*, Fordham University, New York, NY.
- Levy, S. (ed.). (1994). *Louis Guttman on Theory and Methodology: Selected Writings, Dartmouth Benchmark Series*. Aldershot: Dartmouth.
- Lowe, E. G. (2007). *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press.
- Osborne, G. R. (2007). *The Hermeneutical Spiral: A Comprehensive Introduction to Biblical Interpretation*. Westmont, IL: IVP Academic.
- Piaget, J., and Inhelder, B. (1969). *The Psychology of the Child*. New York, NY: Basic Books.
- Porter, S. E., and Robinson, J. C. (2011). *Hermeneutics: An Introduction to Interpretive Theory*. Grand Rapids, MI: Wm. B. Eerdmans Publishing Co.
- Shye, S. (1978). *Theory Construction and Data Analysis in the Behavioral Sciences*. San Francisco, CA: Jossey-Bass.
- Shye, S., and Amar, R. (1985). "Partial-order scalogram analysis by base coordinates and lattice mapping of the items by their scalogram roles," in *Facet Theory: Approaches to Social Research*, ed D. Canter (New York, NY: Springer Verlag), 277–298.
- Shye, S., and Elizur, D. (1994). *Introduction to Facet Theory: Content Design and Intrinsic Data Analysis in Behavioral Research (Applied Social Research Methods)*. Thousand Oaks, CA: Sage Publications, Inc.
- Turvey, M. T., and Moreno, M. A. (2006). Physical metaphors for the mental lexicon. *Ment. Lex.* 1, 7–33. doi: 10.1075/ml.1.1.04tur

Conflict of Interest Statement: 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.

Copyright © 2016 Hackett. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Metacognitive Control of Categorical Neurobehavioral Decision Systems

Gordon R. Foxall*

Cardiff University, Cardiff, UK

OPEN ACCESS

Edited by:

Paul M. W. Hackett,
Oxford University, UK

Reviewed by:

Ekrem Dere,
Pierre and Marie Curie University –
Paris 6, France
Michael F. Dorsey,
Endicott College, USA

*Correspondence:

Gordon R. Foxall
foxall@cf.ac.uk

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 29 November 2015

Accepted: 29 January 2016

Published: 19 February 2016

Citation:

Foxall GR (2016) Metacognitive
Control of Categorical Neurobehavioral
Decision Systems.
Front. Psychol. 7:170.
doi: 10.3389/fpsyg.2016.00170

The competing neuro-behavioral decision systems (CNDS) model proposes that the degree to which an individual discounts the future is a function of the relative hyperactivity of an impulsive system based on the limbic and paralimbic brain regions and the relative hypoactivity of an executive system based in prefrontal cortex (PFC). The model depicts the relationship between these categorical systems in terms of the antipodal neurophysiological, behavioral, and decision (cognitive) functions that engender normal and addictive responding. However, a case may be made for construing several components of the impulsive and executive systems depicted in the model as categories (elements) of additional systems that are concerned with the metacognitive control of behavior. Hence, this paper proposes a category-based structure for understanding the effects on behavior of CNDS, which includes not only the impulsive and executive systems of the basic model but a superordinate level of reflective or rational decision-making. Following recent developments in the modeling of cognitive control which contrasts Type 1 (rapid, autonomous, parallel) processing with Type 2 (slower, computationally demanding, sequential) processing, the proposed model incorporates an arena in which the potentially conflicting imperatives of impulsive and executive systems are examined and from which a more appropriate behavioral response than impulsive choice emerges. This configuration suggests a forum in which the interaction of picoeconomic interests, which provide a cognitive dimension for CNDS, can be conceptualized. This proposition is examined in light of the resolution of conflict by means of bundling.

Keywords: competing neuro-behavioral decision systems, CNDS model, dual- and triple-process models, metacognitive control, temporal discounting, picoeconomics, bundling, categorical system

INTRODUCTION

“...akrasia in rational beings is as common as wine in France”

(Searle, 2001, p. 10)

As I scan my daily newspaper over breakfast, I note the television programs scheduled for the evening. It is easy at so early an hour to vow that I will under no circumstances allow myself to watch what is on offer. Tidying my sock drawer or deadheading the roses seems a more valuable use of my time, and serious reading or writing infinitely preferable. Comes the evening, however, the opportunity to relax and be passively entertained wins out. Am I speaking here of “myself”

as one person whose preferences are reversible simply with the passage of time, or of two separate categories of agents warring to get the upper hand? If the latter, how are these categories related and how do they influence each another? Perhaps there is some superordinate level of decision-making that arbitrates between them; or perhaps they reflect no more than differing histories of operant reinforcement.

The problem of preferences that change with time lies at the heart of many comparatively trivial daily decisions. What seems perfectly reasonable when we begin becomes absurd simply because other options, ultimately less valuable than the initial longer-term objective, have become immediately accessible (Rachlin, 2000a). Apparently for that reason alone, these choices that may be categorically classified as temporarily short-term assume an irresistible level of attractiveness: the result may be excessive consumption leading to obesity or procrastination leading to failure to achieve (Ainslie, 2010). The problem, akrasia or weakness of will, occurs also in the more serious contexts of substance abuse and problem gambling, even when the individuals concerned know from experience the deleterious outcomes of their behavior and have the “best intentions” of changing it. Once again, the questions of the apparently “divided self” or “multiple selves” arise (Elster, 1987; Ainslie, 2001; Ross et al., 2008). It is interesting that Searle (2001) speaks of akrasia

as a common characteristic of *rational* beings: in what sense are we to understand the rationality that underlies such self-defeating behavior?

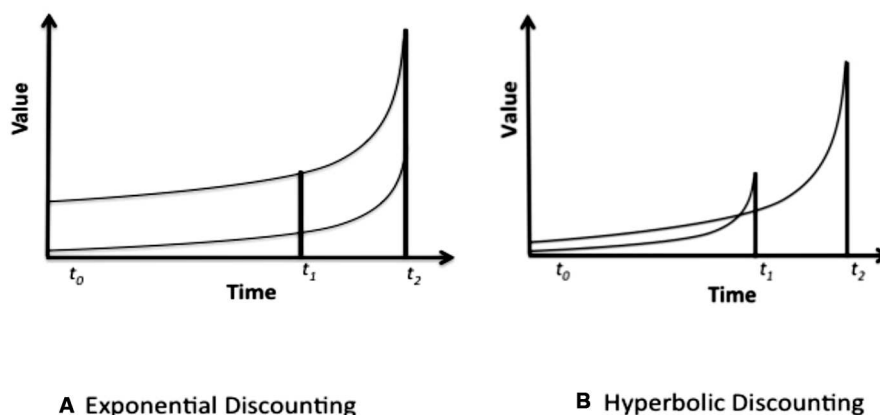
The initially self-controlled and subsequently impulsive behaviors involved in preference reversal can be traced to neurophysiological and cognitive bases of competing decision systems. Jentsch and Taylor (1999) propose that drug seeking stems from amygdala-based reward processing that intensifies the incentive value of potentially addictive substances, accompanied by the weakened capacity of frontal cortical processes to impede such behavior. Bechara (2005) similarly argues that the extent of an individual's willpower to resist drugs depends on the relationship between an impulsive system based on the amygdala which indicates the immediate outcomes of behavior and a reflective system, based on ventromedial prefrontal cortex (vmPFC) which indicates delayed outcomes (Bickel and Yi, 2010). This relationship has been most comprehensively described, however, in the competing neuro-behavioral decision systems (CNDS) model which hypothesizes that two competing neural systems, respectively, exert excitatory or inhibitory control over potentially addictive behavior (Bickel et al., 2013).

The CNDS model proposes that imbalance between an individual's “impulsive” and “executive” categorical systems

BOX 1 | Temporal discounting and preference reversal.

A reward that is to be received at some time in the future – say, \$100 in a year's time – does not seem right now to be worth waiting that long for unless there is some extra bonus attached to it. If someone owes me this amount and offers to let me have in 12 months, I am inclined to say that I will require, say, \$110 at that time. Reward for which one has to wait are devalued or discounted. We say that temporal discounting is concerned with the *current subjective value* of a reward that will be received in the future, i.e., the value of that future reward rated in the present moment. Financial professionals discount exponentially, i.e., at a constant rate regardless of the time elapsed. Their behavior can be expressed as $V_t = A_t e^{-kD_t}$ where V_t is the present value of a delayed reward, A_t the amount of delayed reward, k a constant proportional to the degree of temporal discounting, D_t the delay of the reward, and e the base of natural logarithms. Because this behavior is based on a constant rate of discounting, a larger, later reward (the LLR, available at t_2) *always* has a value greater than that of a smaller reward available sooner (the SSR, available at t_1). This is shown in the first segment (A) of the figure, where the two lines, representing the relative values of the reward, never cross.

Often, however, human behavior is marked by a style of discounting in which the value of a reward changes radically as the time remaining before it becomes available is reduced. While the LLR is preferred at t_0 , indicated by the initially higher line in segment (B) of the figure, just prior to t_1 , when the SSR will become available, its value markedly increases, the curves cross, and the individual opts for the poorer reward. This form of temporal discounting and the preference reversal it involves is described by a hyperbolic function: $V_d = A/(1 + kD)$ in which V_d is the discounted value of a reward of a particular magnitude or amount, A , received after a delay, D (Mazur, 1987; Madden and Bickel, 2010). Rate of discounting varies with the amount of delay (Ainslie, 1992, 2001; Rachlin, 2000a; Rick and Loewenstein, 2008).



influences his/her rate of temporal discounting (see **Box 1**).¹ Hyperactivity of the impulsive system, based on limbic and paralimbic brain regions, coupled with hypoactivity of the executive system, based in the prefrontal cortex (PFC), results in a tendency to discount the future steeply and to engage in addictive behavior (Bickel and Yi, 2008). A major premise of the CNDS model is, therefore, that the impulsive and executive systems must be in some respects antipodal categories and yet contribute in a complementary manner to the determination of the individual's temporal discounting behavior and valuation of currently and potentially available reinforcers. These have been concerns of the CNDS model's authors who also emphasize the role of metacognition (i.e., "cognition about cognition" or "thought about thought") in the regulation of inter-system connectivity (Jarmolowicz et al., 2013). In attempting to clarify further the factors responsible for the achievement of relative balance between the impulsive and executive systems, this paper explores further the antipodality of the model's categorical component decision systems and, in particular, the nature and role of metacognition in their relationships.

The CNDS model has two important implications for the resolution of the question of multiple selves. First, by incorporating *cognitive* or decision-making contributors to the extent of an individual's temporal discounting tendency, it links to the capacity to regulate behavior through goal setting and maintenance, social cognition (understanding why others behave as they do), and insight (taking one's own imperfections into account in judging behavioral outcomes). Second, the model's incorporation of operant behavioral economics and neuroeconomics (Bickel et al., 2007, 2011, 2012a) facilitates its integration with the economic reasoning which underlies another significant contribution to the explanation of multiple selves and their interaction, namely *picoeconomics* (Ainslie, 1992, 2001; Ross, 2012; Foxall, 2014a,b).

Ainslie (1992) speaks of the problem of akrasia by reference to separate *interests* that are in conflict: one concerned with our gaining long term benefit such as engaging in productive work, the other with short-term pleasures like undemanding amusement. One's experience as the locus of this clash of interests is often marked by a sturdy resolve to undertake the more rewarding activity, followed by a lapse into the other just as it becomes available, followed by regret, further resolution and perhaps inevitable relapse. This cycle is characteristic of addiction but it also marks many everyday switches of preference involved in less extreme behavior. What is so preferable when we make our plans is edged out by an alternative that is initially unthinkable but of immense value as it arrives in sight. Even though we know full well that the activity which we were determined to undertake when we set out will bring greater benefit, the fact that it is delayed while the less beneficial can be obtained immediately raises the value of the latter sharply till it exceeds the current

worth of the other (see **Box 1**). An intriguing facet of Ainslie's approach is the possibility that, by "bundling" together (**Figure 1**) the combined benefits of a series of later-appearing reward and comparing these *in toto* with the immediate benefit of a current less valuable choice, it is possible to overcome the temptation to make a sub-optimal decision (i.e., to exercise "willpower" or "self-control"). Hence, picoeconomics has implications for the role of cognition and metacognition in relationships between neuro-behavioral decision systems and the place of agency in understanding their interaction.

Some of these implications are taken up by Ross (2009) who defines the situation in economic terms by reference to two reward available at different times such that *a* is, for example, taking a short vacation starting in a week [*t*₁], and *b* is, for instance, starting a 2-years course of study for a higher degree, [*t*₂]. Looking well into the future, the person's utility function indicates that *b* is preferable to *a*. At this point, the person discounts the future rather gently. However, as the time for the vacation to occur becomes closer, the person's utility function indicates a preference for *a* over *b*. Ross (2012) models the various picoeconomic interests in two ways depending on whether these interests are conceived as acting synchronously or diachronically. In the first case, they may be seen as subagents that have either conflicting utility functions or divergent time preferences. Agents with conflicting utility functions may be modeled in terms of a Nash equilibrium game among these agents. Modeling the behavior of subagents whose time preferences diverge adverts to the sub-personal level of neurophysiology in which a hyperbolic time preference emerges from "competition between steeply exponentially discounting

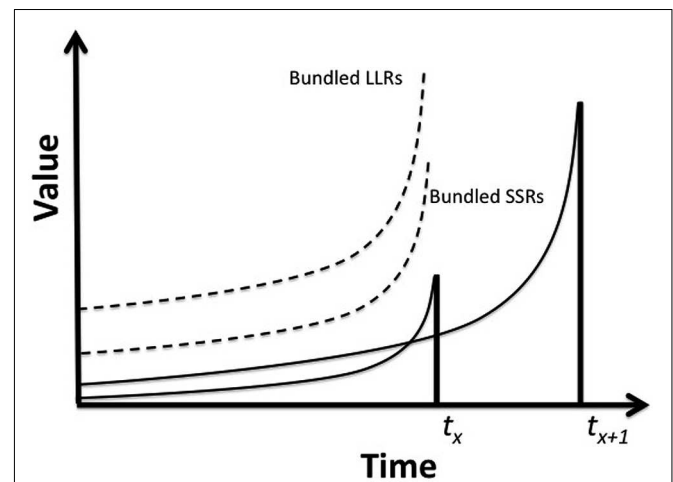


FIGURE 1 | The Principle of Bundling. The solid lines represent any of the individual members of the stream of paired SSR/LLR choices that will be available to the individual over time. (Hence, t_x is the time of any occurrence of an SSR which is paired with an LLR that occurs at t_{x+1} , and we are assuming a sequence of such SSR/LLR pairings over time). The dashed lines represent the individual's imagined aggregation of these reward if they were all brought forward to a point just prior to the appearance of the first SSR. In this case, LLR will always exceed SSR and a decision to select it exclusively on subsequent occasions can be more easily made.

¹Behavior analysts have conducted a large volume of research on temporal discounting and the matching phenomena which underlie it (Ferster and Skinner, 1957; Herrnstein, 1997). For reviews of this research, see Madden and Bickel (2010), Dallery and Soto (2012), Grace and Hucks (2012), Hursh et al. (2012), Jacobs et al. (2012), Foxall (2016). See also Foxall (2015) and the special issue on Operant Behavioral Economics this introduces.

'limbic' regions and more patient (less steeply exponentially discounting) 'cognitive' regions" (Ross, 2012, p. 720). Like the CNDS model, this piceoeconomic portrayal depends heavily on the findings a key experiment in neuroeconomics based on fMRI scans of humans choosing between SSR and LLR (McClure et al., 2004). In the process of scrutinizing immediate reward, participants activated brain regions that involve emotion, namely medial orbitofrontal cortex, medial prefrontal cortex/pregenual cingulate cortex, and ventral striatum. However, while examining longer-term payoffs, they activated areas of the lateral PFC (implicated in higher cognitive functioning), and part of the parietal cortex related to quantitative reasoning. In his modeling of piceoeconomic conflict in terms of diachronically appearing multiple selves, Ross (2012) speculates briefly about the cognitive demands of such a portrayal: each subagent is portrayed as temporarily in control of the person's behavior, with its own utility function and incomplete knowledge of the other, though its utility is constrained by the investments made by earlier-appearing agent(s).

In seeking to clarify the issue of multiple selves, this paper draws on recent investigations of antipodality between the categorical components of the impulsive and executive systems (Bickel et al., 2012b). This work is invaluable for identifying the elements of a theory of behavior that would account for both normal and excessive (addictive) consumption of substances such as alcohol and other drugs and activities such as gambling. Importantly, it demonstrates which elements of the impulsive system are antipodal to elements of the executive system (and can, therefore, be properly considered categorical components of these antithetical tendencies), as well as those which play a broader role in the execution of appropriate behaviors. Prominent among the latter are what the CNDS model identifies as metacognition and the goal-directed regulation of behavior (Jarmolowicz et al., 2013).

The paper builds on the results of this work to propose a model of cognitive functioning in addiction that places the impulsive and executive systems in a framework consistent with recent developments in multi-process theories of cognition (Stanovich and West, 2000; Stanovich, 2009, 2011). It thereby incorporates a broader domain of theory on the cognitive control of behavior which acknowledges a long-standing division of thinking into that which is rapid and intuitive as opposed to that which is slow and deliberative (Evans and Stanovich, 2013). This dichotomy of categories is similar to that which marks the distinction between impulsive and executive systems and is consistent with the account of behavioral control that the examination of the CNDS model in terms of antipodality reveals. The advantage of such a framework is that it allows for a forum within which the competing demands of the impulsive and executive systems interact so that conflict is resolved and behavior that generates more acceptable long-term consequences is selected over short-term expediency.

The central focus is, therefore, on the structure of the CNDS model and, in particular, its incorporation of metacognitive control of behavior. In the next sections, the CNDS model is described in greater detail and the implications of antipodality for the construal of decision systems is discussed. The questions

of how metacognition is depicted in the model and a potential tripartite model are the foci of the following sections. Finally, the implications of the analysis for understanding the multiplicity of selves involved in the decision process are discussed by critically examining the emergent framework in terms of piceoeconomic bundling behavior.

THE COMPETING NEURO-BEHAVIORAL DECISION SYSTEMS MODEL

The Neurophysiological Dimension

At the neurophysiological level, the CNDS hypothesis (Bickel et al., 2012a, 2013) assumes that normal and addictive behaviors reflect the balance between the relative hyperactivity of the limbic and paralimbic systems that are differentially implicated in emotional responding and the relative hypoactivity of prefrontal cortical areas that are differentially implicated in judgment, planning, and other cognitive activities. Hence, the degree of addictiveness exhibited in behavior reflects the balance of activity in brain regions, the first of which, the *impulsive system*, based on the amygdala and ventral striatum, involves the distribution of dopamine during reinforcement learning, while the second, the *executive system*, residing in the PFC, is implicated in the evaluation of reward and their outcomes. The competing systems that comprise the model are more broadly based than these neurophysiological regions, embracing in addition behavioral and cognitive components which justifies their being called neuro-behavioral decision systems. (Bechara, 2005, nominates these the impulsive system and the reflective system, respectively; see also Verdejo-Garcia and Bechara, 2009).

The Impulsive System

The impulsive system incorporates the amygdala and ventral striatum, a midbrain region concerned with the valence of immediate results of action, and is liable to become hyperactive as a result of "exaggerated processing of the incentive value of substance-related cues" (Bechara, 2005, p. 1459). Drug-induced behaviors correlate with enhanced response in this region when the amygdala displays increased sensitization to reward (London et al., 2000; Bickel and Yi, 2008). The receipt of positive reinforcers of all varieties causes the release of dopamine in the nucleus accumbens. This is true of both utilitarian reinforcers such as drugs of abuse, and the receipt of informational reinforcers such as social reward or self-esteem (Foxall, 2011). It is also the case for the receipt of money which has both utilitarian and informational aspects. In the case of a drug of abuse, such brain reward is acute. The effect of the drug in inculcating LTP at specific synapses is recorded in the hippocampus as the result of experience (memory).¹ In the amygdala is involved in the

¹Recent research indicates a far more complicated picture than this. LTP is currently understood as no more than a possible molecular mechanism of learning (see, for instance, Migaud et al., 1998; Uetani et al., 2000). The role of LTD, which has been correlated with learning, is of particular importance. Present knowledge on synaptic plasticity and learning performance, incorporating the learning of drugs' capacity to reward, indicates that LTP and LTD tend toward an optimal balance which may influence memory performance. Memory formation may also

creation of a learned (conditioned) response to the stimuli that accompany the use of the drug. These accompanying stimuli might take the form of informational (social) reinforcers and discriminative stimuli. (For discussion of these points, see, *inter alia*, McGaugh, 2004; Phelps, 2006; Gruber and McDonald, 2012.)

The resulting focus of research has been on the mesolimbic dopaminergic system and other brain regions such as the amygdala and ventral striatum involved in emotional responses. But there is recent evidence that the insula is important because of its relation to conscious craving for drugs (Naqvi and Bechara, 2008). This role has been revealed by correlation-based fMRI studies which show the increased activity of the insula during self-reported urges to ingest drugs. Such activity is related to the emergence of the secondary reinforcers which tie drug use to specific behavioral and contextual factors and to the cognitive drivers of drug use. "Over time, as addiction increases, stimuli within the environment that are associated with drug use become powerful incentives, initiating both automatic (i.e., implicit) motivational processes that drive ongoing drug use and relapse in addition to conscious (i.e., explicit) feelings of urge to take drugs" (Naqvi and Bechara, 2008, p. 61; see also Naqvi et al., 2006). The ritualistic practices involved in the preparation of drugs, associated with specific places, apparatus, packages, lighters, and so on, thus become sources of the pleasure that reinforces not only those activities but the consummatory acts of drug ingestion. These processes, which elicit specific memories of encounters with the contexts and the drugs, are also responsible for differences in the subjective experience of urges for various drugs be they cigarettes, cocaine or gambling. By ensuring that the individual keeps particular goals "in mind," the insula is also involved in (thwarting) the executive functions that might overcome drug urges (cf. Tiffany, 1999). The learning process includes the development of neural plasticity through DA-priming with respect to the impending chain of appetitive events; Naqvi and Bechara (2008) propose that this DA-dependency invokes activity in the insula and associated regions such as the VMPFC and amygdala. The plasticity involves the establishment of representations of the interoceptive outcomes of using drugs and thus engender relapse even after long periods of non-use.

The Executive System

The executive system, which includes the PFC is normally associated with the executive functions of planning and foresight (Barkley, 1997, 2012), and is hypothesized to become hypoactive in the event of addiction. In the absence of its moderating function, effects of the hyperactive dopaminergic reward pathway are exacerbated, leading to an imbalance which is implicated in the enactment of dysfunctional behavior. The behavioral concomitant of these neurophysiological processes is observable in the rate at which individuals discount the value of future reward in favor of more-immediately appearing reinforcers (Bickel and Yi, 2008). In the context of addiction, the CNDS hypothesis posits that drug seeking results from "amplified

incentive value bestowed on drugs and drug-related cues (via reward processing by the amygdala) and impaired ability to inhibit behavior (due to frontal cortical dysfunction)" (Bickel and Yi, 2010, p. 2).

Analysis of the neurobiological pathway proposed to account for the acquisition by the PFC of the capacity to control the higher-level cognitive functions involved in the regulation of behavior in the face of environmental programming reveals a two-stage process (Miller and Wallis, 2009). The first stage is the impingement of signals generated via reinforcement learning on the PFC circuitry: reinforced operant behavior is accompanied by the production of signals that associate PFC functioning with aspects of the stimulus field (the setting in which the behavior takes place), the nature of the behavioral response enacted, and the reinforcing and punishing consequences that are its outcomes. Repeated responding in these circumstances is capable of generating strong PFC representations of the contingencies of reinforcement that maintain such behavior. The second stage in the argument is to account for these signals and the actions of dopaminergic neurons of the midbrain. In the course of learning through the repeated performance of behavioral responses, reinforcers initially activate the dopaminergic neurons themselves, but subsequently the stimuli that predict the reinforcers, rather than the reinforcers themselves, come to activate the dopaminergic neurons. Should an expected reward not appear, the rate of firing of the dopaminergic neurons is reduced. The discrepancy between the expectation of reinforcement and its non-appearance, coded by the dopaminergic neurons' activity is known as the *reward prediction error* and is instrumental in the organism's subsequent ability to direct its actions more effectively toward the achievement of reinforcement (Miller and Wallis, 2009, p. 103–104; see also Foxall, 2014b).

The fundamental assumptions that reinforcement is coded by dopaminergic neurons (Schultz, 1992; Robbins and Everitt, 2002) and that RPEs are also reflected in the firing rates of dopaminergic neurons (Schultz et al., 1997) ground the relationship between neoclassical micro-economics and neuroscience on which neuroeconomics rests (Glimcher, 2011). For present purposes, they serve to integrate operant psychology with these disciplines by promoting a causal connection among reinforcement, neuronal activity, and behavior (Schultz and Dickinson, 2000; Schultz, 2010).

The Behavioral Dimension

The operations of these systems combine to generate behavior that reflects the individual's valuation of future events, his/her degree of "temporal discounting." Hyperbolic temporal discounting is the procedure in which the later-occurring of two reward is diminished in an individual's subjective estimation even though it is the larger, with the result that the more immediate reward is selected in preference despite its being by definition the smaller of the two. This "impulsive" behavior is described by the hyperbolic discounting function

$$V_d = A/1 + kD \quad (1)$$

be subserved by hippocampal network oscillations in the theta, gamma and high-frequency range and the activity of place-, grid-, time-, and reward-related cells in the brain (Pan et al., 2005; Eichenbaum, 2014; Moser et al., 2015; Sanders et al., 2015).

where V_d is the discounted value of a reward of a particular magnitude or amount, A , received after a delay, D (Mazur, 1987; Madden and Bickel, 2010). The k parameter indicates the extent to which the value of the LLR diminishes compared to that of the SSR over time (Stein and Madden, 2013). The major behavioral characteristic of choice described hyperbolically is that the individual is likely to reverse preferences as time advances, an observation which is highly relevant to the extreme drug-use and gambling already mentioned, the making of resolutions to change, and the yielding to temptation that may follow. Behavior that discounts the future is of central importance to the CNDS model insofar as temporal discounting is an index of the extent to which behavior is under the control of the tendency toward disinhibited impulsivity (the selection of an SSR rather than an LLR) as opposed to the inhibiting influence of the executive functions which results in the choice of LLR over SSR ("self-controlled" behavior) (Bickel and Marsch, 2000; Bickel and Yi, 2008; Barkley, 2012).

It is reasonable to inquire how the valuation an individual attaches to the outcomes of his/her future behavior should be understood. The CNDS model argues that the neurophysiological tendencies of the impulsive and executive systems eventuate in an individual's degree of temporal discounting behavior which is explicable in operant terms that translate readily into economic considerations (Bickel et al., 2007, 2011). Bickel et al. (2012a) argue that addiction can be conceptualized as an outcome of "reinforcer pathologies" that can be analyzed in terms of behavioral economics, specifically the inelasticity of demand (manifesting in a willingness to pay an extraordinarily high price for a drug reward) and extremely steep discounting of the future (manifesting as over-valuation of an immediately available reward). These elements which reveal an excessive valuation of one reinforcer in comparison with other available reward and impulsivity, respectively, are consistent with the pattern of behavior found in addicts who may accordingly be defined as "people for whom the transient benefits of the addictive behavior persistently outweigh the significant short- and long-term costs of these choices" (Bickel et al., 2012a, p. 334–335). The portrayal of these benefits in terms of positively- and negatively reinforced behaviors is confirmed by the neurophysiological analysis of addiction which depicts addicts' initial drug administration as determined by the pleasures this confers and their later drug use as a means of avoiding or escaping from deleterious consequences such as withdrawal symptoms (Koob, 2013). Some aversive consequences cannot be avoided by further drug administration, however; the social isolation and damage to health that often result from persistent addiction are examples of the punishing outcomes of such behavior (Rachlin, 2000b; Foxall and Sigurdsson, 2011).

The Need for a Cognitive Dimension

There are several reasons for thinking more formally about the place and function of a cognitive dimension within neurophysiologically based models of decision-making. The principal reason in the current context stems from the fact that the "valuation" involved in temporal discounting is a

mental construct which requires explanation in terms of cognitive representation and evaluation. This, in turn, raises the concern that the present dual process structure of the model may be inadequate to the task of accounting for the metacognitive processes involved in the exercise of self-control. In the discussion that follows, the distinction between the sub-personal level of exposition, that concerning brains and neuronal activity, and the personal level of exposition, that which involves intentionality (e.g., desires, beliefs, emotions, and perceptions) and behavior (Dennett, 1969), is not only of primary importance but inviolate (Foxall, 2007).

ANTIPODALITY

Bases of Antipodality

The CNDS model is an example of a dual process theory, i.e., one that builds on a substantial volume of social scientific argument that human cognition is characterized by two categorical styles of processing (Frankish and Evans, 2009). Type 1 processing is *autonomous*: its execution is *rapid* and *mandatory*, economizes on central processing capacity and higher-level control systems, and employs *parallel* processing so that it avoids interfering with other cognitive operations. These characteristics illustrate the *computational ease* that makes Type 1 processing the default processing mode: unless it is overridden, it will automatically generate a category of responses to environmental conditions. Type 1 processing includes the regulation of behavior by the emotions, encapsulated modules that solve adaptive problems, implicit learning processes and the automatic firing of overlearned associations. By comparison with the Type 1 thinking that characterizes the *autonomous mind* (Stanovich, 2009), the second category of processing, type 2, is *slow* and makes *heavy computational demands*. It requires attention, which is costly, and is involved in conscious problem solving, eventuating in behavior that is directed toward achieving long-range consequences.

The distinction between the impulsive and the executive systems, and that between their respective styles of processing, suggests at least two categorical bases of evaluation and judgment that have opposing tendencies toward behavioral outcomes. This implies that at the systems can be construed as antithetical in important respects that can be related to their interaction to produce particular observed behavior patterns. These antipodal tendencies of the impulsive and executive systems ought ideally to indicate why behavioral imbalance would result from the hyperactivity of one system simultaneously with the hypoactivity of the other, a possibility which Bickel et al. (2012b) have investigated. These authors propose eight executive functions relevant to the CNDS model: Attention, Inhibitory control, Valuing future events, Cognitive behavioral flexibility, Working memory, Planning, Emotional activation and self-regulation, and Metacognitive processes. The first four are categorized as concerned with the *cross-temporal organization of behavior* (CTOB). *Emotional activation and self-regulation* (EASR), comprising two elements: Processing of emotional information and Initiating and maintaining

goal-related responding, and *metacognition* (MC) comprising two more: Social cognition (consisting in theory of mind, empathy) and Insight (or self-awareness). In addition, they propose as elements in the impulsive system, two trait impulsivities: Sensation seeking and Sensitivity to reward, and four state impulsivities: Behavioral disinhibition, Attention deficit impulsivity, Reflection impulsivity, and Impulsive choice (preference for SSR over LLR). The impulsive and executive systems, delineated in terms of the components identified by Bickel et al. (2012b), are shown as interacting systems in **Figure 2**.

Bickel et al. (2012b) assess antipodality by reference to four criteria: definition, measurement, overlapping of clinical populations, and commonality of neural substrates of the elements of impulsivity and executive function that comprise the CNDS model. It emerges from their analysis that the four state impulsivities are definitionally antipodal to four of the executive functions. Attentional deficit impulsivity and attention are clearly opposites, while the definitions of behavioral inhibition and behavioral disinhibition contain common characteristics that set them apart. In addition, reflection impulsivity tends toward the opposite of planning. Finally, the selection of SSR over LLR is antipodal to the capacity to value future outcomes. Note that the four executive functions identified as having antipodal impulsivities all belong to the CTOB grouping.

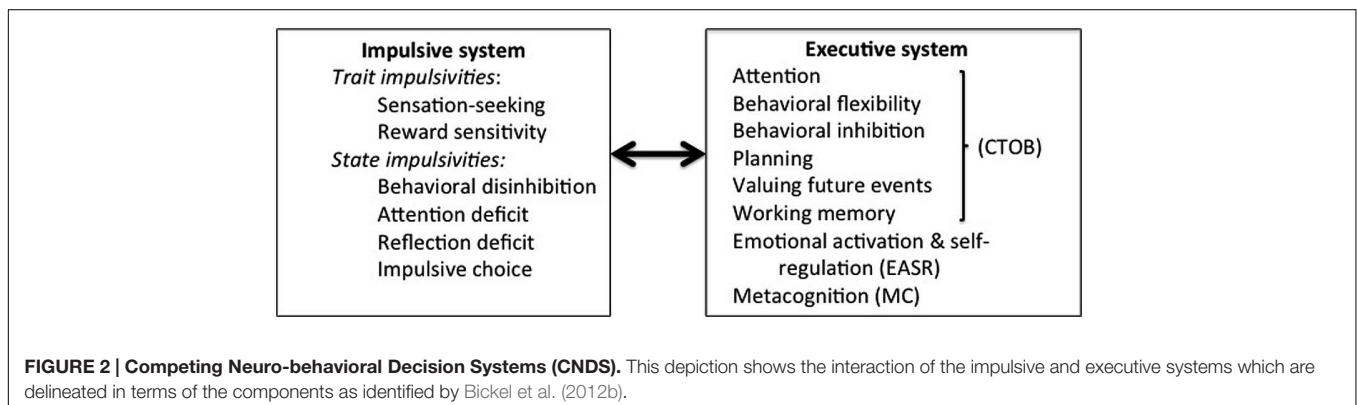
The analysis of antipodality also reveals a categorically distinct though coterminous measures of impulsive system and executive system items in the case of attention-deficit impulsivity and attention, and for behavioral inhibition and behavioral disinhibition. Reflective impulsivity and planning are less similar in their measurement. Finally, the delay discounting methods employed to measure impulsivity have recently come to be used in the measurement of executive functions. Note once more that these results establish CTOB as the seat of executive function which is the antipode to impulsivity.

The third source of evidence is the overlap of clinical populations whose members suffer from addiction and who show either hypoactive executive function or hyperactive impulsivity. Some substance users/abusers for instance demonstrate response inhibition deficits and excesses in behavioral disinhibition. When the substance is alcohol, this tends to be accompanied by lack of attention on the one hand and exaggerated attention deficit impulsivity on the other. Deficits in planning and

high levels of reflection impulsivity are found in users of amphetamines, cigarettes, and opiates. Finally, addicts to alcohol, cigarettes, cocaine, and heroin display steeper discounting of delayed reward more than controls do. Executive function deficits are also closely related to drug addiction.

Finally, in terms of the overlap in neural substrates of brain regions implicated variously in the functioning of the impulsive and executive systems, it is noteworthy that the insula and parts of the PFC are implicated in both behavioral disinhibition and behavioral inhibition. Moreover, since choice impulsivity and the valuation of future events are measured by means of delay discounting assignments they must recruit the same brain areas; they also cite the strongly emerging evidence that the limbic and paralimbic areas are implicated in immediate choice whilst parts of the PFC are implicated in the selection of delayed reward (and therefore with the valuation of future events). Again, it is noteworthy that all of these executive functions belong to the CTOB category. There is, however, little evidence of any overlap between the neural substrates of reflection impulsivity and planning other than the observation that individuals with lesions to the frontal cortices exhibit high reflection impulsivity which supports the view that DLPF and DMPFC are concerned with planning. There is also a paucity of evidence for any neural overlap for attention and attention deficit impulsivity. Nor is impulsivity antipodally related to working memory, EASR or MC even though impediments to these are found variously in addiction. Overall we may conclude that CTOB is antipodally related to the state impulsivities by evidence that they implicate similar neural substrates but that there is little evidence that the other elements employed in the categorization of executive functions shown in **Figure 2** are similarly related to impulsivity.

This does not constitute an original critique of the CNBDS model; indeed, the points made are all acknowledged by Bickel et al. (2012b). These authors specifically note that working memory answers no antipodal aspect of impulsivity and they draw attention to the lack of antipodal relationship between EASR and MC on one hand and impulsivity on the other. Such a relationship would be expected if EASR and MC belonged to executive functions. However, this examination of the findings suggests an alternative depiction of how the impulsive system and executive system are related.



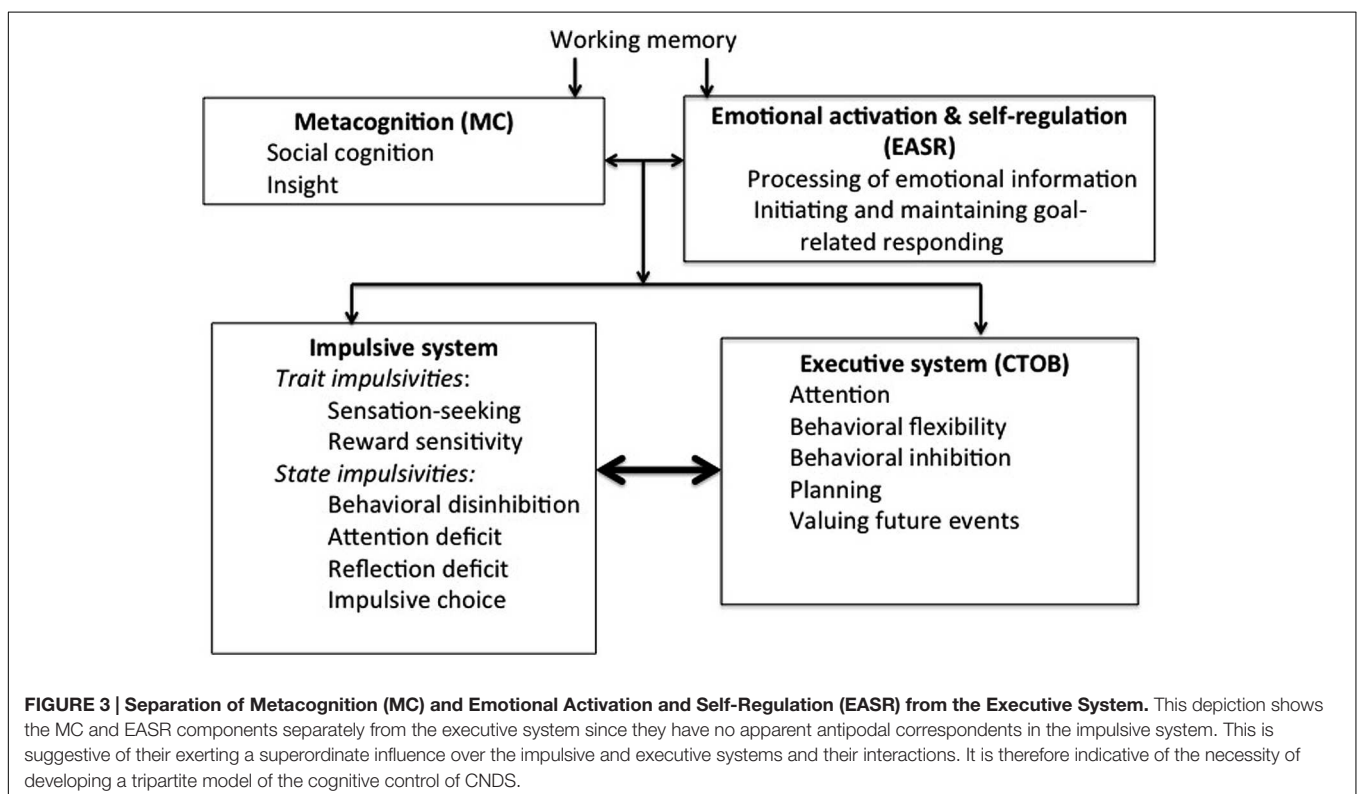
Locating Metacognition

Since the evidence for antipodality locates the competing classes of activity of the impulsive and executive systems firmly within neurophysiological bases, the task remains of placing the decision-making or cognitive elements of CNDS, currently located in the executive System (**Figure 2**), for which no evidence of corresponding and antipodal functions within the impulsive system has been adduced. There are in fact good reasons for separating metacognition (MC) and emotional activation and goal regulation (EASR), in which these cognitive or decision-making functions inhere, from the neurophysiological dimensions which are demonstrably antipodal. Because both MC and EASR involve thought or feeling about thought or feeling, I shall refer to them collectively as *metacognition*, though there may not exclusively fulfill this role in explaining human behavior. The justification for treating these cognitive variables as involved in the explanation of behavior is as follows.

If the CNDS model were conceived solely in terms of the neurophysiologically defined impulsive and executive systems that have been shown to be antipodal, then the individual's behavior manifested in a degree of temporal discounting peculiar to him/her would be the outcome of a sub-personal battle between opposing biological forces. Behavior would be starkly determined by innate neurophysiological capacities resulting from phylogenetic evolution, modified by a learning history that results in neural plasticity formed in a process of Hebbian or similar learning (Rolls, 2008). Behavior would be no more than contingency-shaped, determined in its totality by contingencies of natural selection and operant conditioning. However, this

would be to ignore the rule-governance of behavior, the possibility of an influence of reflective thought on responding. By including MC and EASR in their model, Bickel et al. (2012b) take this into account. Their inability to find or suggest functions of the impulsive system that are antipodal to these cognitive functions, which in any case repose uneasily among the other elements of the executive system depicted in **Figure 2**, argues for their separate consideration. The resulting re-conceptualization is shown in **Figure 3** in which MC and EASR are shown separately from the executive system which contains only those elements that are demonstrably antipodal to elements of the impulsive system. The impulsive system retains *pro tem* the trait impulsivities, sensation-seeking and reward sensitivity, that have no demonstrable correspondents in the executive system. **Figure 3** indicates also the reliance of MC and EASR on working memory.

The empirical outcome of the search for antipodality between impulsive and executive systems represented by Bickel et al.'s (2012b) research suggests the outcome shown in **Figure 3**. But there is a theoretical imperative for the proposal that metacognition occupy a superordinate position to the competing impulsive and executive systems. If the conflict of these systems is to be resolved by means of "cognition about cognition" or "thinking about thinking," it follows that such metacognitive activity must take place in a forum separate from the categorical systems themselves: how else could such activity decide between the interests these systems underpin? As a judge always sits apart from and acts independently of the advocates of plaintiff and defense, the realm of mediation, intercession, and arbitration in



decision-making cannot be incorporated within either the short-range interest that tends toward immediate gratification or the long-range interest that seeks a wider echelon of optimization.

A tri-PROCESS MODEL

Structure of the tri-Process Model

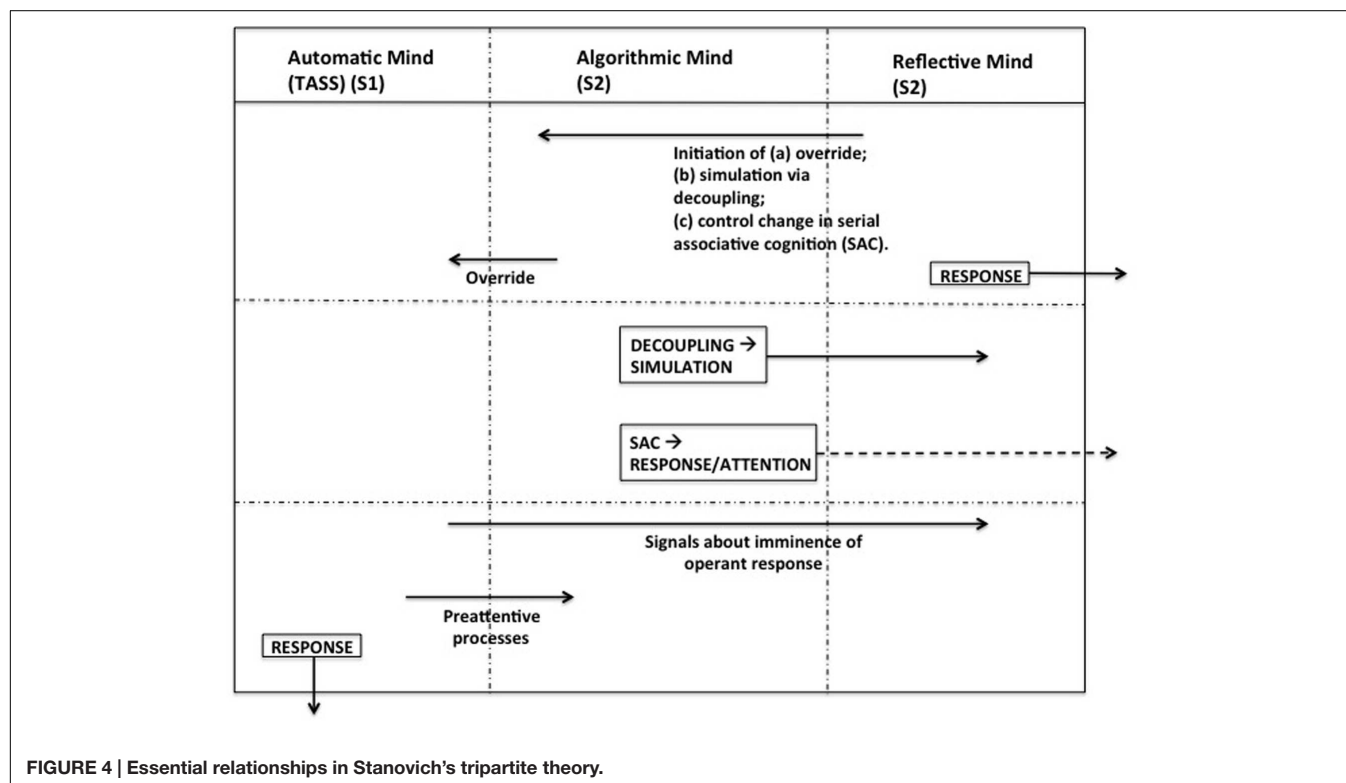
Recent theoretical development in multi-process theory suggest that metacognitive processes such as MC and EASR are most appropriately positioned as superordinate to the interactions of impulsive and executive systems as well as other systems that influence their interrelationship (Stanovich, 2009). Since no antipodal relationship between MC and EASR on one hand and components of the impulsive system on the other suggests that this possibility at least be considered as a means of understanding more fully the import of the CNDS model.

The similarity of the Type 1/Type 2 dichotomy to that of the Impulsive system/Executive system distinction is readily apparent. But any conclusions about the structure of the CNDS model in terms of these different styles of processing should take account of Stanovich's proposal for a tri-process theory (Figure 4). In proposing such a structure, Stanovich (2009) extends his earlier model both conceptually by adding a level of processing as well as by increasing the number of systems that comprise each level of processing. So, instead of a single Autonomous Mind, Stanovich (2009, p. 56) proposes "a set of systems in the brain that operate autonomously in response to their own triggering stimuli and are not under the control of the analytic processing system [i.e., System 2]". This heterogeneous

set, to which he refers as The Autonomous Set of Systems (TASSs), contains systems that are related in terms of their style of functioning (i.e., automaticity) rather than related by modularity. The proposed tri-process CNDS model incorporates two systems of Automatic Mind: the state-impulsive system comprising the state impulsivities and the trait-impulsive system comprising sensation-seeking and reward/reinforcement sensitivity (See also Kahneman, 2003, 2011; Shea et al., 2014).

Type 2 processing is divisible into two sorts of operation, each characteristic of a "kind of mind" (Dennett, 1996). The *Algorithmic Mind* involves individual differences in fluid intelligence, that which is measured by IQ tests, while the *Reflective Mind* involves individual differences in rational thinking dispositions. Rationality is broader than intelligence, requiring well-formulated desires (goals), highly calibrated beliefs and the ability to act on them in order to achieve the goals. It is, therefore, closely associated with the elements that Bickel et al. (2012b) position as components of the Executive system which, in their analysis, found no corresponding antipodal response in the Impulsive system.

The distinction between the two Type 2 systems posited by the tri-process theory rests on several functional differences. The key function of the Reflective Mind is the inauguration of the call to begin cognitive simulation or hypothetical reasoning. The key operation of the Algorithmic Mind in this is the decoupling it carries out. Decoupling is cognitively demanding, assisted by language which provides "the discrete representational medium that greatly enables hypotheticality to flourish as a culturally acquired mode of thought" (Stanovich, 2009, p. 63). Hypothetical thought requires



the representation of assumptions for instance and linguistic forms like conditionals readily allow this. Decoupling abilities differ in their recursiveness and complexity. Decoupling makes it possible to distance oneself from the representations so they can be reflected on and improved. Decoupling is therefore the key function of the algorithmic mind. It is clearly a System 2 operation in that its operation occurs serially and incurs high computational expense. The literature on executive function and working memory, he argues, supports the view that the main function of Algorithmic Mind is to achieve decoupling among representations while conducting cognitive simulation.

The cognitive control exerted by elements of the tri-processual model is justified as follows. System 2, representing analytic mind, contains two levels of functioning: the *algorithmic* level and the *reflective* level. TASS systems will function on a short range basis unless this is overridden by the Algorithmic System which give precedence to the long range goals of the analytical system. These latter reflect the goals of the person and the “epistemic thinking dispositions”. But these goals and dispositions must arise at a level superior to that of the Algorithmic System namely in the Reflective System “a level containing control states that regulate behavior at a high level of generality” (Stanovich, 2009, p. 57). This distinction of analytical systems gives rise to a tripartite system of cognitive processing. The Algorithmic Mind and the Reflective Mind share properties (such as capacity-limited serial processing) that distinguish them from the Automatic Mind (Stanovich, 2009, p. 58) But Algorithmic Mind and Reflective Mind can still be distinguished from one another, especially if we think in terms in relation to the impulsive executive systems. If these two systems of Automatic Mind are in conflict or competition, as the CNDS model proposes them often to be, any adjudication between them that results in a compromise or balanced influence on behavior will have to be done at a superordinate level of processing. It must draw on system goals and strategic procedures that are not the property of either of these systems but of a level of processing that is superior to both of them. This is the Reflective Mind.

Stanovich (2009) argues that measures of the executive functions actually draw upon elements of the Algorithmic Mind rather than the Reflective Mind. While the term “executive” seems superficially to suggest that these functions concern the highest level of mind, Reflective Mind, the tasks used by cognitive scientists to assess executive function actually test skills that result from Algorithmic Mind. Research in cognitive psychology in particular has been concerned with tasks that involve algorithmic level decoupling abilities: stanovich mentions “stop signal paradigms, working memory paradigms, time sharing paradigms, inhibition paradigms” which are highly suggestive of the components of executive function that Bickel et al. (2012b) found to be antipodal to state impulsivities. Individual differences in Reflective Mind capabilities are scarcely involved in these tasks if at all. The Reflective Mind, especially with respect to its involvement in epistemic regulation and cognitive allocation is involved in cognitive control at a level beyond that of the computational capacity to maintain decoupling. Stanovich (2009, p. 66) argues, therefore, that the executive functions

have been misnamed: they are essentially *supervisory processes*, he maintains, based on eternally provided rules rather than internally inaugurated decision-making. By contrast, Reflective Mind is involved in setting “the goal agenda” or in operating at the level of epistemic regulation which he defines as “directing the sequence of information pickup”. Executive functions are not engaged in this kind of work.

While the executive system belongs to the Algorithmic Mind, however, it does not constitute the Algorithmic Mind exclusively. The executive system is fundamentally involved in the overriding of the Automatic Mind but other functions of Algorithmic Mind such as the execution of decoupling are not carried out by the executive system. Similarly, the impulsive system is not the sole element of TASS; the trait impulsivities (sensation seeking and reinforcement sensitivity) also belong to TASS and are involved in moderating the tendency toward impulsivity or self control at the behavioral level. Hence, even the Type 1/Type 2 dichotomy recognizes a complexity that goes beyond that of the original CNDS model. However, Stanovich (2009, 2011; Stanovich et al., 2012) argues for a further distinction, this time between the kinds of processing for which Type 2 systems are severally responsible, which if accepted complicates the division between impulsive and executive systems made by the CNDS model. The interaction of Type 1 and Type 2 processing is evinced by the capacity of the second to prevent the automatic responses inherent in Type 1 processing to engender impulsive behaviors that result in suboptimal outcomes. “Better” responses depend on Type 2 *hypothetical reasoning* in which the individual builds models of the world and performs cognitive simulations on them. Stanovich et al. (2012, p. 787) comment, “When we reason hypothetically, we create temporary models of the world and test out actions (or alternate causes) in that simulated world,” words reminiscent of Popper’s observation that “our conjectures, our theories, die in our stead!” (Popper, 1977) In order to effect this cognitive functioning, Type 2 processes can override those of Type 1, interrupting and suppressing Type 1 functioning and then substituting alternative responses. Moreover, in order to form simulations, it is necessary to *decouple* simulated models from the real world so that they can be manipulated independently. This initiation of decoupling secondary representations from the world and maintaining them while simulation occurs is a Type 2 operation.

Having “taken TASS offline,” the Algorithmic Mind initiates decoupling which enables cognitive simulation to take place. The outcomes of this are reviewed by Reflective Mind which initiates change in serial associative cognition which influences Algorithmic Mind to develop a response. The initiation of serial associative cognition illustrates that while all hypothetical thinking involves analytical mind, not all the actions of analytic mind involve hypothetical thinking. Serial associative cognition is somewhat shallow thinking, “cognition that is not rapid and parallel such as TASS processes, but is nonetheless inflexibly locked into an associative mode that takes as its starting point a model of the world that is *given* to the subject” (Stanovich, 2009, p. 68, 70). Serial associative cognition “is serial and analytic . . . in style, but it relies on a single focal model that triggers all subsequent thought.” Hypothetical thinking constitutes a

vital *reasoning* function. The reflective and algorithmic processes of the analytic mind each have a key function within this process. Hypothetical thinking is closely related to the notion of TASS override. The analytic system must take TASS-initiated tendencies toward behavior offline and replace them with a more appropriate response. Such better responses come from cognitive simulation where they can be tested; only if they survive that will they be adopted.

Triple Processing in the Context of CNDS

It is feasible, therefore, to develop the CNDS model by incorporating MC and EASR as components of a level of processing superordinate to those of the impulsive and executive systems (**Figure 4**). This figure depicts two Type 1 impulsive or TASS systems: the first comprises the state impulsivities that Bickel et al. (2012b) showed to be antipodal to components of the Executive system; the second is composed of the two trait impulsivities, sensation seeking and reward sensitivity, that are not linked antipodally to elements of the Executive system. They are shown here as exerting modifying influences on the relationship between the State Impulsive System and the Executive System. This key relationship is shown by the bold arrow. The Executive System exerts Type 2 influence on this relationship which is modified also by the action of the Type 2 Reflective System which promotes balance between the State Impulsive System and the Executive System. The Type 2 systems draw upon Working Memory, another element ascribed to the Executive system in the original CNDS model (**Figure 1**) which has no antipodal complement in the Impulsive system, for their operations. For this reason, it is shown separately from the Type 1 and Type 2 systems in **Figure 5**. The relationship between the State Impulsive System and the Executive System (bold arrow) is the immediate precursor of the degree of temporal discounting exhibited in the individual's behavior.

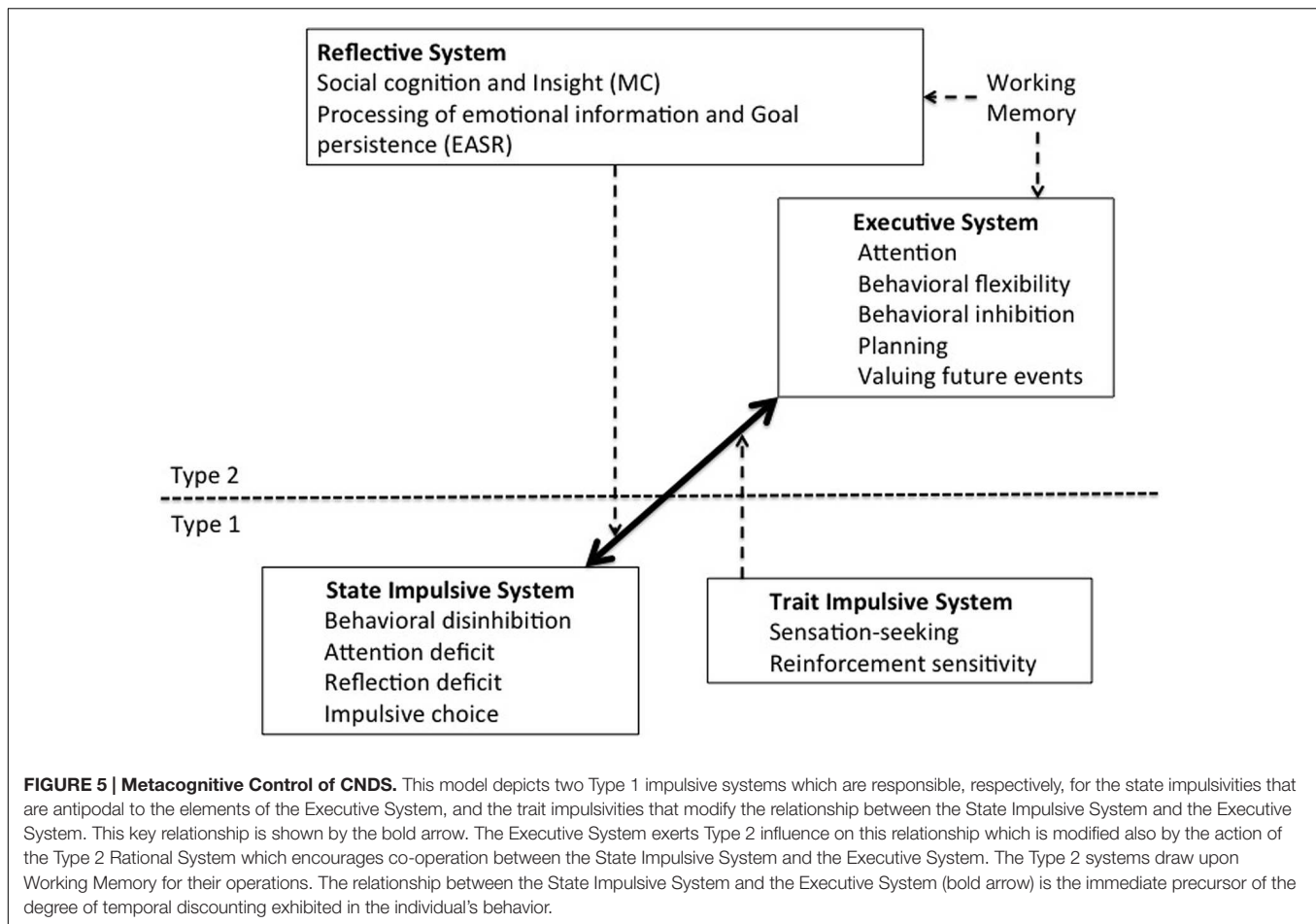
Individual differences in sensation seeking and reward or reinforcement sensitivity, which may derive from the individual's neurophysiology and/or learning history, are posited as moderating the relationship between the impulsive and executive systems. Sensation seeking is understood by Zuckerman (1979, 1994) as a preference for sensations and experiences that embody variation, novelty, and complexity, together with a willingness to incur physical and social risks in order to gain such experience. Reinforcement sensitivity reflects individual differences in susceptibility to reinforcing and aversive stimuli. Reinforcement sensitivity theory (RST; Corr, 2008; Smillie, 2008) relates propensity to behavior not only to the stimuli that have been consequential on such behavior in the past but also to the mediating neurophysiological events that are the immediate precursors of responding.

Left to itself, the Automatic Mind will act via the state impulsivities, in the absence of any influence of the behavioral inhibition, planning, and attention-maintaining tendencies of the Algorithmic Mind: the result will be a failure to reflect on the longer-range outcomes of immediate behavior, so that the resulting behavior reflects a preference for SSR over LLR. Pursuit of this short-range interest can be overcome only by an intervention of the Reflective Mind which initiates override

of the Automatic Mind via the Algorithmic Mind. Acting in response to the Reflective Mind's initiation of override, the Algorithmic Mind activates its executive functions that counter impulsivity (paying attention, drawing on behavioral flexibility and disinhibition, planning, and valuing future events) and which enable longer-term interests to be explored and pursued. Override, which thus consists in the countering of the immediate short range of the Automatic Mind by exercise of the executive functions, does not of itself result in the formulation of a plan for longer-term behavior, however. Planning with foresight entails that the Reflective Mind also initiate the decoupling of the representations for which the Automatic Mind and Algorithmic Mind are responsible so that simulation of alternative courses of action can take place. Simulation makes possible the hypothetical thinking that permits these alternatives to be generated and tested: an apparently satisfactory plan (one that is strategically and consistent with long-term goals and capabilities) engenders a response from Reflective Mind such as the pursuit of a longer-term objective in place of the impulsive action which unencumbered Automatic Mind would have produced.

The trait impulsivities can promote or impede the operations of either the Automatic Mind or the Reflective Mind, working toward the generation of either the short- or long-range interest. Trait impulsivities, sensation-seeking and reward sensitivity, are based on individual differences which are susceptible to learning history as well as the neurophysiological basis of behavior. How the trait impulsivity system works is debatable but it may be responsible for the *style* of thinking characteristic of or preferred by an individual, his/her tendency toward an analytical or intuitive approach to problem solving (Sadler-Smith, 2009). This would set limits to an individual's range of actual behaviors. Imagine a hypothetical range of behaviors from the most impulsive to the most executively controlled which contains all the actual ranges of behavior of which individuals in the population are capable. The actual range of any individual will be a subset of this. The extent of the actual subset that is the behavioral range of any individual will reflect his/her cognitive style especially as it is determined by sensation-seeking and reinforcement sensitivity, the propensity of his/her behavior to be reinforced by highly arousing stimuli and immediate reward.

The tri-process configuration captures well the requirements of the CNDS model, especially in portraying those of its elements that have been shown to be antipodal, those that remain after the establishment of antipodality has been exhausted, and the relationships among them. The tri-process model comprises an Automatic Mind which responds rapidly to environmental circumstances (which captures well the imperatives of the impulsive system posited by the CNDS model). This Automatic Mind can, however, be checked by the Algorithmic Mind (that includes the executive system which has precisely the antithetical imperatives required to counter the impulsive tendencies of Automatic Mind). The Algorithmic Mind's countering the tendencies of the Automatic Mind relies in turn on its being directed by the Reflective Mind to override the Automatic Mind in order to inaugurate the decoupling of representations based on reality so that the procedure of simulation via hypothetical thought can occur. In simulation, alternative behaviors that



might be enacted can be examined in terms of their outcomes in the short and long term. The information so gained is fed back to the Reflective Mind which inaugurates action. The Reflective Mind has additional functions which include monitoring environmental circumstances and being aware of the likely response of the Automatic Mind to them in order to initiated decoupling and simulation. These are not the functions of the Algorithmic Mind of the tri-process theory or the executive system of the CNDS model.

A tri-Process Framework

The CNDS model portrays normal and addictive behaviors as the outcomes of interaction between an impulsive system based on limbic and paralimbic brain regions and an executive system based in PFC. The interaction is indexed behaviorally by the steepness of the temporal discounting an individual's decision-making exhibits. Several of the elements of the impulsive and executive systems in the model are antipodal in terms of their definition, measurement, application to populations of addicts, and neurophysiological substrates. Specifically, the state impulsivities of the impulsive system and the elements of the executive system responsible for the CTOB display antipodality. Configuring the remaining elements of the model according to developments in multi-process theories of cognition does

not detract from the CNDS model but extends its capacity to explain normal and addictive behaviors. It has, therefore, been argued that elements of the impulsive and executive systems that do not correspond in this way, constitute additional systems that provide a more comprehensive understanding of the ways in which the interaction of the competing systems eventuate in behavioral responses. In the tri-systems theory of Stanovich (2009), a third level of processing (Reflective Mind) provides a mechanism through which the conflicting imperatives toward impulsivity and restraint of more basic systems can be managed and superseded. *Metacognition* (MC) and EASR, elements of the CNDS model's executive system which find no antipodal correspondents in the impulsive system, contribute to this third level. The state-impulsive system belongs to what Stanovich terms Automatic Mind, while the executive system belongs to Stanovich's Algorithmic Mind. The trait impulsivities, *sensation-seeking* and *reward sensitivity*, which have no antipodal correspondents in the executive system form an additional system within Automatic Mind. This trait-impulsivity system moderates the individual's behavioral output which manifests in a rate of temporal discounting. This move receives support from Stanovich and West's (2003) argument that there are individual differences in how effective Algorithmic Mind is in overriding Automatic Mind. The removal of the trait impulsivities from

the impulsive system to form another TASS makes these variables' influence more coherent; if they act negatively on the relationship between the impulsive and executive systems, they can undermine overriding, decoupling and perhaps simulation. Ross et al. (2008) note that some individuals may simply be incapable of bundling. One of the causes of this deficiency may be the overvaluation of reinforcers that arises from a tendency toward seeking unusually high levels of arousal and the particularly strong sensitivity to reward.

A Proposal for Empirical Research

The composite model summarized in **Figure 5** proposes that the outcome of conflict between the State Impulsivity System and the Executive System is the immediate precursor and cause of the rate of temporal discounting exhibited in the individual's behavior. The model suggests further that the Reflective System is responsible for the extent to which social insight and emotional control exert an inhibiting influence on the tendency toward impulsive behavior (choice of SSR over LLR); it suggests also that the Trait Impulsivity System is responsible for the degree to which the individual is inclined to control his/her impulsivity and that sensation seeking and reinforcement sensitivity are especially potent in this regard. The variables that compose the Reflective System and the Trait Impulsivity System reflect individual differences in self-control versus impulsivity. The precise measures of the effects of the Reflective System and Trait Impulsivity System remain to be empirically determined. The task of empirical research inspired by the model is to determine how and to what extent these variables impact the rate at which temporal discounting occurs and therefore the degree of balance the individual exhibits between the operation of the State Impulsive System and the Executive system. The following suggestion for empirical research is indicative in general terms of the feasibility of a research program that would facilitate the critical examination of hypotheses drawn from the model. Its principal objective at this stage is to demonstrate that the model is amenable to empirical investigation and is falsifiable in principle.

Stanovich argues that Reflective Mind involves the exercise of a cognitive style that influences the overarching approach an individual assumes in the pursuit of problem solving and decision making. As I have discussed elsewhere (Foxall, 2014b, 2016), one approach to the empirical delineation of cognitive style is provided by Kirton's (2003) adaption-innovation theory and measure. On this theory, extreme adaptors pursue solutions to problems within tried and tested frameworks of experience-based analysis and conceptualization and are likely to discount the future less steeply than innovators who seek solutions in novel and outlandish proposals which entail steeper discounting. The adaptor is likely therefore to exhibit greater capacities for social cognition and insight, to process emotional information in a more constrained fashion, and to persist in the pursuit of a goal once it has been adopted. The innovator is more likely to rely more on his/her own notions of how pursuit of a specific behavior would generate effective consequences, to be more emotionally involved in the advocacy of his/her ideas, and to be more easily deflected from current goals in favor of novel objectives. The

behavior patterns typical of adaptors and innovators may also be grounded in separate neurophysiological regions (van der Molen, 1994). There is therefore scope for empirical research which seeks to test the hypotheses (1) that adaptors exhibit a lower rate of temporal discounting on specified decision tasks than will innovators, and (2) that these cognitive styles are associated with the innervation of distinct neurophysiological regions that reflect the brain bases of high and low levels of temporal discounting.

Similar investigation is feasible by means of psychometric measures of sensation seeking and reinforcement sensitivity which may be employed to monitor the trait impulsivity of individuals engaged on tasks involving decisions that reflect differing rates of temporal discounting. Higher levels of both of these traits would be expected to associate with steeper temporal discounting and also to be linked to distinct brain regions. To the extent that sensation seeking and reinforcement sensitivity are captured by the adaption-innovation spectrum, Kirton's measure of cognitive style may also suffice for the investigation of these dimensions of trait impulsivity. The ultimate aim of research of this kind is to establish double dissociations (a) between components of cognitive functioning and rate of temporal discounting, and (b) between cognitive functions and neurophysiological activation. Initial investigation (Foxall and Yani-de-Soriano, 2011) suggests that the thorough empirical examination of the model would require psychometric investigation of the individual traits that comprise the Reflective System and the Trait Impulsive System in order to present a more fine-grained analysis of the relationships proposed by the model. Situational variables, notably the specific nature of the decision under investigation would likely influence the extent to which consumers discount the future in addition to the contribution of their fundamental cognitive styles.

DISCUSSION

Bundling in tri-Processual Perspective

Of the three components of the CNDS model – neurophysiology, decision-making (cognition), and behavior – cognition probably has received least attention. The foregoing discussion supports the conclusion, however, that if the elements of the CNDS model are configured in accordance with multi-process theories of cognitive control such as that of Stanovich (2009) the cognitive implications of the model can be made explicit. This proposition can be tested by applying the framework presented in **Figure 5** to explicate the idea of bundling.

Bundling involves an individual's adoption of a rule in order to overcome the tendency to select the inferior of two rewards as a result of discounting the future hyperbolically. The rule prescribes that one consider all of one's choices between pairs of reward of this kind in a way that makes one's present choice the precedent for later choices. In this way the individual precommits him/herself to act in a particular manner by recognizing that selecting an entire series of LL alternatives motivates him/her to avoid temporary preferences for SS options when they arise (Ainslie, 1992, 2001, 2007; Ainslie and Monterosso, 2003).

An individual who considers each choice between an SSR and a LLR as it arises is likely as we have seen to initially prefer the latter but to switch preferences when the value of the former is magnified by the fact of its imminent availability. This pattern is recursive: good resolution is followed by akrasia, not once but repeatedly. If this person resolves to consider the sum of all future SSRs in relation to the sum of all future LLRs, the conclusion is that future LLRs will cumulatively outstrip future SSRs. Crucially, the sum of the series of LLRs is also greater than the first SSR that will be encountered in the series, i.e., when it becomes immediately available, a comparison which makes acceptance of the LLR on this first occasion easier. Insofar as this first choice is predictive of later choices, following this rule makes a series of LLR choices more probable. The bundling strategy is not only theoretically defensible but also practically efficacious (Ross et al., 2008).

Both the Automatic Mind (embodying the impulsive system) and the Algorithmic Mind (executive system) are involved in this process. However, bundling requires in addition an array of mental operations which can be most appropriately understood by reference to the tri-process model we have considered. These operations include (i) holding the immediately available behavioral option in mind, (ii) holding the array of long range behaviors and their outcomes in mind, (iii) summing the outcomes of the long range behaviors, (iv) bringing the summed outcomes into comparison with that of the short range outcome, and (v) adjudicating between them. These operations cannot be carried out within either the impulsive system or the executive system. Neither has the capacity to undertake these tasks. Moreover, since the short-term and long-term interests depicted in terms of temporal discounting by piceoeconomics, exist by definition at different times, the only way in which they can be brought together is mentally, specifically through the medium of imaginative or hypothetical thinking. Representations of the two interests must be created and allowed to impinge on one another. Hence, the process of bundling is that described by Stanovich as requiring the decoupling of the Automatic Mind and the simulation by means of hypothetical thinking of alternative scenarios for future behavior. These operations require the monitoring of the behavioral tendencies of Automatic Mind in light of environmental contingencies (which must also be monitored beyond the level of the impulsive system), the initiation of override of the Automatic Mind, and the initiation of simulation via decoupling in which alternatives to the immediate uncritical pursuit of short-term gratification are hypothesized and evaluated. The only area of mind that can initiate these procedures is the Reflective Mind. The Algorithmic Mind cannot undertake such monitoring and initiating. Its functions are regulatory and supervisory rather than innovative, and bundling depends on hypothetical thinking that brings a multiplicity of long range outcomes and short range outcomes into the same arena and allows them to impinge on one another so that a calculation based on the valuation of the separate outcomes and a selection the appropriate action can be made.

There is another reason why the tri-process model is particularly relevant to the analysis of normal and addictive

decision-making by means of the CNDS model and piceoeconomics. Neither the state-impulsive system element of the Automatic Mind nor the executive system of the Algorithmic Mind can adjudicate between the imperatives of immediate gratification that fulfill the short-range interest embodied in the former and the delayed benefits that fulfill those of the long-range interest. Both the CNDS model and piceoeconomics are enhanced by the inclusion of Reflective Mind in the overall system of decision-making they posit. The Reflective Mind is a kind of *present self* that can hypothesize about the behaviors of one's *past self* and *future self*. Hypothetical reasoning requires that representations of the real world not interfere with representations of imaginary situations. In comparing the pleasure to be obtained by ingesting a recently acquired drug with the deleterious consequences of a series of binges in the future, it is necessary to differentiate clearly the monetary cost of the newly obtained supply of the drug from the imagined emotional and social as well as financial costs of sustained consumption that would be the outcome of bingeing. These abstract operations require the participation of a Reflective Mind.

Multiple Selves or Incompatible Interests?

This account of bundling operations implies that the rational decision-making element of the model shown in **Figure 5** is Reflective Mind. the impulsive and executive systems (inherent in Automatic Mind and Algorithmic Mind, respectively) appear to be largely neurophysiological systems that are under the ultimate control of Reflective Mind. It is here that personhood or agency is located: while it may experience the conflict of having to decide between alternative interests by determining the content of the utility function that will be the outcome of its behavior, it is a single person.

Dual process models such as CNDS contain the conflict between short- and long-range interests within warring impulsive and executive systems; piceoeconomics, whilst open to multiple selves, also tends in practice largely to confine its deliberations to these two categories of mental operation. But there are elements of the executive system, as defined in **Figure 2**, that tri-process theories such as that of Stanovich (2009) suggest play an overarching role in the relationships between the systems it characterizes as Automatic Mind and Algorithmic Mind and which, respectively, embrace the impulsive and executive systems. The restructuring of the model components, achieved in **Figure 3**, proposes that MC and EASR would constitute part of this higher level system, the Reflective Mind, which would be involved in the regulation of the Automatic Mind which otherwise would respond to environmental stimuli spontaneously or impulsively. The regulation imposed by Reflective Mind would take the form of its "innervating" Algorithmic Mind to initiate the overriding of Automatic Mind and the decoupling of mundane mental representations so that the simulation of

hypothetical futures can be accomplished. Reflective Mind is also portrayed in Stanovich's theory as receiving the outputs of simulation and effecting a response at the level of the entire organism.

Caution is essential on the part of psychologists, whether cognitive or behavior analytical, in their treatment of models of this kind. Models that depict cognitive operations necessarily deal in unobservables and there is a danger that these will be multiplied without a firm empirical basis being provided for them. In applied areas such as the treatment of addiction and other forms of excessive consumption it may be necessary, however, to use ascriptions of thought processes to individuals in order to understand fully their behavior possibly as a prelude to predicting and/or modifying it. There is, moreover, little in the modeling which has been the subject of this paper that could not be captured, though perhaps less economically, in the language of behavior analysis and in particular that of verbal behavior and rule-governance. There need be no internecine conflict between adherents of different vocabularies as long as the analysis is comprehensible to all in terms of their several theoretical viewpoints. This is especially the case if the new analysis proves useful at the level of effecting the prediction and control of behavior but it is also justified if its principal contribution is the furtherance of understanding of the processes involved in shaping and maintaining that behavior.

Where does this model-building lead in terms of resolving the question of multiple selfhood that opened this paper? Ross (2009; see also Ross, 2005) argues that there are present at t_1 and t_2 , respectively, different *agents* as is demonstrated by there being different utility functions at each time. However, useful this agential distinction process is as an analytical device, the degree of difference between *selves* or *persons* must not be exaggerated. To argue that different persons or selves exist at t_1 and t_2 may be self-defeating since it is only by establishing a considerable degree of continuity between the person who exists at t_1 and the person who exists at t_2 that we can comprehend why intrapersonal conflict arises. If the t_1 and t_2 persons or selves are remote from one another we can argue that neither is bound by what the other has done (Hanson, 2009). If the selves can establish this degree of moral separateness, it is difficult to see how the motivational conflict engendered by the contemporaneous existence of competing interests necessary for picoeconomics can come about.

What exists at t_1 and t_2 is the organism; without the assumption of at least this degree of continuity there would be no conflict. The question is whether this self that persists is the agent or whether "interests" or "selves" existing within the person can be thought of as agents each of which has a separate utility function. I would argue that the person who exists at t_1 and t_2 has different interests or motivations on each occasion because he/she is facing different contingencies of reinforcement and punishment. It is unnecessarily metaphysical to argue that two persons or agents are involved: it is one person or agent with conflicting interests. The interests have neurophysiological correlates within the person which form a central aspect of the impulsive and executive systems; these neurophysiological

events are to be understood at the sub-personal level of exposition. The cognitive dimensions of the Automatic Mind and Algorithmic Mind which include these and other systems are to be understood, however, at the personal level of beliefs, desires, perceptions, and emotions. Not only, contra Dennett (1969, 1987), can the intentionality that properly belongs at the personal level not be ascribed to sub-personal entities (Foxall, 2007); it is also not possible for the interests to have utility functions of their own and thus be considered agential. Rather, a person's interests determine his/her preference structure (as revealed in his/her choice *behavior*) which eventuates in his/her utility function (the final configuration of the consequences of his/her behavior (Rachlin, 1994, 1989). What changes from t_1 to t_2 is the contingencies of reinforcement with which the individual is faced; these have the effect of modifying his/her utility function. The preferences embodied in his/her behavior reflect the dominance of either the impulsive and executive system (or the Automatic Mind and Algorithmic Mind of which they are subsets) at that time. The Reflective Mind is the personal forum within which the deliberations regarding the alternative behaviors available and their likely outcomes takes place. It is within the person that conflict occurs and is felt.

The rational individual for whom *akrasia* is a commonplace experience is not, therefore, two persons, or agents in the course a day, but a single person who encounters differing situations and changes his preferences accordingly; these preferences are revealed in the earlier verbal behavior which values highly constructive work and in the subsequent physical behavior which values recreation. These contingencies of reinforcement are each advantageous in their way at different times and insofar as he/she is conscious of them they signal the benefits that will derive from his/her behavior. These benefits form my interests. There is no need to translate these extrapersonal interests into intrapersonal homunculi that compete. What compete are (i) the differing contingencies of reinforcement (at the super-personal level), and (b) the differing neurophysiological tendencies (at the sub-personal level). The personal level is concerned with acting upon one or other of these, and/or adjudicating between them. In the course of debating different courses of action, any rational person may experience cognitive discomfort, feel as though they are being wrenched first by one alternative then the other as they participate in "making a choice." But at no time is this individual anything other than a person facing incompatible options who finds this situation aversive. He/she has only one utility function at a time, by virtue of being able to perform only one behavior at a time and his/her utility function is the outcome of that behavior.

CONCLUSION

This paper has argued that the CNDS model's obvious strengths can be enhanced through consideration of the categorical structure of the model and the functions of its components. Hence, the formulation shown in **Figure 5** seems more closely

aligned with the results of the investigation of the antipodality of the impulsive and executive systems undertaken by Bickel et al. (2012b). The question that arises from that exercise is where the components of the impulsive and executive systems (Figure 2) that have no antipodal correspondents should be positioned within the CNDS model. The proposal to remove MC and EASR from the Executive System and accord it a superordinate role in cognitive control of (a) the impulsive and executive systems and (b) overt behavior suggested in Figure 3 is borne out by the theoretical reasoning advanced by multi-process modelers such as Stanovich (2009). Recent theoretical development suggests also that several systems comprise Automatic Mind (Stanovich, 2011) and this offers a role to the state impulsivities that do not exhibit antipodality with any of the elements of the Executive System depicted in Figure 2. A final tri-process model which incorporates the restructured components of the CNDS hypothesis is put forward in Figure 5. Although this model does not essentially contain any components not already present in the original CNDS model (Figure 2), it aims to present their interrelationships in a way that is consistent with their functions in controlling the imperatives of impulsivity and self-control and the individual's behavior.

REFERENCES

- Ainslie, G. (1992). *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. Cambridge: Cambridge University Press.
- Ainslie, G. (2001). *Breakdown of Will*. Cambridge: Cambridge University Press.
- Ainslie, G. (2007). "Emotion: the gaping hole in economic theory," in *Economics and the Mind*, eds B. Monterosso and M. D. White (New York, NY: Routledge), 11–28.
- Ainslie, G. (2010). "Procrastination: the basic impulsive," in *The Thief of Time*, eds C. Andreou and M. White (Oxford: Oxford University Press), 11–27.
- Ainslie, G., and Monterosso, J. (2003). Building blocks of self-control: increased tolerance for delay with bundled rewards. *J. Exp. Anal. Behav.* 79, 37–48. doi: 10.1901/jeab.2003.79.37
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unified theory of ADHD. *Psychol. Bull.* 121, 65–94.
- Barkley, R. A. (2012). *Executive Functions: What They Are, How They Work, and Why They Evolved*. New York, NY: Guilford Press.
- Bechara, A. (2005). Decision-making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. *Nat. Neurosci.* 8, 1458–1463. doi: 10.1038/nn1584
- Bickel, W. K., Jarmolowicz, D. P., MacKillop, J., and Epstein, L. H. (2013). "What is addiction?," in *Addictions: A Comprehensive Guidebook*, 2nd Edn, eds B. S. McCrady and E. E. Epstein (New York, NY: Oxford University Press), 3–16.
- Bickel, W. K., Jarmolowicz, D. P., MacKillop, J., Epstein, L. H., Carr, K., Mueller, E. T., et al. (2012a). "The behavioral economics of reinforcer pathologies: novel approaches to addictive disorders," in *APA Addiction Syndrome Handbook*, ed. H. J. Schaffer (Washington, DC: American Psychological Association), 333–363.
- Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Gatchalian, K. M., and McClure, S. M. (2012b). Are executive function and impulsivity antipodes? A conceptual reconstruction with special reference to addiction. *Psychopharmacology* 221, 361–387. doi: 10.1007/s00213-012-2689-x
- Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., and Gatchalian, K. M. (2011). The behavioral economics and neuroeconomics of reinforcer pathologies: implications for etiology and treatment of addiction. *Curr. Psychiatry Rep.* 13, 406–415. doi: 10.1007/s11920-011-0215-1
- Bickel, W. K., and Marsch, L. A. (2000). "The tyranny of small decisions: origins, outcomes, and proposed solutions," in *Reframing Health Behavioral Change with Behavioral Economics*, eds W. R. Bickel and R. E. Vuchinich (Mahwah, NJ: Erlbaum), 341–391.
- Bickel, W. K., Miller, M. L., Yi, R., Kowal, B. P., Lindquist, D. M., and Pitcock, J. A. (2007). Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. *Drug Alcohol Depend.* 90S, S85–S91. doi: 10.1016/j.drugalcdep.2006.09.016
- Bickel, W. R., and Yi, R. (2008). Temporal discounting as a measure of executive function: insights from the competing neuro-behavioral decisions systems hypothesis of addiction. *Adv. Health Econ. Health Serv. Res.* 20, 289–309. doi: 10.1016/S0731-2199(08)20012-9
- Bickel, W. R., and Yi, R. (2010). "Neuroeconomics of addiction: the contribution of executive dysfunction," in *What is Addiction?*, eds D. Ross, H. Kincaid, D. Spurrett, and P. Collins (Cambridge, MA: MIT Press), 1–25.
- Corr, P. J. (2008). "Reinforcement sensitivity theory (RST): introduction," in *The Reinforcement Sensitivity Theory of Personality*, ed. P. J. Corr (Cambridge: Cambridge University Press).
- Dallery, J., and Soto, P. L. (2012). "Quantitative description of environment-behavior relations," in *APA Handbook of Behavior Analysis: Methods and Principles*, Vol. 1, ed. G. J. Madden (Washington, DC: American Psychological Association), 219–249.
- Dennett, D. C. (1969). *Content and Consciousness*. London: Routledge and Kegan Paul.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
- Dennett, D. C. (1996). *Kinds of Minds: Towards an Understanding of Consciousness*. London: Weidenfeld and Nicholson.
- Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.* 11, 732–744. doi: 10.1038/nrn3827
- Elster, J. (1987). "Deception and self-deception in *stendhal*: some sartrean themes," in *The Multiple Self*, ed. J. Elster (Cambridge: Cambridge University Press).
- Evans, J. S., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241. doi: 10.1177/1745691612460685
- Ferster, C. B., and Skinner, B. F. (1957). *Schedules of Reinforcement*. New York, NY: Prentice-Hall.
- Foxall, G. R. (2007). Intentional behaviorism. *Behav. Philos.* 35, 1–56.

Three themes emerge from this analysis. The first is the implications for the CNDS model of considering cognitive control of behavior in light of the tri-process theory. The second is the capacity of the tri-process depiction of neuro-behavioral decision-making to enhance understanding of addictive behavior and its resolution. This is discussed by Foxall (2016) in terms of the mental operations that are presupposed by piceoeconomic bundling and their incorporation in the model presented in Figure 5, as is the nature of the multiple agents apparently involved in the breakdown of will and its resolution is discussed in the context of the model.

AUTHOR CONTRIBUTION

GF is entirely responsible for the paper.

ACKNOWLEDGMENTS

I should like to express my immense gratitude to two reviewers for incisive comments on an earlier draft. Portions of this paper are to appear in Foxall (2016).

- Foxall, G. R. (2011). *Interpreting Consumer Choice: The Behavioral Perspective Model*. New York: Routledge.
- Foxall, G. R. (2014a). "Neurophilosophy of explanation in economic psychology: an exposition in terms of neuro-behavioral decision systems," in *Routledge Companion to the Future of Marketing*, eds L. Moutinho, E. Bigne, and A. K. Manrai (London: Routledge), 134–150.
- Foxall, G. R. (2014b). Cognitive requirements of competing neuro-behavioral decision systems: some implications of temporal horizon for managerial behavior in organizations. *Front. Hum. Neurosci.* 8:184. doi: 10.3389/fnhum.2014.00184
- Foxall, G. R. (2015). Operant behavioral economics. *Manage. Decis. Econ.* doi: 10.1002/mde.2712
- Foxall, G. R. (2016). *Addiction as Consumer Choice: Exploring the Cognitive Dimension*. London: Routledge.
- Foxall, G. R., and Sigurdsson, V. (2011). Drug use as consumer behavior. *Behav. Brain Sci.* 34, 313–314. doi: 10.1017/S0140525X11000707
- Foxall, G. R., and Yani-de-Soriano, M. (2011). Influence of reinforcement contingencies and cognitive styles on affective responses: an examination of Rolls's theory of emotion in the context of consumer choice. *J. Appl. Soc. Psychol.* 41, 2506–2535. doi: 10.1111/j.1559-1816.2011.00823.x
- Frankish, K., and Evans, J. St. B. T. (2009). "The duality of mind: an historical perspective," in *Two Minds: Dual Processes and Beyond*, eds J. St. B. T. Evans and K. Frankish (Oxford: Oxford University Press), 1–29.
- Glimcher, P. (2011). *Foundations of Neuroeconomic Analysis*. New York, NY: Oxford University Press.
- Grace, R. C., and Hucks, A. D. (2012). "The allocation of operant behavior," in *APA Handbook of Behavior Analysis: Methods and Principles*, Vol. 1, ed. G. J. Madden (Washington, DC: American Psychological Association), 307–337.
- Gruber, A. J., and McDonald, R. J. (2012). Context, emotion, and the strategic pursuit of goals: interactions among multiple brain systems controlling motivated behavior. *Front. Behav. Neurosci.* 6:50. doi: 10.3389/fnhum.2012.00050
- Hanson, C. (2009). *Thinking about Addiction: Hyperbolic Discounting and Responsible Agency*. Amsterdam: Rodopi.
- Herrnstein, R. J. (1997). *The Matching Law: Papers in Psychology and Economics*, eds H. Rachlin and D. I. Laibson (Cambridge, MA: Harvard University Press).
- Hursh, S. R., Madden, G., Spiga, R., Iser, G. D., and Francisco, M. T. (2012). "The translational utility of behavior economics: the experimental analysis of consumption and choice," in *APA Handbook of Behavior Analysis. Translating Principles into Action*, Vol. 2, ed. G. J. Madden (Washington, DC: American Psychological Association), 191–224.
- Jacobs, E. A., Borrero, J. C., and Vollmer, T. R. (2012). "Translational applications of quantitative choice models," in *APA Handbook of Behavior Analysis: Translating Principles into Action*, Vol. 2, ed. G. J. Madden (Washington, DC: American Psychological Association), 165–190.
- Jarmolowicz, D. P., Mueller, E. T., Koffarnus, M. N., Carter, A. E., Gatchalian, K. M., and Bickel, W. K. (2013). "Executive dysfunction in addiction," in *The Wiley-Blackwell Handbook of Addiction Pharmacology*, eds J. MacKillop and H. de Wit (Chichester: Wiley), 27–61.
- Jentsch, J. D., and Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behaviour by reward-related stimuli. *Psychopharmacology* 146, 373–390. doi: 10.1007/PL00005483
- Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720. doi: 10.1037/0003-066X.58.9.697
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.
- Kirton, M. J. (2003). *Adaption-Innovation: In the Context of Diversity and Change*. London: Routledge.
- Koob, G. E. F. (2013). "Neuroscience of addiction," in *Addictions: A Comprehensive Guidebook*, 2nd Edn, eds B. S. McCrady and E. E. Epstein (New York: Oxford University Press), 17–35.
- London, E. D., Ernst, M., Grant, S., Bonson, K., and Weinstein, A. (2000). Orbitofrontal cortex and human drug abuse: functional imaging. *Cereb. Cortex* 10, 334–342. doi: 10.1093/cercor/10.3.334
- Madden, G. J., and Bickel, W. K. (eds) (2010). *Impulsivity: The Behavioral and Neurological Science of Discounting*. Washington, DC: American Psychological Association.
- Mazur, J. (1987). "An adjusting procedure for studying delayed reinforcement," in *Quantitative Analysis of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value*, Vol. 5, eds M. Commons, J. Mazur, J. Nevin, and H. Rachlin (Hillsdale, IL: Erlbaum), 55–73.
- McClure, S. M., Laibson, D. L., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507. doi: 10.1126/science.1100907
- McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.* 27, 1–28. doi: 10.1146/annurev.neuro.27.070203.144157
- Migaud, M., Charlesworth, P., Dempster, M., Webster, L. C., Watabe, A. M., Makhinson, M., et al. (1998). Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95 protein. *Nature* 396, 433–439. doi: 10.1038/24790
- Miller, E. K., and Wallis, J. D. (2009). "Executive function and higher-order cognition: definition and neural substrates," in *Encyclopedia of Neuroscience*, Vol. 4, ed. L. R. Squire (Oxford: Academic Press), 99–104.
- Moser, M.-B., Rowland, D. C., and Moser, E. I. (2015). Place cells, grid cells, and memory. *Cold Spring Harb. Perspect. Biol.* 7, a021808. doi: 10.1101/cshperspect.a021808
- Naqvi, N., Shiv, B., and Bechara, A. (2006). The role of emotion in decision-making. *Curr. Dir. Psychol. Sci.* 15, 260–264. doi: 10.1111/j.1467-8721.2006.00448.x
- Naqvi, N. H., and Bechara, A. (2008). The hidden island of addiction: the insula. *Trends Neurosci.* 32, 56–67. doi: 10.1016/j.tins.2008.09.009
- Pan, W. X., Schmidt, R., Wickens, J. R., and Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25, 6235–6242. doi: 10.1523/JNEUROSCI.1478-05.2005
- Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.* 57, 27–53. doi: 10.1146/annurev.psych.56.091103.070234
- Popper, K. (1977). *Natural Selection and the Emergence of Mind*. Cambridge: Darwin College.
- Rachlin, H. (1989). *Judgment, Decision, and Choice: A Cognitive/Behavioral Synthesis*. New York, NY: Freeman.
- Rachlin, H. (1994). *Behavior and Mind: The Roots of Modern Psychology*. Oxford: Oxford University Press.
- Rachlin, H. (2000a). *The Science of Self-Control*. Cambridge, MA: Harvard University Press.
- Rachlin, H. (2000b). "The lonely addict," in *Reframing Health Behavior Change with Behavioral Economics*, eds W. K. Bickel and R. E. Vuchinich (Mahwah, NJ: Erlbaum), 145–164.
- Rick, S., and Loewenstein, G. (2008). "The role of emotion in economic behavior," in *Handbook of Emotions*, 3rd Edn, eds M. Lewis, J. M. Haviland-Jones, and L. F. Barrett (New York, NY: Guilford Press), 138–156.
- Robbins, T. W., and Everitt, B. J. (2002). "Dopamine: its role in behavior and cognition in experimental animals and humans," in *Dopamine in the CNS II*, ed. G. Di Chiara (Berlin: Springer), 173–211.
- Rolls, E. T. (2008). *Memory, Attention, and Decision-Making*. Oxford: Oxford University Press.
- Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. Cambridge, MA: MIT Press.
- Ross, D. (2009). "Economic models of procrastination," in *The Thief of Time*, eds C. Andreou and M. White (Oxford: Oxford University Press), 28–50.
- Ross, D. (2012). "The economic agent: not human, but important," in *Philosophy of Economics*, ed. U. Mäki (Amsterdam: Elsevier), 691–736.
- Ross, D., Sharp, C., Vuchinich, R. E., and Spurrett, D. (2008). *Midbrain Mutiny: Picoeconomics and Neuroscience of Disordered Gambling*. Cambridge, MA: MIT Press.
- Sadler-Smith, E. (2009). "A duplex model of cognitive style," in *Perspectives on the Nature of Intellectual Styles*, eds L. Zhang and R. J. Sternberg (New York, NY: Springer), 3–28.
- Sanders, H., Rennó-Costa, C., Idiart, M., and Lisman, J. (2015). Grid cells and place cells: an integrated view of their navigational and memory

- function. *Trends Neurosci.* 38, 763–775. doi: 10.1016/j.tins.2015.10.004
- Schultz, W. (1992). Activity of dopamine neurons in the behaving primate. *Semin. Neurosci.* 4, 129–138. doi: 10.1016/1044-5765(92)90011-P
- Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6:24. doi: 10.1186/1744-9081-6-24
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500. doi: 10.1146/annurev.neuro.23.1.473
- Searle, J. (2001). *Rationality in Action*. Cambridge, MA: MIT.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., and Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends Cogn. Sci.* 18, 186–193. doi: 10.1016/j.tics.2014.01.006
- Smillie, L. D. (2008). What is reinforcement sensitivity? Neuroscience paradigms for approach-avoidance process theories of personality. *Eur. J. Pers.* 22, 359–384. doi: 10.1002/per.674
- Stanovich, K. E. (2009). “Distinguishing the reflective, algorithmic, and autonomous minds: is it time for a tri-process theory?,” in *Two Minds: Dual Processes and Beyond*, eds J. St. B. T. Evans and K. Frankish (Oxford: Oxford University Press), 55–88.
- Stanovich, K. E. (2011). *Rationality and the Reflective Mind*. Oxford: Oxford University Press.
- Stanovich, K. E., and West, R. F. (2000). Individual differences in reasoning: implications for the rationality debate? *Brain Behav. Sci.* 23, 645–726. doi: 10.1017/S0140525X00003435
- Stanovich, K. E., and West, R. F. (2003). “Evolutionary versus instrumental goals: how evolutionary psychology misconceives human rationality,” in *Evolution and the Psychology of Thinking: The Debate*, ed. D. E. Over (New York, NY: Psychology Press).
- Stanovich, K. E., West, R. F., and Toplak, M. E. (2012). “Intelligence and rationality,” in *Cambridge Handbook of Intelligence*, eds R. Sgternberg and S. B. Kaufman (Cambridge: Cambridge University Press), 784–826.
- Stein, J. S., and Madden, G. J. (2013). “Delay discounting and drug abuse: empirical, conceptual, and methodological considerations,” in *The Wiley-Blackwell Handbook of Addiction Psychopharmacology*, eds J. MacKillop and H. de Wit (Chichester: Wiley), 165–208.
- Tiffany, S. T. (1999). Cognitive concepts of craving. *Alcohol Res. Health* 23, 215–224.
- Uetani, N., Kato, K., Ogura, H., Mizuno, K., Kawano, K., Mikoshiba, K., et al. (2000). Impaired learning with enhanced hippocampal long-term potentiation in PTP8-deficient mice. *EMBO J.* 19, 2775–2785. doi: 10.1093/emboj/19.12.2775
- van der Molen, P. P. (1994). “Adaption-innovation and changes in social structure: on the anatomy of catastrophe,” in *Adaptors and Innovators: Styles of Creativity and Problem Solving*, ed. M. J. Kirton (London: Routledge), 137–172.
- Verdejo-Garcia, A., and Bechara, A. (2009). A somatic marketing theory of addiction. *Neuropharmacology* 56(Suppl. 1), 48–62. doi: 10.1016/j.neuropharm.2008.07.035
- Zuckerman, M. (1979). *Sensation Seeking: Beyond the Optimum Level of Arousal*. Hillsdale, IL: Erlbaum.
- Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge: Cambridge University Press.

Conflict of Interest Statement: 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.

Copyright © 2016 Foxall. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Web Ontologies to Categorical Structure Reality: Representations of Human Emotional, Cognitive, and Motivational Processes

Juan-Miguel López-Gil^{1*}, Rosa Gil² and Roberto García²

¹ Department of Computer Languages and Systems, University of the Basque Country, Vitoria-Gasteiz, Spain, ² Department of Computer Science and Engineering, Universitat de Lleida, Lleida, Spain

OPEN ACCESS

Edited by:

Paul M. W. Hackett,
University of Oxford, UK

Reviewed by:

Hiroaki Wagatsuma,
Kyushu Institute of Technology, Japan
Maya Ivanova Dimitrova,
Bulgarian Academy of Sciences,
Bulgaria

*Correspondence:

Juan-Miguel López-Gil
juanmiguel.lopez@ehu.es

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 01 December 2015

Accepted: 04 April 2016

Published: 25 April 2016

Citation:

López-Gil J-M, Gil R and García R
(2016) Web Ontologies to Categorical
Structure Reality: Representations of
Human Emotional, Cognitive, and
Motivational Processes.
Front. Psychol. 7:551.
doi: 10.3389/fpsyg.2016.00551

This work presents a Web ontology for modeling and representation of the emotional, cognitive and motivational state of online learners, interacting with university systems for distance or blended education. The ontology is understood as a way to provide the required mechanisms to model reality and associate it to emotional responses, but without committing to a particular way of organizing these emotional responses. Knowledge representation for the contributed ontology is performed by using Web Ontology Language (OWL), a semantic web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language such that computer programs can exploit knowledge expressed in OWL and also facilitates sharing and reusing knowledge using the global infrastructure of the Web. The proposed ontology has been tested in the field of Massive Open Online Courses (MOOCs) to check if it is capable of representing emotions and motivation of the students in this context of use.

Keywords: ontology, upper ontologies, emotion, cognition, motivation, Massive Open Online Courses

INTRODUCTION

Ontology has been a field of philosophy since Aristotle and from its beginnings it has been characterized as a study of existence, a compendium of all there is in the world. Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology is the “branch of metaphysics that concerns itself with what exists” (Blackburn, 1996). Ontology deals with questions concerning what entities exist or may be said to exist, and how such entities may be grouped, related within a hierarchy, and subdivided according to similarities and differences.

Although, ontology as a philosophical enterprise is highly theoretical, the use of ontologies has expanded considerably in recent years in order to reflect different real-world concepts. One practical application is ontology engineering in information science and information technology, where an ontology is a formal naming and definition of the types, organized taxonomically, plus their properties and interrelationships that exist for a particular domain of discourse. In fact, an often-cited definition of ontologies in computer and information sciences is that “an ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words” (Guarino, 1998).

One of the most recent and potentially disrupting contributions of computer science to ontology has been moving them to the Web. In a Web Ontology, all defined concepts are identified using Web Uniform Resource Identifiers (URI), like:

<http://sw.opencyc.org/concept/Mx4rvZLWajwpEbGdrcN5Y29ycA>

This URI corresponds to the concept for the field of study “Psychology.” This simplifies reusing this concept in other Web ontologies to just pointing to this URI from the ontology reusing the concept. For instance, if we want to define in our ontology the concept “Educational psychology,” we do not need to fully define it. We can define a subconcept that points using the *subClassOf* relation to the nearest one in OpenCyc ontology using its URI and just define the particularities of the new one. For instance:

<http://myspy.org/concept/EducationalPsychology>
subConceptOf

<http://sw.opencyc.org/concept/Mx4rvZLWajwpEbGdrcN5Y29ycA>

In order to make Web Ontologies really interoperable, basic relations like *subClassOf* have to be normalized so their meaning can be understood across different Web Ontologies. To this end, the World Wide Web Consortium (W3C) has standardized the Web Ontology Language (OWL) (Hitzler et al., 2009) for defining ontologies in web environments. OWL ontologies have formally defined meaning and provide classes, properties, individuals, and data values and are stored as shareable documents in the Internet. Web ontologies are formalized vocabularies of terms, often covering a specific domain and shared by a community of users. They specify the definitions of terms by describing their relationships with other terms in the ontology. Besides, they are especially useful when dealing with complex conceptual frameworks, as they provide an unambiguous representation of a conceptual framework and are expressive enough to make it possible to automate to a great extent sophisticated information processing services. Furthermore, ontology integration involves the creation of bridging modules between ontologies that accurately reflect the shared understanding of the semantic relationships between the different entities in the different ontologies (Hastings et al., 2014). The standard defines a set of primitives that constitute the building blocks to represent ontologies in the Web, like the *owl:subClassOf* relationship or the *owl:Class* concept. OWL also defines how these building blocks should be interpreted, their semantics, so all tools processing OWLs interpret them in a coherent way and automated reasoning is possible.

The hypothesis explored in this work is that it is possible to develop a functional Web ontology capable of linking categorical structures representing reality and the emotional, cognitive and motivational states people associates to these representations. The emotional state of people is important as it modifies their perception of the world, so it is important not only to adequately describe categorically structured ways of understanding the world around us, but also to describe the emotional, cognitive and motivational processes of people to understand how they

perceive and interpret the world around them. Besides, both descriptions of reality and emotion, cognition and motivation can also be modeled by means of Web ontologies and all the knowledge shared in a common framework. This article describes the current state of development of an ontology that meets the previous goals, in which significant improvements have been introduced regarding the cognitive model in order to represent mechanisms that have proven to be relevant when it comes to recognize and generate emotions. Motivation has been introduced as a key element in generating emotional responses (Sartre, 1939; Lazarus, 1991).

In addition to explaining performed improvements, and in order to validate the proposal, it is also shown how the proposed ontology has been used in the field of Massive Open Online Courses (MOOC). Adequate representation of emotions and motivation is especially important in MOOCs in order to ensure their success among people who use them. A virtual agent was developed in order to gather information about how users interact with the system and assess how they felt and perceived everything surrounding the MOOC. Presented ontology was designed to link gathered user interaction data with the description of the MOOC environment, the concepts deployed and people interacting with the platform.

The rest of the paper is structured as follows. Next section describes materials and methods, i.e., the ontologies in the field of information sciences, the notion of upper ontologies and OWL, together with the ontology engineering methodologies applied to develop the contributed ontology. Section Results: the Emotions & Cognition Ontology presents resulting ontology for linking reality with its perception by human beings using emotion, cognitive and motivational processes, including information about existing models on how emotion, cognitive and motivational processes affect the perception of the surrounding world by individuals. Section Evaluation: Massive Open Online Courses presents the evaluation of this ontology in the context of MOOCs, where it allows determining what users perceived and felt while interacting with a MOOC. Finally, Section Conclusions outlines the conclusions.

MATERIALS AND METHODS

Philosophers classify ontologies in various ways using criteria such as the degree of abstraction and field of application:

- **Upper ontology:** concepts supporting development of an ontology, meta-ontology. Also known as a top-level ontology or foundation ontology. It describes very general concepts that are the same across all knowledge domains.
- **Domain ontology:** concepts relevant to a particular topic or area of interest, for example, information technology or computer languages, or particular branches of science.
- **Interface ontology:** concepts relevant to the juncture of two disciplines.
- **Process ontology:** inputs, outputs, constraints, sequencing information, involved in business or engineering processes.

Recently, ontology has evolved a lot in the computer science and artificial intelligence fields. In these fields, an ontology is viewed as a formal, explicit specification of a shared conceptualization (Studer et al., 1998). “Formal” in the sense that it is an abstract model of a portion of the world and “explicit specification” because it is machine-readable and understandable. “Shared” implies that ontologies are based on a consensus and “conceptualization” that they are expressed in terms of concepts, properties, etc.

Ontologies were first used in Artificial Intelligence to facilitate knowledge sharing and reuse. Currently, their use is expanding to other disciplines related to information technologies and are starting to play an important role in supporting the information exchange processes, as they provide a shared and common understanding of a domain.

In computer science, ontologies are constructed using knowledge representation languages and logics. This allows automatic reasoning using the knowledge captured by ontologies. A great part of the meaning of expressions can be captured combining simpler concepts and conceptual relations. At the end, some preliminary set of fundamental concepts and relations is found. This set must have a rich semantic grounding in order to make powerful and valid automatic reasoning. Moreover, if it is shared among a great community, it may permit a great level of understanding.

As previously introduced, and also in the computer science field, the kind of ontologies providing fundamental concepts and relations are called upper ontologies. Upper ontologies, also known as foundational or top-level ontologies, try to formalize the more general concepts in our conception of the world and reality. These ontologies are fundamental to facilitate information and knowledge integration by automatic means. Thus, there have been many attempts to produce upper ontologies as detailed in **Table 1**.

Focusing on how upper ontologies can be used as Web ontologies, the previous example defining our custom concept representing “Educational Psychology” in terms of “Psychology” in OpenCyc can be represented using OWL as shown in **Table 2**. The first block, from lines 1 to 7, defines the other ontologies and parts of the OWL standard that are going to be reused. Then, lines 9–11 define the new ontology, including the URI that will be its global identifier (<http://myspy.org/concept/>) and a human-friendlier label, “My Psychology Terms”. Finally, from line 13 to 16, the class for “Educational Psychology” is defined as a subclass of the “Psychology” class in OpenCyc. This example finishes with the closing mark in line 18, though a real ontology would include many more class definitions together with properties representing relationships among them, like the property “isTopicFor” relating concepts like the previous one and the class “Course.”

TABLE 1 | Summary of the analyzed Upper Ontologies.

Ontology Name	Description
Cyc	One of the biggest foundational ontologies is Cyc (Lenat, 1995), a project started in 1984 and that currently defines more than 239,000 concepts. A subset of that ontology has been released as an open ontology under the name OpenCyc and a more complete one for research purposes called ResearchCyc. The main value of this ontology is the enormous coverage it has gained over the years.
BORO Business Objects Reference Ontology	BORO is a reference ontology, designed for developing ontological or semantic models for large complex operational applications. It is based on a 4 Dimensional approach, where time is treated as another dimension, making it easier to capture change patterns. BORO also facilitates reuse because it is conceived as a framework to develop other ontologies under the same foundations.
UMBEL Upper Mapping and Binding Exchange Layer	It is an ontology of 28,000 reference concepts that maps to a simplified subset of the OpenCyc ontology. It provides near wide coverage of OpenCyc without the complexity of the knowledge representation languages used to define Cyc.
BFO Basic Formal Ontology	This ontology is specially conceived for the sciences, though it is kept really small because it does not enter into the particularities of any scientific domain. On the other hand, it is very generic because it incorporates both three-dimensionalist and four-dimensionalist perspectives on reality within a single framework.
DOLCE Descriptive Ontology for Linguistic and Cognitive Engineering	DOLCE (Gangemi et al., 2002) is an upper ontology with a clear cognitive bias, in the sense that it aims at capturing the ontological categories underlying natural language and human commonsense. Consequently, it is in many cases easier for non-ontology experts. For instance, the fundamental distinction between enduring and perduring entities, i.e. continuants and occurrents, is motivated by our cognitive bias.
SUMO Suggested Upper Merged Ontology	Many upper ontology initiatives were merged in the IEEE SUO effort (Standard Upper Ontology). The ontologies resulting from this effort, and specially the SUMO ontology (Suggested Upper Merged Ontology; Niles and Pease, 2001), are characterized by strong logical foundations that facilitate the implementation of sophisticated reasoning mechanisms on top of them. On the other hand, however, logical subtleties might make modeling more complex and time consuming.
UFO	UFO is a reference ontology of endurants, which is based on a number of different theories such as philosophy of language, formal ontology, linguistics, cognitive psychology and philosophical logics. Since UFO is a 3D ontology, it focuses less on processes and events. In order to deal with time and changes, additions to UFO have been made. It is called UFO-B, an ontology of perdurants.

From this Web Ontology, an automated reasoner processing it using OWL semantics can infer, without any additional knowledge, that it makes sense to recommend courses about “Educational Psychology” when someone is looking for courses about “Psychology,” because the latter includes all instances of the former from a logic standpoint. Moreover, the reasoner can

follow the URI to OpenCyc to retrieve additional information about this class, for instance labels in different languages or how it is related to other concepts.

Modeling Cognition, Motivation, and Emotion Using Web Ontologies

One of the great challenges of artificial intelligence has been to conceptualize a model of human behavior using technological agents. There are different theories combining emotion and cognitive concepts. Scherer et al. (2010) classify them as discrete, dimensional and appraisal theories of emotion.

Focusing on the models of emotions used in this work, appraisal theories can be translated as evaluation or estimate. They are seeking to detail the underlying mental processes related to elicitation of emotions. That is to say, reflects the person-environment relationship, do not take into account only one aspect. This relationship is characterized by size (appraisal variables). A possible example would be: Is this a desirable event or a desired objective? Who caused it? Or do you expect? The results are mapped on emotions. Some derived models describe in detail how the resulting emotions influence individual cognitive and behavioral responses. Concepts from traditional artificial intelligence BDI models (beliefs, desires and intentions) are mapped to the dimensions from the appraisal theories. The computer models used are the Emotion and Adaptation models.

Each of them has resulted in more refined theories on emotional computer models, as shown in **Figure 1**. The following

TABLE 2 | Example of Web Ontology using the Web Ontology Language (OWL) standard.

```

1 <!DOCTYPE Ontology [<!ENTITY cyc "http://sw.opencyc.org/concept/">]>
2 <rdf:RDF
3   xmlns="http://mypsy.org/concept/"
4   xmlns:cyc="http://sw.opencyc.org/concept/"
5   xmlns:owl="http://www.w3.org/2002/07/owl#"
6   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
7   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
8
9 <owl:Ontology rdf:about="http://mypsy.org/concept/">
10   <rdfs:label>My Psychology Terms</rdfs:label>
11 </owl:Ontology>
12
13 <owl:Class rdf:ID="EducationalPsychology">
14   <rdfs:label>Educational Psychology</rdfs:label>
15   <rdfs:subClassOf rdf:resource="&cyc;Mx4rvZLWajpEbGdrcN5Y29ycA"/>
16 </owl:Class>
17
18 </rdf:RDF>

```

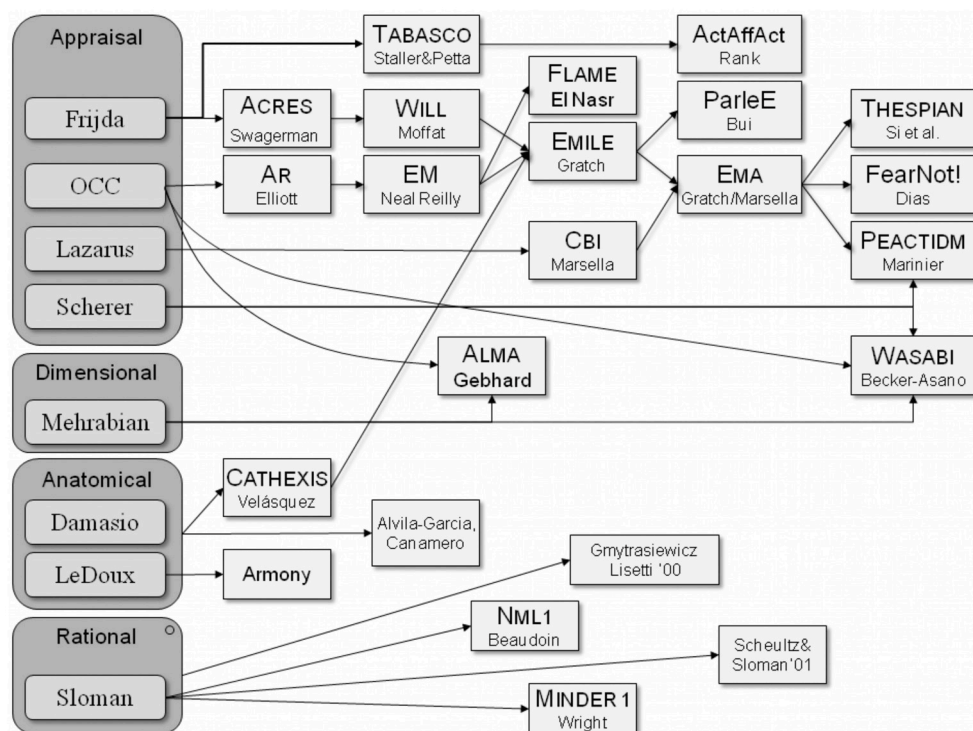


FIGURE 1 | Emotional models in artificial intelligence (Scherer et al., 2010).

issues are aspects of these models, we see that sometimes are hybrids as they include aspects of two theories in order to show that all are included in the ontology proposal (see next section). One of the referenced models, ALMA (A Layered Model of Affect), is referred because it is a hybrid model (Gebhard, 2005) as also shown in **Figure 1**.

Anatomical models emphasize the neuronal basis, so that we can talk about different ways (low-road, high-road) for the elicitation of an emotion, so that these models focus on low level of perceptual-motor tasks encoding a dual process as seen emotion (perception ontology). In the first case, there is a rapid and automatic response while, secondly, a slower response comes from reasoning processes (cognitive processes in the ontology). Consequently, affect models allow modeling cognitive or behavioral consequences, while (Chiew and Braver, 2011) also explores the motivational part related to cognitive and emotional processes.

Cognitive neuroscience aims at mapping mental processes onto brain function, which begs the question of what “mental processes” exist and how they relate to the tasks that are used to manipulate and measure them. Poldrack et al. (2011) proposed that cumulative progress in cognitive neuroscience required a more systematic approach to representing the mental entities that are being mapped to brain function and the tasks used to manipulate and measure mental processes. As a result, Cognitive Atlas¹, an ontology that characterizes the state of current thought in cognitive science was developed.

The Emotions Ontology (Hastings et al., 2014) is an ontology of emotion based on the BFO (Basic Formal Ontology) presented in Section Materials and Methods. Like BFO, this ontology is specially intended for the scientific domain and particularly to the biological sciences and human health. For instance, in combination with biology ontologies also based on BFO, it is capable of representing neurotransmitters and their influence in emotional processes and responses. Moreover, due to being based on BFO, it is a sophisticated ontology with strong logical foundations capable of modeling complex logical expressions. However, from our perspective, this makes it more difficult for people without an ontology modeling background.

Web Ontology Engineering

Starting from the previous building blocks (ontologies, technologies, theories, etc.), and in order to generate a consistent ontology that satisfies the requirement, an ontology engineering methodology has been applied. The Methontology methodology (Fernández-López et al., 1997) has been chosen because it provides guidance for ontology development process but also for other support and management activities. Moreover, it is extensively based on “classical” software engineering methodologies and this fact makes it easier to learn and apply for people with some software engineering experience.

Methontology proposes some ontology management activities, which include scheduling, control and quality assurance. There are also ontology support activities, which are performed at the same time as the development-oriented

activities, namely: knowledge acquisition, evaluation, reuse (merging or aligning other ontologies), documentation and configuration management. These are all support activities while the main ontology creation work is performed in development process.

The development process is composed by the following phases: specification, conceptualization, formalization, implementation and maintenance. The specification phase corresponds to the pre-development aspects, where the development requirements are identified. The maintenance phase is a post-development activity, which is performed once the ontology is developed. During the conceptualization activity, the domain knowledge is structured as meaningful models. Moreover, if a formal language is used to build the model, it is possible to automate the formalization and implementation activities. In our case, as OWL is a formal language, we can benefit from this feature and existing ontology development environments implementing it, like Protégé (Musen, 2015). Moreover, as it is detailed in Section Evaluation: Massive Open Online Courses, it is also possible to use automated logical reasoners to check the consistency of the resulting ontology.

RESULTS: THE EMOTIONS & COGNITION ONTOLOGY

First of all, for the development of our ontology Emotions & Cognition Ontology, the chosen knowledge representation for the contributed ontology is the Web Ontology Language (OWL), which also facilitates sharing and reusing knowledge using the already global infrastructure of the Web. This, compared to existing pre-Web ontologies, facilitates sharing and also reusing existing ontological frameworks as it is detailed next.

Our approach has been to reuse as much as possible existing ontologies, especially upper ontologies and other more specific ones related to cognition and emotion. This approach reduces the cost of developing an ontology but also strengthen it because it is based on more solid foundations, provided by already proven and widely used ontologies.

The first choice has been to reuse the upper ontology Cyc. The main motivation has been to benefit from its wide coverage. This way, it is usually possible to find amongst the 239,000 concepts it provides the required ones to model the real situations for which we want to capture their perception taking into account cognitive, motivational and emotional factors.

Basically, whenever a particular term is needed, we can search Cyc, locate the relevant concept based on the provided descriptions and relations to other concepts and, finally, refer to that concept using its reference. This is facilitated by the fact that we are using a Web ontology and that OpenCyc is also available as in that form.

However, our experience showed that beyond providing a lot of base concepts where we can root the ones introduced by our ontology, OpenCyc was too normative and rigid when trying to model the glue among these concepts that capture the parts of reality we are interested in modeling. OpenCyc is based on a strong use of logic geared toward automated reasoning

¹<http://www.cognitiveatlas.org/>

that requires a profound knowledge and effort. We required a modeling approach less abstract and more similar to what we were trying to capture, human cognition.

Based on these requirements, the clear choice was DOLCE, whose aim is precisely to capture the ontological categories underlying natural language and human common sense. DOLCE does not commit to strictly referentialist metaphysics related to the intrinsic nature of the world like 4D ontologies do. Rather, the categories it introduces are thought of as cognitive artifacts, ultimately depending on human perception, cultural imprints and social conventions. In this sense, they intend to be just descriptive notions that assist in making already formed conceptualizations explicit.

For instance, the distinction between enduring and perduring entities is simplified in DOLCE to the relation of participation: an enduring “lives” in time by participating in some perdurants. For example, a person, which is an enduring, may participate in a discussion, which is a perdurant. A person’s life is also a perdurant, in which a person participates throughout its all duration. Using this approach, we have rooted the contributed ontology on the fundamental terms provided by DOLCE. This facilitates the process of modeling real world situations and their perception. On the other hand, given the limited scope of DOLCE, when specific terms for concepts like “Psychology” are needed to build a representation, then we look into OpenCyc and refer to them using their URIs.

DOLCE is also a Web ontology. Consequently, this approach makes it easy for agents to process the new ontology, as they just need to follow its links in order to retrieve additional facts about the reused concepts. The vision of this approach is that, this way, ontologies can grow and evolve organically through the web in an unrestricted and self-organized way, like the Web did with great success. Another example of this vision is the rhizome metaphor proposed by Deleuze and Guattari (1987).

In addition to these advantages, other key features of Emotions & Cognition Ontology are:

1. The underlying conceptual model, implemented by the ontology, is independent from any specific emotion theory. It provides a set of building blocks (concepts) that are selected and combined as required in order to capture the subtleties of a particular model of cognition or emotion.
2. The ontology is capable of dealing with different emotion communication modalities. The model includes generic concepts like *Sensor* or *Emotion Expression System*, which are then refined to specific kinds, like biological (eye, taste...) or artificial sensors (camera, microphone...).
3. Reality is represented by means of different ontologies, which are used combined with proposed ontology to represent the world around us.
4. Reality models are based on the notion of context, which provides flexible and accurate ways of modeling situations, as detailed next.

Fundamental Building Blocks

The previous features are based on the use of the DOLCE upper ontology, which provides the fundamental building blocks like

the *Context* or *Sensor* concepts. Consequently, our proposed ontology extends the DOLCE upper ontology and particularly the *Description* and *Situation* concepts. *Perception* generates *Descriptions* that represent *Situations*, configurations of the real world. These *Descriptions* may trigger and be associated with emotions.

Another fundamental feature of the proposed ontology is that it does not commit to a particular emotions theory. For instance, an example of emotions ontological modeling might be just to represent using an ontology and as a taxonomy the categorical theory of emotions by Ekman (1984). However, this limits the proposed ontology just to the application of this particular emotions model.

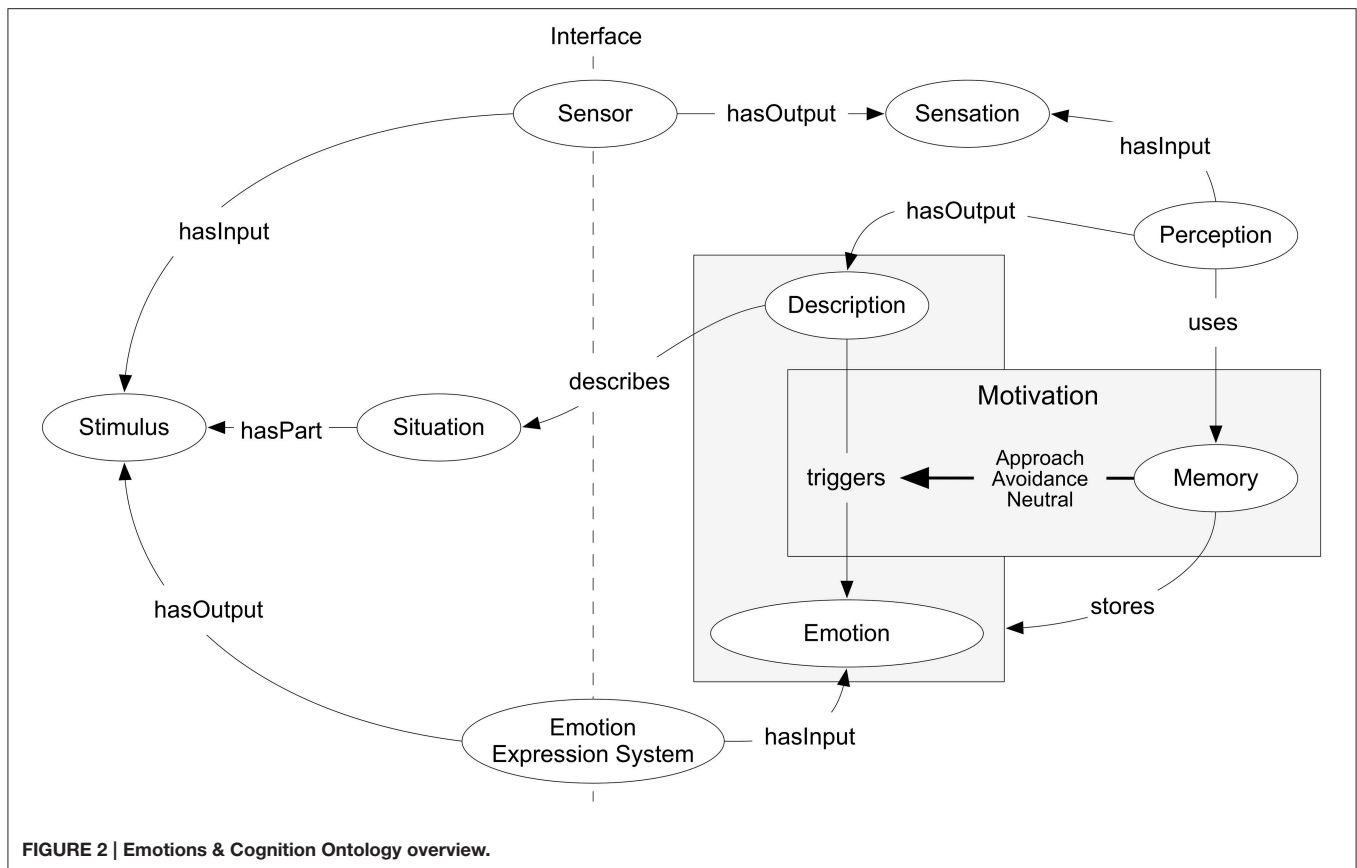
Consequently, even from the initial steps of defining an ontology for describing just emergent emotions (López et al., 2008), Emotions & Cognition Ontology has been planned as emotion theory agnostic model. Thus, it is capable of providing the required mechanisms to model reality and associate it to emotional responses, but without committing to a particular way of organizing these emotional responses. This approach is improved in this new proposal as we have now consolidated the ontological foundations provided by DOLCE and Cyc, and other resources reused to facilitate *Descriptions* modeling like FrameNet. Moreover, the ontology includes now the appraisal aspects described in Section Results: the Emotions & Cognition Ontology.

The flexibility of Emotions & Cognition Ontology is due in great measure to the inclusion of the generic concepts reused from OpenCyc and specially DOLCE, combined with the mechanisms that the ontology provides to model the interactions between an agent and its environment, something that is fundamental in emotion theories based on appraisal.

In this regard, from DOLCE we reuse the concepts of *Description* and *Situation*, which constitute the basic building blocks to model the relationship among agents and their environment, cognitive processes and motivation. The cognitive process of *Perception*, as shown in **Figure 2**, generates *Descriptions*, which are representations made by the agent holding that cognitive process of the perceived *Situations*, configurations of the real world identified by the agent. These representations, the *Descriptions*, are what the agent associates to emotions as a result of its cognitive and motivational processes.

The fundamental building blocks also include other cognitive process and related aspects, which can be used and combined depending on the particular emotion and cognition theories to work with, and the required detail level. As it is shown in **Figure 2**, the ontology also includes a generic concept *Emotion* that can be directly extended and refined. This is usually enough when working with dimensional theories of emotion. Additionally, if theories based on appraisal are considered, the ontology also provides mechanism to model context using concepts like *Perception* or *Memory*, combined with the separation between *Environment* and *Agent*.

The interface between the *Environment* and the *Agent*, from an emotions point of view, is defined by *Sensors* and the *Emotion*



Expression System. As the ontology defines *Agent* as a generic concept, which includes both human agents and artificial ones, the sensors include human senses but also artificial sensors. The same applies to the emotion expression systems under consideration.

On the other hand, the Environment corresponds to the agent “reality,” a real world or a virtual one in which virtual agents interact. The latter might include the Internet or a particular Web application like a specific social network. In this case, the Web application is what determines the available sensors and emotions expression systems, for instance emoticons.

Finally, in the *Agent* side, in addition to *Perception*, the ontology defines other cognitive processes like *Memory* and *Motivation*. These additional cognitive processes allow modeling the key aspects of the appraisal factors, from previous agent experiences that define its beliefs to the desires and intentions that configure its motivations.

All the cognitive processes have been linked to the main cognitive science ontology identified in Section Results: the Emotions & Cognition Ontology, Cognitive Atlas. For instance, *Memory* has been linked to the corresponding concept in Cognitive Atlas *memory*², which also provides access to specific kinds of memory if such level of detail is required, like *context memory* or *emotional memory*. Other relevant concepts from

Cognitive Atlas the ontology is linked to are: *perception*³ or those related to *Motivation*⁴.

Memory stores past associations of *Descriptions* and the *Emotions* triggered by the corresponding *Situations*. These memories are fed into *Motivation*, which matches the current *Description* to past memories. If the match strength is low, because the corresponding situations have little in common and consequently the associated *Descriptions* too, the motivation is *Neutral Behavior* and the effect of motivation on the triggered *Emotion* is low or inexistent.

On the contrary, if the match is high, because the *Descriptions* of the current situation and the past one are similar, then the motivation is non-neutral. In this case the effect on emotion triggering might be positive or negative, positive if the *Memory* associated the matched *Description* to a positive *Emotion*, or negative if it was a negative one. The former corresponds to an *Approach* behavior and the latter to an *Avoidance* one, from a motivational point of view.

What constitutes a positive or negative *Emotion*, and its effect on *Motivation*, will depend on the particular emotion theory to be applied and on the available mechanisms to characterize emotions, as it is detailed in the practical scenario described in Section Conclusions.

²*Memory* in Cognitive Atlas, <http://www.cognitiveatlas.org/concept/memory>

³*Perception* in Cognitive Atlas, <http://www.cognitiveatlas.org/concept/perception>
⁴*Motivation* category in Cognitive Atlas, <http://www.cognitiveatlas.org/concepts/category/Motivation>

Once the previously introduced ontological building blocks have been described, the next subsection details how the main building blocks to model reality, i.e., *Descriptions*, are defined. These descriptions capture the aspects of reality considered relevant by the agent and link them to emotional responses.

Modeling Descriptions of Reality

In order to cope with the enormous range of different situations that might need to be associated with emotions, their descriptions are modeled using concepts from OpenCyc, as previously introduced. In addition to OpenCyc concepts, we have also included terms from FrameNet (Scheffczyk et al., 2008). This is not an upper ontology but a big lexical database, with more than 10,000 word senses, structured following Frame Semantics (Fillmore, 2006).

Frames fit really well with situations modeling as they try to explain words meaning by building a description of a type of event, relation, or entity and the participants in it. This way, DOLCE provides the ontological foundations, FrameNet the glue to structure situations modeling and OpenCyc the anchors to the specific concepts involved in situations and descriptions modeling. Section Conclusions provides examples of how these ontologies are reused to this end. These examples will use the STUDYING frame, which is presented in Table 3.

For a complete list of all the concepts defined in Emotions & Cognition Ontology, it is available online as a Web Ontology⁵. In the next section, this ontology is put into practice in a particular scenario, online education.

EVALUATION: MASSIVE OPEN ONLINE COURSES

The aim of this section is to evaluate Emotions & Cognition Ontology in the context of a real use case. Web ontologies can be evaluated from a purely logical standpoint using a reasoner capable of processing OWL. There are many OWL reasoners available and we have used one of them, Pellet (Sirin et al., 2007) to validate the consistency of the ontology. However, the consistency of an ontology is just a lightweight evaluation. We have the guarantee that it is not going to generate contradictory conclusions but we don't know if it is going to be capable of modeling the required knowledge and producing the expected conclusions.

For this kind of evaluation, it is necessary to put the ontology in use in a real or simulated scenario. We have applied previous versions of this ontology to gather emotional common sense (Gil et al., 2015) and in the context of tangible user interfaces (López-Gil et al., 2014). More recently and as detailed next, we have also started to apply the enhanced version including cognition in the context of online education and Massive Open Online Courses (MOOCs).

MOOCs are a recent development in distance education that allow the participation of a big amount of users and that can be accessed using the Internet. They have become popular since

TABLE 3 | FrameNet description of the frame STUDYING.

STUDYING

Definition: A Student enrolls and then remains at an Institution for the purposes of education within a Subject. They may receive instruction from a particular Teacher at the Institution.

Frame Elements

Core:

- **INSTITUTION:** an educational establishment, such as a school or college.
- **STUDENT:** one who receives instruction from a Teacher or Institution.
- **SUBJECT:** the area of knowledge or skill which is taught to a Student.
- **TEACHER:** one who instructs a Student in some area of knowledge or skill.

Non-Core:

- **CO-PARTICIPANT:** an entity that participates in a coordinated way in the event.
- **DEPICTIVE:** a state which describes the Student during the process of study.
- **DURATION:** the amount of time for which the process of study is ongoing.
- **EXPLANATION:** the reason for which the Student studies the Subject.
- **INTENDED ACADEMIC CREDIT:** the Student's motivation for studying.
- **LEVEL:** identifies the Level of a student in his/her education.
- **MANNER:** the manner of studying that the Student has.
- **PLACE:** the Place is the location within which the studying takes place.
- **TIME:** the time when the student is studying.

2012, when some courses vendor platforms such as Coursera, in which prestigious universities participated, emerged. In addition to course materials, such as videos with lectures, readings or sets of problems, MOOCs also provide interactive forums and online communication tools to promote interaction between students and teachers.

Despite their popularity, the MOOCs also have disadvantages and associated challenges, including that relying on user-generated content can generate a chaotic learning environment, necessary knowledge of the online platform to make appropriate use of it, the time and effort required by the participants, the difficulty of controlling the course trajectory once it has been released due to the amount of different students and self-regulation of users to obtain the expected educational benefit. All these aspects are strongly influenced by the characteristics of the users and their expectations, which may result in different emotional and motivational states depending on how the course is elapsing.

In this type of systems motivation is especially important. It is an important factor to improve the performance of students and to improve the ratio of pupils that successfully complete the courses that are enrolled in. In addition, modeling the emerging emotions that a person expresses is also important in such environments in order to learn how they are feeling.

Emotions & Cognition Ontology can represent different emotions expression systems that can be used in these environments as a basis for recognizing the emotions of users and also to analyze their motivation. In the case of MOOCs, the agent expressing emotions is the human being, so different emotions expression systems are considered, including natural language,

⁵<http://rhizomik.net/ontologies/emotions&cognitionontology/>

speech, facial expressions, and even galvanic skin response, brain activity, heart rate, or blood pressure.

However, before we can start representing emotional responses, we need to model the real world situations to which they are associated. As presented in Section Evaluation: Massive Open Online Courses, we will use DOLCE *Descriptions* as the representations of the real world *Situations*. Moreover, we are going to use FrameNet frames and OpenCyc concepts to provide the required level of detail to these *Descriptions*.

In our scenario, one frame that is particularly relevant is the one shown in **Table 3** in the previous section, the frame STUDYING. We will use this frame to illustrate how the ontology can be used to model a *Description*. For instance, a situation in this scenario might be “*The second grade student Stuart Adams has been studying educational psychology online course for 2 h today.*” The *Description* for this situation is based on the STUDYING frame, where the frame elements are filled with different parts of the situation as follows:

[STUDYING] → [STUDENT: **Stuart Adams**]
 → [LEVEL: **second grade**]
 → [SUBJECT: **educational psychology**]
 → [MANNER: **online course**]
 → [DURATION: **2 hours**]
 → [TIME: **today**]

The idea is that, given the previous *Situation*, an agent perceives it through its sensors, sense in the case of a human agent, and its cognitive processes generate the corresponding *Description*, the representation that the agent builds for its environment. With the ontology, and for the online education scenario, the objective is to try to mimic this behavior so we can first model the context of the student being monitored, and then associate an emotional response so we can improve the student experience.

However, before we connect the *Description* to emotions, we can detail it further, going beyond the use of FrameNet frames. The *Description* can be enriched with concepts from Web ontologies like OpenCyc, which help defining, for instance, what “educational psychology” refers to. It might be the case this particular concept is not present in OpenCyc, we can then define it as we did in the example in Section Evaluation: Massive Open Online Courses as a subconcept of “psychology,” which is defined in OpenCyc as shown in **Table 4**.

This way, the frame element for the previous description can be further detailed to:

→ [SUBJECT: <http://myspy.org/concept/EducationalPsychology>
subConceptOf
<http://sw.opencyc.org/concept/Mx4rvZLWajwpEbGdrcN5Y29ycA>]

In this case, the refinement makes it less ambiguous what Stuart Adams is studying, specially from the point of view of an artificial agent who has to monitor and respond to his emotional responses, for instance to adjust the learning pace of this particular student. It might be the case that the agent does not have any particular knowledge about the concept EducationPsychology, but it can at least follow its definition as

TABLE 4 | OpenCyc definition for the concept Psychology

OpenCyc Individual: **Psychology**

Unique ID: [Mx4rvZLWajwpEbGdrcN5Y29ycA]

English ID: [Psychology]

The **FieldOfStudy** of psychology.

• Instance of: **FieldOfStudy**

• Wikipedia: <http://en.wikipedia.org/wiki/Psychology>

• Same as:

• <http://umbel.org/umbel/sc/Psychology>

• http://data.linkedmdb.org/resource/film_subject/369

• <http://dbpedia.org/resource/Psychology>

a subconcept of the OpenCyc concept for Psychology, and from there learn that it is a field of study also defined in other Web ontologies such as the DBpedia, which is the Web ontology version of Wikipedia.

This refinement allows the agent to recognize that the current subject is related to previous ones he has also studied, which appeared to be especially frustrating for the user given the records of existing descriptions and emotional responses. Consequently, it might be convenient to anticipate and adjust the pace to improve the learning experience in this case. This is supported by another cognitive process also modeled by the ontology, *Memory*. The MOOC agent can use these sensitive memories to represent past *Descriptions* and their associated *Emotions*.

This way, we can use *Memories* to also model *Motivation* and its influence on the emotional response to the current *Situation*. The proposed approach is based on matching the *Description* for the current situation with *Descriptions* associated to past *Memories*, which in the case of the MOOC agent will be stored by the agent to try to anticipate the motivations of the student.

As mentioned in the case of the Educational Psychology subject, the matching does not need to be direct. Though the current *Description* refers to the concept *EducationalPsychology* and there are no previous memories with descriptions referring to this same subject, it might be the case that there are past memories that are associated to subject related to OpenCyc concept *Psychology* like *EducationalPsychology*.

Therefore, we need a matching algorithm that takes into account, for instance, the amount of concepts that the compared descriptions share. However, this algorithm can be further sophisticated to take into account structure and semantics (Gallagher, 2006). In any case, what is needed is a mechanism to derive for each memory its behavior associated to *Motivation*.

First of all, if it is mainly a *Neutral Behavior*, for instance if the current description and the past one just share a small amount of concepts. If the behavior is not neutral, then two behaviors can be derived: *Approach* and *Avoidance*. The former, if the recorded emotional response for the memory had at least a positive valence, is even clearer if the arousal was also positive. On the other hand, the behavior anticipated by *Avoidance* would be clearer if the response had a negative valence.

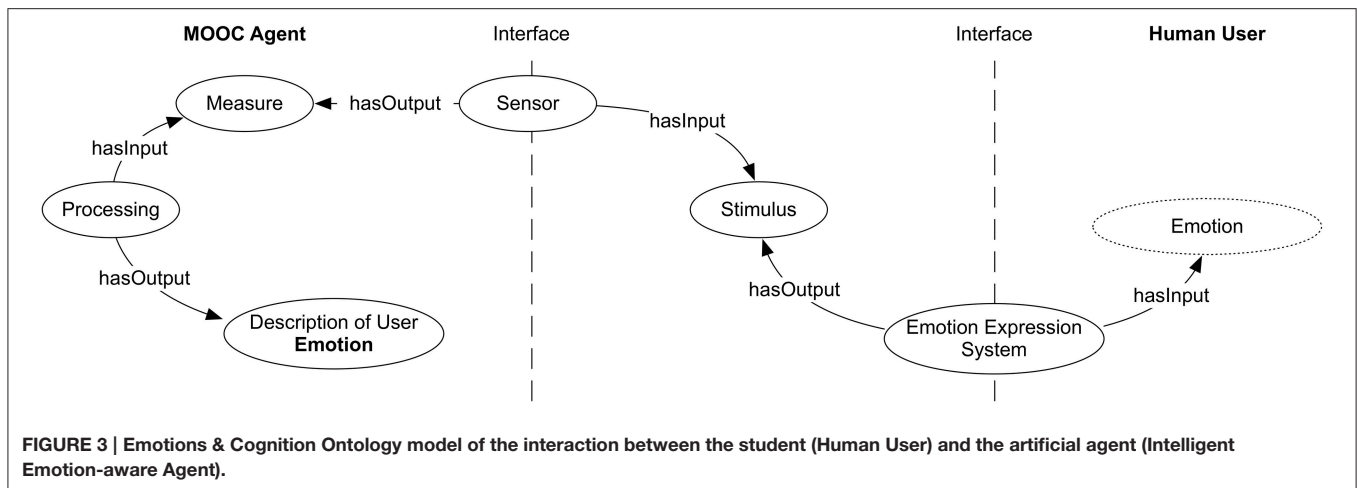


FIGURE 3 | Emotions & Cognition Ontology model of the interaction between the student (Human User) and the artificial agent (Intelligent Emotion-aware Agent).

Thus, the MOOC agent will use the combined set of behaviors for the non-neutral motivations to try to anticipate the motivation of the student for the current situation. For instance, if there is *Memory* for a subject is related to OpenCyc *Psychology* with an emotional response characterized by a positive valence and arousal, as registered by available sensors, the agent can infer a non-neutral motivation of *Approach*. This information can be then used to adjust the student pace.

The previously mentioned sensors, used by the agent, can be also modeled using Emotions & Cognition Ontology. The agent uses them to monitor emotions expression systems of students, as shown in **Figure 3**, where the agent is displayed in the left side and the student on the right. The environment in this case, represented in the center of the figure, consists of the MOOC system itself and the physical environment in which the student interaction is carried out.

The range of sensors available to the MOOC agent will depend on the devices available to each student. Our current experimentation setting includes the following devices:

- **Microphone:** Captures vocal parameters and natural language.
- **Keyboard:** Captures the natural language.
- **Webcam:** Captures facial expressions.
- **Wristband:** Wristband with sensors to capture the galvanic skin response, heart rate and blood pressure.
- **Neuroheadset:** Headset with sensors to capture brain activity by gathering data from EEG channels.
- **Eye Tracker:** Captures the focus of the user within the given user interface.

All the information provided by the sensors feeds the artificial agent. With the Emotions & Cognition Ontology, we are able to set what the agent needs at the conceptual level, which is the aim of this paper. From now on, we need to face what can be called the emotion semantic gap between the signals captured by the sensors the agent includes and the conceptual representation of the recognized emotion at the conceptual level, i.e., as an *Emotion*.

We have already tested the feasibility of this approach for some of the sensors in our experimentation setting. For instance, we have used a combination of techniques and applications to infer the emotional state of the student and render it using valence and arousal dimensions, following the PAD model mentioned in Section Results: the Emotions & Cognition Ontology. The electro-physiological experiments were carried out according to the principles of the declaration of Helsinki and approved by the ethics committee on clinical research of the Arnau de Vilanova University Hospital. With appropriate software for processing EEG and other electrophysiological data, such as EEGLab⁶, we process the neuroheadset signal and derive the arousal from the EEG signal, while the valence is derived from a wristband. These values are fed to the MOOC agent so it can associate them to the *Description of the Situation* the student is faced at that particular moment. A sample dataset based on the Emotions & Cognition Ontology for the MOOC scenario presented in this section is available online⁷.

CONCLUSIONS

This article presents an ontology for linking reality with its perception by human beings. As the emotional state of people is important as it modifies their perception of the world, it is important not only to adequately describe categorically structured ways of understanding the world around us, but also to describe the emotional, cognitive and motivational processes of people to understand how they perceive and interpret the world around them. Besides, as both descriptions of reality and emotion, cognition and motivation are modeled by means of ontologies, all this knowledge is shared in a common framework.

It is also shown how proposed ontology has been used in the field of MOOCs, environments where adequate representation of emotions and motivation is especially important to ensure its success among people who use them. Testing the ontology in real

⁶<http://sccn.ucsd.edu/eeglab/>

⁷Sample dataset for the MOOC scenario: <http://rhizomik.net/ontologies/2015/12/mooc-sample.ttl>

scenarios has allowed the validation of one of the main aims of the ontology: it is relatively simple to apply even for non-experts in ontology modeling.

Additionally, the rest of the intended contributions detailed in Section Fundamental Building Blocks have been also addressed. First, as the scenarios are based on different theories of emotion, it has been possible to test that the Emotion & Cognition Ontology provides the required building blocks to accommodate these different views, from discrete or dimensional theories considered in the context of Tangible User Interfaces (TUI) (López-Gil et al., 2014) to those based on appraisal, as illustrated in this paper. The scenarios have also shown that the ontology is capable of dealing with different emotion communication modalities, from physical sensors and emotion expression systems, available in the case of TUIs, to virtual ones like in the case of MOOCs.

On the other hand, the approach based on Web Ontologies has facilitated reusing many different ontologies from upper ontologies like DOLCE or OpenCyc to the reuse of frames from FrameNet, which has considerably reduced the modeling effort. The latter has also highlighted the advantages of including the notion of context in the core of the ontology through DOLCE's Descriptions and Situations, which have been smoothly connected with the notion of frames to facilitate Descriptions modeling.

All these findings corroborate the contribution beyond existing emotions ontology, specially comparing to the Emotions Ontology (Hastings et al., 2014), which is the main ontology in this field and was introduced in Section Results: the Emotions & Cognition Ontology. Emotions Ontology is more sophisticated than the proposed one, as it is based on an upper ontology that makes use of logic formalisms to enable elaborate reasoning. However, this introduces too much unnecessary complexity when working with simpler emotion theories or scenarios where just a simple modeling of emergent emotions is required. Moreover, Emotions Ontology is quite tied to a vision of emotion based on the concept of appraisal.

REFERENCES

- Blackburn, S. (1996). *The Oxford Dictionary of Philosophy*. Oxford; New York, NY: Oxford University Press.
- Chiew, K. S., and Braver, T. S. (2011). Positive affect versus reward: emotional and motivational influences on cognitive control. *Front. Psychol.* 2:279. doi: 10.3389/fpsyg.2011.00279
- Deleuze, G., and Guattari, F. (1987). *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis, MN: University of Minnesota.
- Ekman, P. (1984). "Expression and nature of emotion," in *Approaches to Emotion*, eds K. Scherer and P. Ekman (Hillsdale, NJ: Erlbaum), 319–343.
- Fernández-López, M., Gómez-Pérez, A., and Juristo, N. (1997). "METHONTOLOGY: from ontological art towards ontological engineering," in *Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series* (Palo Alto, CA).
- Fillmore, C. J. (2006). Frame semantics. *Cogn. Ling. Basic Read.* 34, 373–400. doi: 10.1016/b0-08-044854-2/00424-7
- Gallagher, B. (2006). Matching structure and semantics: a survey on graph-based pattern matching. *AAAI FS* 6, 45–53.

As technology advances, different types of sensors are available to gather information about people's emotions, cognitive processes and emotions, such as EEG, heart rate, electrodermal activity, facial expressions or speech. Still, such information is not enough by itself to determine such complex aspects and must be considered in the frame of established models and theories. The ontology formalizes a common view about how theories and models are mapped, which are then used to facilitate data integration. If these mappings would not be provided by the ontology, the semantic gap would remain and interoperability among ontology components would be seriously compromised.

Neuroscience is advancing day by day in the knowledge of how humans manage emotions. There are many emotional computing models that relate the abstract concepts that included in the ontology. However, in the not too distant future, emotions will not be restricted to humans, as it seems that machines and virtual agents in general will also be able to recognize and synthesize emotions. Exposed line of work aims to set out a general framework for all kinds of emotional interactions, including the ones with such emotion-aware devices and virtual agents.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

This work has been supported by the Basque Government (IT421-10 and IT722-13), the Gipuzkoa Council (FA-208/2014-B) and the University of the Basque Country (PPV12/09). It has also been supported by InDAGuS (Spanish Government TIN2012-37826-C02) and INSPIRES, the Polytechnic Institute of Research and Innovation in Sustainability, Universitat de Lleida, Spain.

- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). "Sweetening ontologies with DOLCE," in *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. LNCS, Vol. 2473* (Heidelberg: Springer), 166–181.
- Gebhard, P. (2005). "ALMA: a layered model of affect," in *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems* (New York, NY: ACM), 29–36.
- Gil, R., Virgili-Gomá, J., García, R., and Mason, C. (2015). Emotions ontology for collaborative modelling and learning of emotional responses. *Comp. Hum. Behav.* 51, 610–617. doi: 10.1016/j.chb.2014.11.100
- Guarino, N. (1998). "Formal ontology and information systems," in *Formal Ontology in Information Systems*, ed N. Guarino (Amsterdam: IOS Press), 3–15.
- Hastings, J., Frishkoff, G. A., Smith, B., Jensen, M., Poldrack, R. A., Lomax, J., et al. (2014). Interdisciplinary perspectives on the development, integration, and application of cognitive ontologies. *Front. Neuroinform.* 8:62. doi: 10.3389/fninf.2014.00062
- Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P. F., and Rudolph, S. (2009). *OWL 2 Web Ontology Language Primer. W3C Recommendation*. Cambridge, MA: World Wide Web Consortium.

- Lazarus, R. S. (1991). Cognition and motivation in emotion. *Am. Psychol.* 46, 352–367. doi: 10.1037/0003-066X.46.4.352
- Lenat, D. B. (1995). CYC: a large-scale investment in knowledge infrastructure. *Commun. ACM* 38, 33–38. doi: 10.1145/219717.219745
- López, J. M., Gil, R., García, R., Cearreta, I., and Garay, N. (2008). Towards an ontology for describing emotions. *WSKS* 5288, 96–104. doi: 10.1007/978-3-540-87781-3_11
- López-Gil, J. M., Gil, R., García, R., and Collazos, C. A. (2014). EmotionsOnto: an ontology for developing affective applications. *J. Univ. Comp. Sci.* 20, 1813–1828. doi: 10.3217/jucs-020-13-1813
- Musen, M. A. (2015). The Protégé project: a look back and a look forward. *AI Matt.* 1, 4–12. doi: 10.1145/2757001.2757003
- Niles, I., and Pease, A. (2001). “Towards a standard upper ontology,” in *FOIS '01: Proceedings of the International Conference on Formal Ontology in Information Systems*, eds C. Welty and B. Smith (New York, NY: ACM Press), 2–9.
- Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., et al. (2011). The cognitive atlas: toward a knowledge foundation for cognitive neuroscience. *Front. Neuroinformat.* 5:17. doi: 10.3389/fninf.2011.00017
- Sartre, J.-P. (1939). *Sketch for a Theory of the Emotions*. New York, NY: Philosophical Library.
- Scheffczyk, J., Baker, C. F., and Narayanan, S. (2008). “Ontology-based reasoning about lexical resources,” in *Ontologies and Lexical Resources for Natural Language Processing*, eds N. Calzolari, A. Gangemi, A. Lenci, A. Oltramari, and L. Prévot (Cambridge: Cambridge University Press), 1–8.
- Scherer, K. R., Bänziger, T., and Roesch, E. B. (2010). *Blueprint for Affective Computing: A Sourcebook*. Oxford: Oxford University Press.
- Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., and Katz, Y. (2007). Pellet: a practical OWL-DL reasoner. *Web Semant.* 5, 51–53. doi: 10.1016/j.websem.2007.03.004
- Studer, R., Benjamins, R., and Fensel, D. (1998). Knowledge engineering: principles and methods. *Data Knowl. Eng.* 25, 161–197. doi: 10.1016/S0169-023X(97)00056-6
- Conflict of Interest Statement:** 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.

Copyright © 2016 López-Gil, Gil and García. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Free Associations Mirroring Self- and World-Related Concepts: Implications for Personal Construct Theory, Psycholinguistics and Philosophical Psychology

Martin Kuška^{1*}, Radek Trnka^{1,2}, Aleš A. Kuběna³ and Jiří Růžička¹

¹ Science and Research Department, Prague College of Psychosocial Studies, Praha, Czech Republic, ² Faculty of Humanities, Charles University in Prague, Praha, Czech Republic, ³ Department of Econometrics, Institute of Information Theory and Automation of the Czech Academy of Sciences, Praha, Czech Republic

OPEN ACCESS

Edited by:

Paul M. W. Hackett,
University of Oxford, UK

Reviewed by:

Elena Clare Cuffari,
Worcester State University, USA

Rossella Guerini,
University of Roma Tre, Italy

*Correspondence:

Martin Kuška
kuska@volny.cz

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 01 April 2016

Accepted: 14 June 2016

Published: 29 June 2016

Citation:

Kuška M, Trnka R, Kuběna AA
and Růžička J (2016) Free
Associations Mirroring Self-
and World-Related Concepts:
Implications for Personal Construct
Theory, Psycholinguistics
and Philosophical Psychology.
Front. Psychol. 7:981.
doi: 10.3389/fpsyg.2016.00981

People construe reality by using words as basic units of meaningful categorization. The present theory-driven study applied the method of a free association task to explore how people express the concepts of the *world* and the *self* in words. The respondents were asked to recall any five words relating with the word *world*. Afterward they were asked to recall any five words relating with the word *self*. The method of free association provided the respondents with absolute freedom to choose any words they wanted. Such free recall task is suggested as being a relatively direct approach to the respondents' self- and world-related conceptual categories, without enormous rational processing. The results provide us, first, with associative ranges for constructs of the *world* and the *self*, where some associative dimensions are defined by semantic polarities in the meanings of peripheral categories (e.g., Nature vs. Culture). Second, our analysis showed that some groups of verbal categories that were associated with the words *world* and *self* are central, while others are peripheral with respect to the central position. Third, the analysis of category networks revealed that some categories play the role of a transmitter, mediating the pathway between other categories in the network.

Keywords: personal construct theory, free association, world, self, psycholinguistics, philosophical psychology

INTRODUCTION

Theoretical concepts that were developed a few decades ago are often understood rather as theoretical foundation stones and may be even seen as outdated in some cases. Sometimes, however, situations occur when some older theory is still very influential and inspiring for the field. In psychology, this is the case with Kelly's (1955) personal construct theory (PCT), to which the attention of researchers is turned rather frequently. For instance, Home et al. (2010) point out that PCT provides "the theoretical framework to learn which constructs the respondents themselves use, but have possibly not yet articulated, and which avoids introducing constructs that stem from the researcher" (Home et al., 2010, p. 503); Cridland et al. (2014) conclude in their research article on adolescents with high-functioning autism (HFA) that "PCT provides an eloquent and in-depth account of developmental issues for adolescents with HFA"

and “propose PCT as efficacious in doing justice to the complexity of this condition during the particular challenging period of adolescence” (Cridland et al., 2014, p. 116); Horley and McPhail (2009) interpret terrorism from the PCT perspective and name PCT as a “well-established if rather unconventional psychological theory” (Horley and McPhail, 2009, p. 119), which enables understanding an issue through the use of construct systems. Recently, there have been apparent efforts among several authors to update and utilize the timeless thoughts of Kelly and his followers, Bannister and Mair (e.g., Bannister and Mair, 1968) or Fransella (e.g., Fransella, 2005).

In dealing with the meanings of words, more precisely: in measuring the meanings of words, the term semantic space is widely employed and two basic ways of delimiting it are distinguished: (1) the traditional one, which is word-based and focused on the co-occurrence of words (e.g., Lowe, 2006), and (2) the syntax-based, which reflects the mutual relations between words (e.g., Geeraerts, 2010). Kelly’s PCT is focused primarily on a person who is actively engaged in giving meaning to the *world* and the *self*. After many years Kelly’s contribution is still considered to be radical, because his personal constructs psychology means abandoning the mechanistic and reductionist traditions in psychological thinking, and it fits comfortably into more recent developments aiming to see man in a holistic perspective (Winter, 2012).

The goal of the current paper is to explore how people perceive and understand the *world* and the *self*, or, more specifically, how they associate these two given terms employed as cue words in a free association task. The process of associating is comprehended as verbal constructing. As Kelly states, “there are always some alternative constructions available to choose among in dealing with the world” (Kelly, 1991, p. 11). Associations of the terms *world* and *self* acquired in the present study are considered to be basic meaning elements of what Kelly calls constructs: “man creates his own ways of seeing the world in which he lives . . . man might be seen as an incipient scientist . . . each individual man formulates in his own way constructs through which he views the world of events” (Kelly, 1991, p. 9). Kelly also highlights the importance of how people construct and understand the self: “The self is, when considered in the appropriate context, a proper concept or construct. It refers to a group of events, which are alike in a certain way and, in that same way, necessarily different from other events. The way in which the events are alike is the self. That also makes the self an individual, differentiated from other individuals” (Kelly, 1991, p. 91). Categorizing and further analysis of associations of the words *world* and *self*, both of which are fundamental terms, became the key to our study, inspired by Kelly’s PCT.

PERSONAL CONSTRUCT THEORY

If we wish to think about what brought us to remove the dust from Kelly’s thoughts six decades after their first publication, let’s begin with a quotation: “We start with a person. Organisms, lower animals, and societies can wait. We are talking about

someone we know, or would like to know—such as you, or me. More particularly, we are talking about that person as an event—the processes that express his personality” (Kelly, 2003, p. 7). These words open a perspective in which the individual is in the center. But Kelly gets even closer – he is more personal or intimate. Moreover, he turns this perspective inside out and describes a man as his own “scientist.” And this is not only a particular man as the scientist – “The aspirations of the scientist are essentially the aspirations of all men” (Kelly, 1991, p. 30). If we believe this claim that Kelly makes, the respondents in the present study, when writing their associations, gain the possibility of approaching themselves in this manner. “Thus, just as the experimental scientist designs his experiments around rival hypotheses, so each person designs his daily explorations of life around the rival hypotheses which are suggested by the contrasts in his construction system” (Kelly, 1991, p. 90). The associations gathered in the present study are considered to be the basic elements of these constructs.

Kelly named the core of PCT as the “Fundamental Postulate: a person’s processes are psychologically channelized by the ways in which he anticipates events” (Kelly, 1991, p. 32). Kelly states this postulate not as a dogmatic idea, but rather as a thought-provoking one. “The new outlook which a person gains from experience is itself an event; and, being an event in his life, it needs to be construed by him if he is to make any sense out of it” (Kelly, 1991, p. 55). To associate terms such as *world* and *self* means to come out from individual experience, and, based on that experience and through the verbal categories, to show how the *world* and the *self* may be constructed. Kelly explains the terms channelized and anticipating events as depicting the dynamism of human processes and as cutting off from former psychological stimulus-response determinism (Kelly, 1991).

Kelly himself points out that PCT is more a metatheory than a theory. Ten years after the first publication of his book (Kelly, 1955) he also summarized how PCT was categorized among other authors: as cognitive, existential, emotional, learning, psychoanalytic, behaviouristic, pragmatic, reflective and many other theories, including “no theory at all. It has also been classified as nonsense, which indeed, by its own admission, it will likely some day turn out to be” (Kelly, 2003, p. 8). A period critical evaluation of PCT in a submission by Kelly’s direct followers, Bannister and Mair, was published by Eysenck (1968) in *Nature*. Eysenck’s (1968, p. 99) review summarizes the imperfections of PCT, namely, that some significant claims of PCT were “achieved purely on a verbal plane.” On one hand, the lack of empirical evidence, which is criticized as an imperfection, is understandable. On the other hand, results based strictly on quantitative empirical evidence can generate conclusions such as, e.g., “the more psychotherapy, the smaller the recovery rate” (Eysenck, 1952, p. 322), which is Eysenck’s famous conclusion from his research on the issue of (in-) effectiveness of psychotherapy, published in Eysenck (1952). Such an argument implicitly contains a question about whether philosophizing in the area of psychology should be practiced at all. In addition, Kelly explicitly delimits his position as philosophical and calls it constructive alternativism. “The best we can ever do is project our anticipations with frank uncertainty

and observe the outcomes in terms in which we have a bit more confidence” (Kelly, 2003, p. 5). This statement also explains why the task of free associations was employed as a data-mining method in the present study.

METHODOLOGY

Subjects

Data were collected from university students, university graduates and young adults of similar age. The majority of respondents were young people, often college students (68%, age between 18 and 33), educated mostly in the humanities (41%), technical or economic fields (29%), and natural science or medicine (9.6%). The final sample consisted of 251 respondents (156 women and 95 men); the median age was 26 years old. Respondents were recruited from Czech universities of different types and from random data collection in Prague, the capital of the Czech Republic. Participation in the study was fully voluntary with no incentives provided for participation. All respondents were Czech native speakers. The research design was approved by the institutional ethics committee. All respondents signed an informed written consent with their participation in the study.

Materials and Procedure

A free association task (Nelson et al., 2004) with multiple associations to the cue words was used in the present study. The main aim was to explore the conceptual categories that respondents retrieve when they are asked to recall associations to the word *world* and the word *self*. In the free association task, the respondents are asked to write the first word that comes to mind that is related or strongly associated to the presented cue word. The method of free association provides the respondents with absolute freedom to choose any words they wanted. Such free recall is suggested as being a relative direct approach to the respondents' self- and world-related conceptual constructions.

At the beginning, respondents filled in basic demographical characteristics. Then the cue words were presented to respondents on a sheet of paper. Two cue words were used: the word *world* and the word *self*. The respondents were asked to recall any five words relating with the Czech word for *world* with the instruction “Please, write five words that you recall when you hear the word *world* to the following five lines.” Five short horizontal lines placed below each other were offered to respondents below the instruction. Afterward respondents were asked to recall any five words relating with the Czech word for *self* with the same instruction and response design. The time for writing the associations was not limited.

The present research went the opposite way when compared with the method of controlled word association test (e.g., Woodworth, 1921) in that we did not apply predefined categories from which the respondents had to choose particular words. The aim of the present study was to categorize all of the acquired associations by their prevailing meanings into groups in order to subsequently display the distribution of categories and inter-category relations.

In the sense of delimiting the semantic space, the method of the present study is word-based, and because we work with particular associations which were not articulated in form of sentences or narrations, we also visualize the relations between words other than the grammatical. Jackson and Bolger (2014, p. 12) suggest in their recent article about a high-dimensional graph of semantic space that “perhaps the method of calculating relationship, rather than representing relationships, is what differs between relationship types.” We suggest trying to avoid the use of the term “semantic space” in the present study, with respect to the employed method of coding, which reflects only one prevailing meaning of an association in the given context. In this way, each categorized association represents a one-directional vector, which creates no space *per se*. This position seems to be quite radical, especially in the context of contemporary computational psycholinguistics, where the number of dimensions of semantic space has no limit and is usually calculated with multi- or high-dimensionality. Also, the authors of this article would otherwise prefer a multidimensional space construction, if appropriate, as was used in their recent article dedicated to emotional space (Trnka et al., 2016). But here, instead of dealing with the semantic space of particular words or collocations, we (1) describe and explain the meaning of each category of associations and (2) reconstruct the distribution of categories in two dimensions, separately for associations of *world* and *self*.

Data Analysis

Acquired data in form of individual words or short phrases were analyzed separately for the cue words *world* and *self*. Firstly, the data were sorted by frequency of occurrence of particular words. No data pooling was employed, so singulars and plurals, various verb tenses etc. remain intact as gathered. The most frequent words with particular meanings were suggested as category names. Secondly, data were categorized by the meaning of each association. For associations of both cue words, *world* and *self*, two separate categorization systems were developed. The exhaustive categorization process left no associations uncategorized. All categories were derived from an examination of the data (e.g., Aylwin, 1977). In the case, when a word can be assigned to more than one category, the prevailing meaning of the given word was employed. An independent rating scored an almost 82% match in the ex-post sorting of the associations into categories, namely 82.4% for *world* associations ($\kappa = 0.913$) and 81.6% for *self* associations ($\kappa = 0.909$), both at $p < 0.001$ level.

For both cue words, *world* and *self*, we elaborated two 251×18 incidence matrices. In each matrix related to the mentioned cue words, rows represented respondents and columns represented particular categories. The cell gained the value $a_{nk} = 1$ when the respondent n reported at least one association of the category k , and the value $a_{nk} = 0$, when the respondent reported no association. Incidental matrices act as an entry for two types of analysis borrowed from psychometrics and multivariate statistics. These two methods are designated for mapping statistical connections of the mutual occurrence of particular categories. The first method is factor analysis (PCA)

(Kline, 2013), and the second is the technique of Inverse covariance matrix (Whittaker, 1990). The output of factor analysis (PCA) is the mapping of each item to the smaller Euclidean space (Überla, 1968), which is in psychometrics significantly less dimensional due to correlation proximity. Coordinates in this space allow each particular item to be described as the vector of several factors with sufficient individual exactness maintained.

The goal of an inverse covariance matrix is to distinguish direct statistical dependencies of particular items (categories) from transmitted, indirect correlations. The output is a network of direct statistical dependencies on the level of a statistically typical respondent, which allows a pathway (or more pathways) of inter-category dependencies to be reconstructed between more indirectly dependent categories. The number of factors in the factor analysis was defined using the technique of point of curve-break of descending eigenvalues. For both cue-words, the corresponding break-value of the eigenvalue was 1.5 – the factor was accepted when it explained at least 50% more variance than the mean value allocated to the particular category. For both of the cue words, *world* and *self*, this allowed a given items to be mapped by two factors. Varimax rotation was used for the resulting two-dimensional maps.

RESULTS

Associations and Category Frequencies

Two discrete categorization systems were developed, one for associations of the cue word *world*, second for associations of the cue word *self*. Both systems consist of 18 categories (see **Tables 1** and **2**).

Both of these descriptions (in the Czech language) have also been provided to an independent rater for the ex-post sorting of the associations into categories.

The word *world* was associated most strongly to words from categories which relate to nature's essence of the world in a local, global or cosmic dimension. Special meaning was dedicated to the element water. Civilization and its achievements – or an anthropocentric paradigm – were sorted in more less extensive categories, with respect to different meanings of associations.

The word *self* strongly reflected individualism, including personal settings and various states of being. All associations were sorted distinctively in particular categories by their meanings, as **Table 2** displays more in detail.

The Pareto chart shows on one side the frequencies of the most commonly appearing categories, while on the other side it shows that after the initial most common categories, which for the cue word *self* are Individual, Trait, Existence, Embodiment and Social, the level of 60% was achieved within the named categories. It also shows that other categories, starting with Family, appeared rather rarely (see **Figures 1** and **2**).

PCA Analysis

To display the various ranges of how people associate the world and themselves, PCA data analyses were conducted, separately for the cues *world* and *self*.

The component plot in a rotated space chart for *world* associations (**Figure 3**) shows the distribution of categories in two basic dimensions (ranges). The iWorld category in the central position represents the point of origin of vertical and horizontal axes, which correspond to following ranges: a vertical axis with polarities Nature vs. Culture and a horizontal axis, where the range is triangulated by concrete categories (People, Space, Aqua) on the left side and by abstract categories (Challenge, Positive, Negative, Big and Wow!) on the right side.

The vertical range demarcated by the categories Nature on one side and Culture on the other side shows the two opposite directions of how people associate the *world*. The central position of the iWorld category corresponds very well to the basic postulates of PCT, where the individual is the key determinant of how people reflect the *world* and the *self*. Although the iWorld category occurred only rarely, its position marks the center about which personal constructs for the *world* oscillated.

The component plot in a rotated space chart for *self* associations (**Figure 4**) displays the distribution of categories into four groups: (1) categories like Leisure, Fun, Family and Social grouped in the upper part of the data projection are examples representing some aspects of everyday life; (2) the categories Trait, Love, Embodiment, Emotion, Value and Crisis are related with personality structure; (3) alongside, the categories Value, Crisis and Existence in the bottom part of data projection represent the philosophical context of being; and (4) the left part of the chart groups the distinctly personal categories: Individual, Gender and You.

Category Networks

The inverse covariance matrix express the system of statements $\rho_{A,B|X} \sim 0$, where A, B are categories, and X is any kind of knots which separate vertices A and B in the network. $\rho_{A,B|X}$ is coefficient of partial correlation of articulation of A and articulation of B by random subject, controlled for set X. Category network approximates real dependence structures, which are generated from the standardized inverse covariance matrix.

On the basis of the inverse covariance matrix, which analyzes data on the level of individual respondents, two networks of categories were modeled, one for the categories of the *world* associations, second for the associations of *self*. The individual is the numerator of the links. The purpose of the inverse covariance matrix method is to distinguish two types of dependencies, direct and mediated dependencies. An edge connects the nodes directly. If another category is placed between two categories, such a category acts as a transmitter, which mediates a pathway between them and illustrates the mediated dependencies (see **Figures 5** and **6**).

Lines display the direct not-mediated probability dependence. The probability of a common occurrence is independent of other category constellations. As an example, let's take a look at the Linguistic and Trouble categories: if they occur together in the same respondent, the category Negative acts as the mediator here. The only lone-standing category is Culture. This means

TABLE 1 | Description of categories of *world* associations.

Nature	Animals, plants, animate and inanimate nature. Nature's elements. Natural phenomena and processes. Dimension: all on the planet Earth (but the Planet as a whole belongs to the Space category).	Life, air, forest, wind, sand, elements, trees, light, continents, oil, birth, cloud, sky. . .
Aqua	Water. Discrete subcategory of Nature. Everything on water and about water.	Water, sea, ocean, oceans, waterfall, underwater, river. . .
Big	Great. Discrete subcategory of Nature, pointing out the largeness.	Largeness, big, expanse, scope, endless, infinity. . .
Space	Universe. Planet Earth as a whole, its characteristics. Cosmos. Widest possible scope.	Space, planet, Earth, globe, rounded, Sun, everything, sphere, equatorial. . .
People	Human (culture is separated in the next category).	People, population, human, man, mankind, inhabitants. . .
Culture	Culture according to its anthropological definition, which means, in the widest extension. i.e., as the product of human behavior, material and non-material, including values, systems, achievements and products. Including processes, situations, symbolic systems.	Culture, art, politics, money, cities, music, communication, God, society, technology, globalization. . .
Positive	Positive values in general, positive associations.	Beauty, nice, fun, joy, peace, love, happiness, party, harmony. . .
Challenge	The world as the challenge.	Challenge, adventure possibilities, discovering, curiosity, interesting. . .
Negative	Negative values in general, negative associations.	Horrible, pain, evil, cruel, dirty, destruction, madhouse, crisis. . .
Trouble	Serious problems of the world. Discrete subcategory of Negative.	War, catastrophes poverty, violence, corruption, injustice hunger, threatened. . .
Travel	Traveling. Evaluation (positive and negative) does not belong here.	Journey, travel, holiday, maps, vacation, Phileas Fogg, navigation, traveling. . .
Diversity	Variety. If associations reflect differences and diversity, they belong here. Possible overlapping with Culture to be resolved by putting of plurals here, if they refer to diversity (e.g., cultures, languages etc.).	Atlas, nations, other cultures, diverse, complexity, heterogeneity. . .
Social bonds	Social relationships (and their background).	Relationships, family, friends, parents, children, friendship, grandparents. . .
Color	A particular color.	Colors, blue, green, gray, deep blue. . .
Linguistic	A special association within the Czech language (e.g., a rhyme, lexical basis contains the Czech word for world, etc.). Specific Czech facts.	(Not translatable)
Wow!	An expression of amazement, impression. Needs to be balanced with Big, where largeness is the matter (but not astonishment from the largeness).	Miracle, uniqueness, eternity, huge, unbelievable, surprising, magic. . .
Together	Jointly. The expression of mutuality. "We all live in the one world."	Togetherness, we, brotherhood, integrity, wholeness coexistence. . .
iWorld	Me-world, appropriation of the world, I am the world.	My, my life, me, my playground, I, is mine, this around me. . .

that Culture, even if this is the fourth most commonly occurring category, has no probability dependence to be found with another specific category more or less as with others. Such a category does not elicit any category more than any other.

DISCUSSION

The present theory-driven study provides new empirical evidence that contributes significantly to the development of PCT. Aside from this; our results also have implications for other fields,

such as psycholinguistics, associative priming, or philosophical psychology. These implications will be presented within the particular subsections, but at the beginning we will briefly summarize the main findings of the present study. First, the results provide us with the associative ranges for the constructs of the *world* and the *self*. The results showed that some associative dimensions are defined by semantic polarities (e.g., Nature vs. Culture, everyday life vs. philosophical context of being). Second, our analysis showed that some groups of verbal categories associated with the words *world* and *self* are central, while others are peripheral in regard to the central position (see

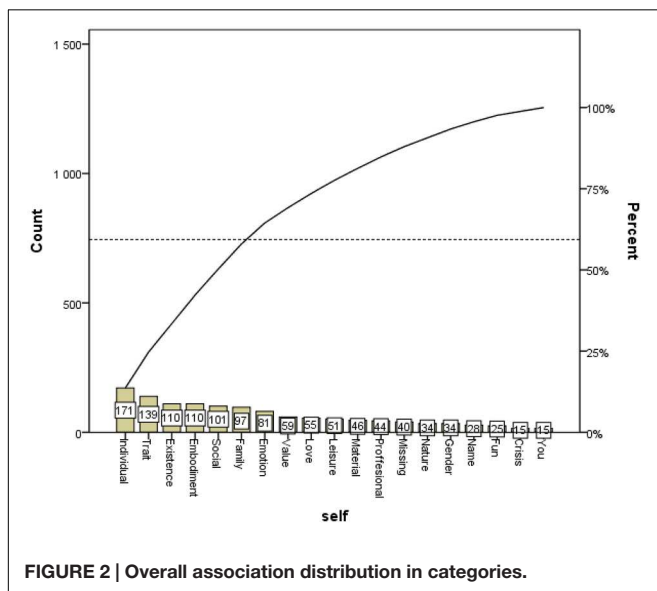
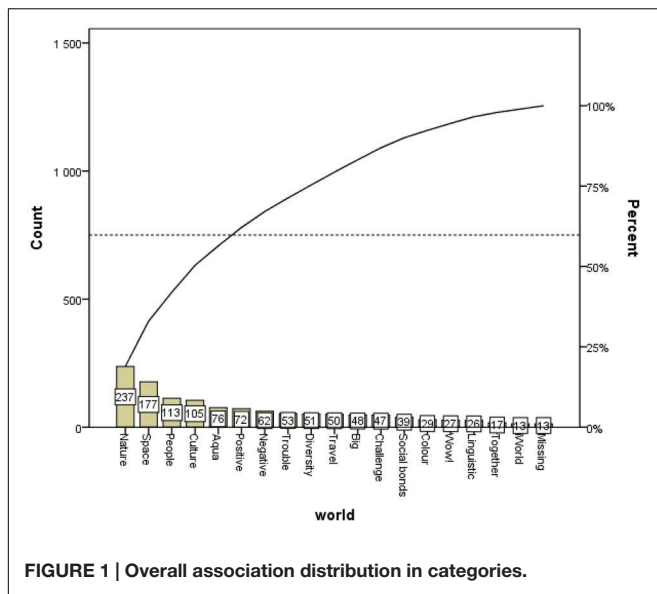
TABLE 2 | Description of categories of self associations.

Trait	Personal attribute, characteristic. Positive and negative self-evaluation, marking of self with a particular characteristic or ability. A particular personality trait or character.	Responsibility, effort, reliability, creativity, intellect, selfish, asshole, slacker. . .
Value	Worth, but because not evaluating, also their opposites. Note: the difference from Trait is often only in that Trait is the accusative form and Value the nominative form. Beware, that Value can be also accusative, if not expressing a Trait directly (example: equal = Value, but not Trait).	Freedom, understanding, health, calm, space, honor, tradition. . .
Individual	I as the initiator (performer), originator and center. Self-highlighting (also in a negative way). This is not a Trait or Value!	Me, ego, human, person, creature, personality, ambition, center, king, individuality. . .
Embodiment	Everything concerning the body. Physical. Evaluation of the body, esthetical and in time. Body activities and abilities. Body reactions. Body language (non-verbal communication).	Smile, beauty, power, old, cry, laugh, brain, skeleton, weight. . .
Existence	Being and consciousness of it, experiencing of it. Questions on being. Trajectory and direction of being. Dimensions of being. My being. Doubts concerning being. Terminality of being.	Life, transience, be, terminality, nothing, who?, paradox, journey, soul, fight. . .
Social	Society. Social bonds of all kinds, including experiencing their absence.	Friendship, relationships, friend, he, she, we, partner, people, together, loneliness, buddy. . .
Family	A distinct subcategory of Social, including types of family bonds, and offspring.	Mother, parents, wife, children, sister, brother, son, daughter, family. . .
Gender	A distinct subcategory of Social. Only the indication of gender, without the possibility of recognizing any social bonds.	Woman, man, girl, gay, boy. . .
You	A distinct subcategory of Social. Separate group of associations, where “I” associates with “you.”	You, you and me, what about you? . .
Name	Own names or names of someone else, incl. Zodiac signs.	(Different names), Brad Pitt, Freud, Ěapek, my name, Asterix, Capricorn. . .
Emotion	And feelings concerning the self. Expressions of like and dislike, Emotionality and its experience. Experienced emotions.	Joy, happiness, fear, sadness, stress, desperation, compassion. . .
Love	A distinct subcategory of Emotion. Everything, where love is explicitly occurring. Including negation of love and its manifestation.	Love, like, flirt, sex, tenderness, I love you, adored, loving. . .
Leisure	Leisure time activities. That which prevailingly associates with free time. That is, I and “something” (a given association) predominantly represent my free time.	Sport, dance, music, football, beer, travel, run, weekend, singing, art, book. . .
Professional	Profession, occupation, job. That which concerns work and preparation for a profession. Professional branches and domains.	Work, student, employment, factory, maid, dentist, sociology, school. . .
Material	Material background, things, tools. Material equipment. Consummated things. Materialism and related point of view on reality, incl. other people.	Home, money, car, household, shopping, train, treasure, welfare, bed, food. . .
Nature	Animals, plants, inanimate nature. Natural elements and resources. Natural phenomena.	Sea, bird, nature, dog, parrot, coast, nests, country, sun, fire, mountains, miaow. . .
Fun	Joy, amusement. “I” is associated with cheerfulness of all kinds.	Humor, fun, giggle, funny, hilarious, amusement, funny. . .
Crisis	Crisis and emergency situations. Mess. Border situations. Consternation or anxieties from the self.	Disaster, illness, chaos, crisis, collision, fall, worry, overpressure. . .

the results of PCA analysis in **Figures 3 and 4**). Third, the analysis of category networks revealed that some categories play the role of a transmitter that mediates a pathway between two other categories in the network (see **Figures 5 and 6**). Interestingly, one category, named “Culture,” was determined to be an independent category in the analysis of category networks, probably without any prevailing dependence with another category. All of these findings help us to better understand the way that our respondents constructed verbal associations with the constructs of the *world* and the *self*, but also the mechanism of personal construction in general.

Implications for Personal Construct Theory

Kelly’s PCT is on one hand considered as essentially important; on the other hand, the lack of empirical evidence supporting his theory has complicated its broader impact. We still consider Kelly’s PCT as a not very empirically verified reservoir of inspiration, which calls for rediscovering. We assume, in line with our results, that such rediscovering, especially when conducted with the use of empirical methods, will broaden PCT and the possibilities of its implications. Even after half a century, some scholars (e.g., Butler, 2006) are pioneering in the gathering of



empirical material which is employed to specify the content of Kelly's theoretical assumptions. Like Butler (2006), who is empirically searching for the content of Kelly's core constructs, the present study aims to provide new insight on the relations between meaningfully categorized words originating from a free association task. The present study used a dynamic way of categorization, which allowed us to work with each arbitrary association, expressed as a word or collocation. Subsequent statistical operations enabled the data to be visualized, separately for associations of *self* and *world*, and opened a space for interpretation of the displayed distributions and relations. Although the present study is not working with a large language corpus, as many recent studies have (e.g., Nelson et al., 2004; Jackson and Bolger, 2014), our data enabled the significant relations between categorized associations to be visualized.

The results of our study led us to reconsider three of the eleven ground stones on which Kelly elaborated the fundamental postulate (see the section PCT) of his PCT. Firstly, we broaden the basic concept of the dichotomous foundations of the construct that Kelly stated and named as a dichotomy corollary; secondly, we reformulate his definition of the choice corollary; and, thirdly, we suggest reducing the number of ground stones of PCT to 10, via complete removal of the range corollary.

Kelly's proposition, which he named the dichotomy corollary, states that people construct their reality on two opposite alternatives. This point was strongly criticized, e.g., by Riemann (1990). Given the results of the PCA analyses conducted within the present study (see Figures 3 and 4), we are able to identify more than only opposite alternatives; moreover, we can reconstruct the entire space where *self* and *world* are emerging as constructs and which are not necessarily based on contrasts (see the section PCA analysis). In this manner, we extend Kelly's original statements.

We assume that Kelly partially anticipates the fact that the foundations of constructs which are not dichotomous also exist. On one hand, he argues that "A construct is the basic contrast between two groups" (Kelly, 2003, p. 10), and on the other, he operates with a "finite number of dichotomous constructs" (ibid.). In other words, the number of dichotomous constructs is limited, which can elicit the question of whether there are also other, non-dichotomous constructs. We guess that here is a hidden space for formulating the fundamentals of constructs other than dichotomous ones, the space which is captured as result of PCA analysis (Figures 3 and 4).

Further, the need for the upgrade of another ground stone, the choice corollary, stems from previous reformulation of the dichotomy corollary, which is considered to be the most important such ground stone. The main contribution of this update is the broadening of the space where alternatives other than the dichotomized can be considered. The distributions of categories, as shown by the results of our PCA analyses, as well as the number of categories – 18 for each of the cue words – show a higher complexity of constructs, where groups of categories were identified, including their central or peripheral positions. Regarding the above-mentioned, we suggest removing the word *dichotomized* from the definition of the choice corollary. The definition will thus read: A person chooses for himself the alternative in a construct through which he anticipates the greater possibility for the elaboration of his system (see Kelly, 2003, p. 11).

Last, but not least, the range corollary, which deals with the issue of a limited number of dichotomous constructs, is groundless in the light of reformulating the previous two fundamentals of the PCT. Kelly stated that the range corollary as "A construct is convenient for the anticipation of a finite range of events only" (ibid), and he subsequently argues that "Not everything that happens in the world can be projected upon all the dichotomies that make up a person's outlook. Indeed I doubt that anyone has ever devised a construct that could cover the entire range of events of which he was aware. There are patches of clouds in every man's sky. This is to say that the geometry of the mind is never a complete system" (ibid). We understand this

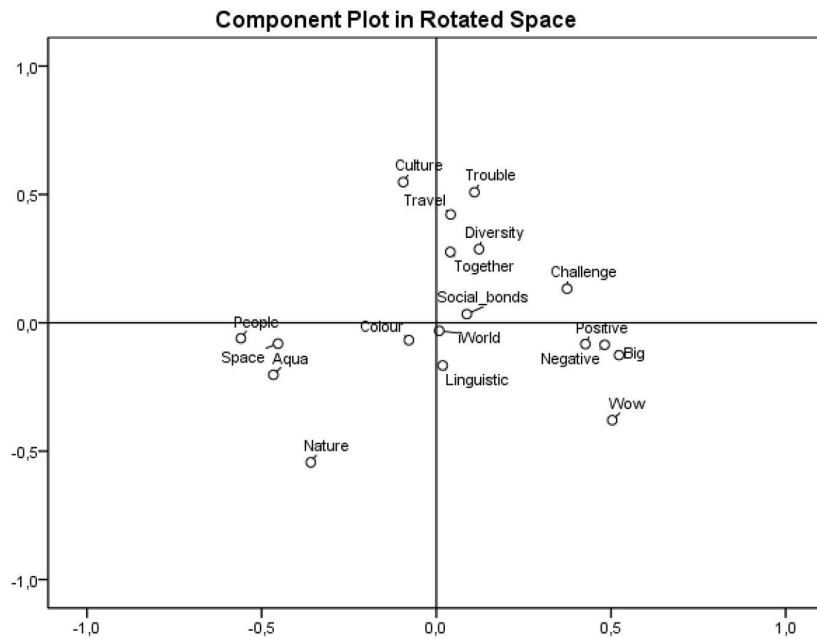


FIGURE 3 | Component plot in rotated space chart for *world*.

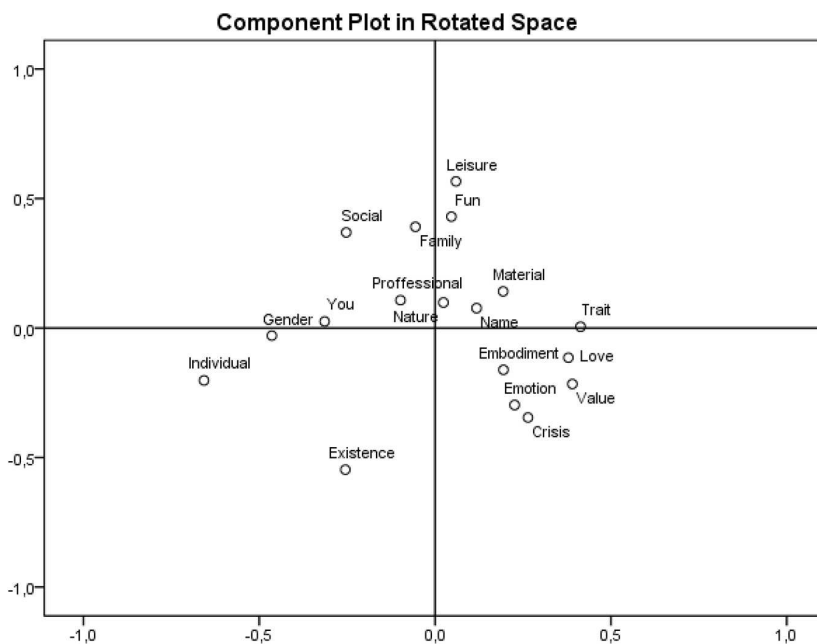
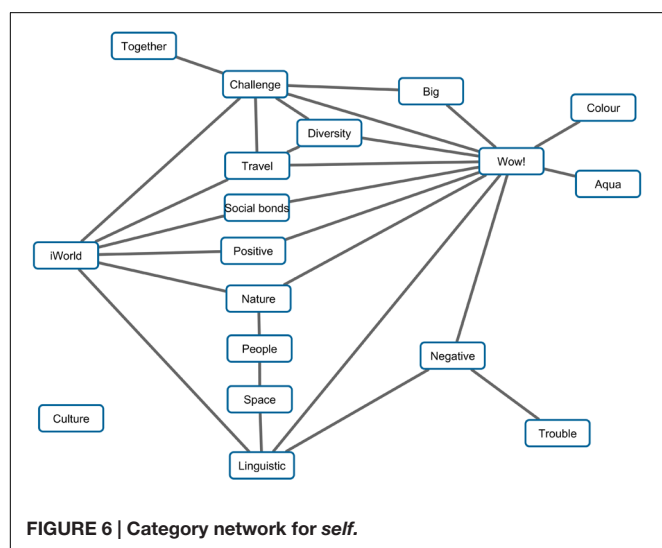
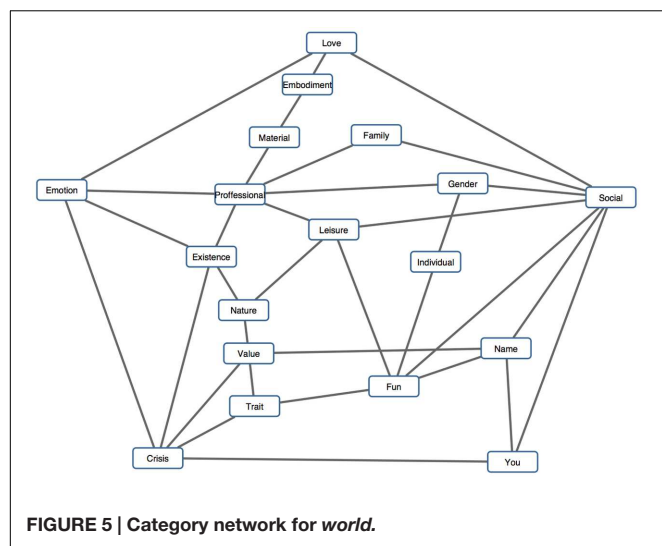


FIGURE 4 | Component plot in rotated space chart for *self*.

quoted paragraph also as a collision with how Kelly delimited the borders of his theory. When we set aside the precondition that constructs are based only on dichotomies, as we suggest above, and what the results of present study illustrate (see results of PCA analysis), we can also consider this Kelly's skepticism as pointless. In this manner, we extend Kelly's original statements; we even reduce his 11 corollaries to 10.

Implications for Psycholinguistics

One of the large topics in contemporary psycholinguistics relates to the term *mental lexicon*. The unflagging interest of scholars about this subject has been elicited by Aitchison (1987), who called it a *human word-store*. To explain the term, the following definition can be employed: "A mental lexicon refers to the words that comprise a language, and its structure is defined here by



“Words are so associatively interconnected with each other they meet the qualifications of a *small world* network, wherein it takes only few associative steps to move from any one word to any other in the lexicon” (Bruza et al., 2009, p. 364). In our study, we revealed the mediated dependencies between categories emerging based on the respondents’ associating, and these dependencies created networks (**Figures 5 and 6**) having characteristics similar to the *small world* networks presented by Bruza et al. (2009). Therefore, our study also extends the theoretical assumption of Bruza et al. (2009) for the case of broader semantic spaces, as coherent categories of meaning are suggested. So, according to our analysis, we may rephrase the above-mentioned definition for the level of categories: categories with coherent semantic meanings emerging from human associating are so interconnected with each other that they meet the qualifications of a network of *small world of associations*, wherein it takes only a few steps to move from any one category to any other category. The links shown between the categories in our analysis showed probability dependences, indeed, as explained in the Category network section. Thus, the present study provides innovative insights into the deep structure of the *mental lexicon*, especially for the instance of associate reasoning and subsequent recall of verbal concepts.

When discussing results concerning the semantic meanings of words, we can also turn our attention to classical authorities in the field. One of first sources of inspiration is the work of Osgood et al. (1957) and his semantic differential. The task of measuring semantic differential requires analyzing the given words in opposite dualities; Kelly (1955) also worked with bipolar dimensions of constructs in his PCT. The results of the present study indicate that some of the associative dimensions that emerged from the associations of our respondents also have the character of opposite dualities. The meanings of categories that are located on the vertical and horizontal peripheral sections of the associative range for the *world* are in semantic opposition (**Figure 3**: Nature vs. Culture, concrete vs. abstract categories). On the other hand, the horizontal dimension in the associative

the associative links that bind this vocabulary together. Such links are acquired through the experience and the vast and semi-random nature of this experience ensures that words within this vocabulary are highly interconnected, both directly and indirectly through other words” (Bruza et al., 2009, p. 363). Within such a delimitation, the present study gathered words from a free association task, where associative recall was primed by the cues *word* and *self*. These association chains were interconnected on the individual, personal level (for examples, see the **Tables 3 and 4**).

Also, subsequent PCA analyses showed relations between categories, where inter-category links were calculated on the basis how individuals articulated their association chains (see the category networks of *word* and *self* in **Figures 5 and 6**). Such data visualization can be considered as an innovative contribution of the present study, because previous analyses in psycholinguistics have been primarily focused on the level of relations between particular words, as in the following example:

TABLE 3 | World: selected association chains (in rows).

Big	Horrible	Cruelty	Injustice	Life
Pain	Ball	Ocean	Life	Death
People	Envy	Nature	Progress	Backstabs
Hazard	Future	Children	Family	Worry
Sea	Courage	Nature	People	Caution
Love	Family	God	Life	Eternity

TABLE 4 | Self: selected association chains (in rows).

Boy	Jerk	Moron	Cool	Brother
Buffoon	Drunkard	Man	Brad Pitt	Superman
Uncertainty	Fear	Confusion	Mom	Good girl
Killer	Ideas	Lie	Mind	Punishment
Speed	Stress	Time	Work	School
Batty	Emotionally labile	Own world	Helper	Life

range for the self does not show such distribution. One may ask: what does this mean? It is necessary to distinguish two levels of analyzed data: first, the level of associations, and second, the superordinate level of categories. On the first level, it is possible to place associations in the form of accusatives and verbs between two polarities, if it is not a polarity itself. In other words, the position of each such association is definable by its position in the range between corresponding polarities, defined by Kelly as the *extreme points of constructs*. In linguistic categorization, nouns, adjectives and various types of verbs are the basis for sorting words on the scale concreteness–abstractness (e.g., Carnaghi et al., 2008). Because the present study acquired associations of all word classes and because of the aim to categorize all of gathered data by the prevailing meaning of the words or short phrases used, another sorting on the first level wasn't employed, even though there is a wide scale of methods by which these data could be analyzed (e.g., Nelson et al., 2004). However, new insight on the concreteness–abstractness dichotomy was provided, when PCA analysis of the categories from the term *self* was conducted (see **Figures 3** and **4**).

Our data are not sufficient to conclude that people construct the *world* and the *self* in the manner of opposite dualities, but analysis of the acquired data is able to show that respondents do employ non-opposite mechanisms of semantic construction when associating the given cue words. Nonetheless, it is too early to make such a general conclusion, and the fact that this interpretation is supported only by the results of one empirical study should be taken into account. On the second level of the analyzed data, some very significant opposite dualities are illustrated among the categories by the PCA method. Of course, not all categories in the PCA projection were located in the sense of opposing dualities, which supports another possible way of interpreting of results, as employed by the distribution of the *self* categories (see **Figures 3** and **4**). When analyzing the data on the level of categories (for example Nature and Culture, see **Figure 3**), their distribution in the given example may also be understood as *extreme points of constructs*, if using Kelly's words.

Further, it is also necessary to stress that the experimental design used in the present study involves the effect of priming. The respondents were provided with the cue words, *world* and *self*, and these cue words influenced their follow-up retrieval of associations from memory. Stacy et al. (2006) pointed out: "Thus, word association seems to measure some sort of relationship that has relevance well beyond the word association task itself" (Stacy et al., 2006, p. 77). Both of the words, *world* and *self*, served as prime words leading to a related target – a chain of word associations in the present study (for examples, see the **Tables 3** and **4**). It is very important to keep in mind this priming effect when exploring the results of the analysis of category networks. This analysis revealed that some categories played the role of a transmitter mediating a pathway between two other categories in the network. Just the presence of transmitters or mediators illustrates very well the sequential nature of our data. Analyzing the problem of sequentiality in chains of words in associative tasks would be an interesting challenge for future studies, but it lies beyond

the main scope of the present study, which was focused mainly on the investigation of associative ranges for the constructs of the *world* and the *self* within the theoretical framework of PCT.

Implications for Philosophies Regarding Self and World

This paper is basically an empirical study, not a philosophical one. Despite this fact, however, some interesting implications for philosophies regarding *self* and *world* can be derived from the acquired data. Generally, we understand the implications introduced below to be a link between the realms of theoretically driven empirical research study on one hand and the field of philosophical thinking on the other. These implications are thought mostly to be preliminary incentives for future developments in the field rather than final statements.

At the beginning, it is important to realize the nature of our experimentally acquired data. Simply put, the respondents recalled their first five associations with the words *world* and *self*. Therefore, we may afford to discuss these associative findings within the context of the phenomenology of Husserl. According to Husserls (1970, 1977), our mental processes are essential for the construction of the objective world. Through and by our mental processes we create our own ideas about our existence *in* the world and belongingness *to* the world. From this point, the associations that were gathered from our respondents showed us the underlying structures that constitute the inner meanings of the world's belongingness. For example, when we look at **Figure 3**, there are many possible concepts of how people may understand their being in the world. The free association task enabled us to approach the inner meanings of the *world* in our respondents. But the question arises: what is the source of these meanings?

Husserls (1970) speaks about *a priori* structures that may be responsible for our belongingness with the world. Categories such as Space, Aqua, Positive, or Negative (see **Figure 3**) are examples of concepts that could even be regarded as some kind of *a priori* structures pre-established in our brains. However, all associations are also primarily subjective. The role of cognition should not be omitted here. Cognition definitely varies between subjects, and free associations are suggested to emerge into the consciousness of a person because of that person's performance of cognition (*Erkenntnisleistung*) (Husserls, 1977). Cognition is generic within the free associative task, because it generates various associative meanings when faced with stimulus priming. As seen in **Figure 3**, some kinds of categories seem to be less general, e.g., Social bonds, Travel or Linguistic. A person's performance of cognition (*Erkenntnisleistung*) (Husserls, 1977) plays an important role in the free associative task, and we argue that the gathered associations may be the expression of *a priori* structures but also influenced significantly by a person's cognition. The concepts that respondents have created in their minds have been constituted (and are still being constituted) in the subject's constitutional history. And because the associations

presented in this study are related to the *world* and the *self*, we may even speculate that these associative concepts can mirror even the core, fundamental concepts of a person's being in the world.

Perhaps it is too early to make such statements, and we should be careful to come to such a conclusion. Therefore, we instead adopt a less daring position here and satisfy our inquiry with the notion that free association may provide us with signs of the deep mental processing of reality. But there is another interesting implication relating to the mutual inter-relations of associations in our results. As shown by the analysis of category networks (see **Figures 5 and 6**), we may suggest that the concepts are not disconnected from each other in the respondents' memories. The analysis of category networks revealed a high number of mediated dependencies that are depicted in **Figures 5 and 6**. This informs us about the deep structuring of the inner perception of the world in people's minds. According to Husserl's (1970), all objects are experienced in consciousness, and therefore, all meanings are always constituted in the constitutional history of a subject. Our results showed (see **Figure 5**) that concepts relating to the world are mostly inter-dependent, and that one concept often acts as a transmitter for another concept or concepts. At this point, it would be speculative to consider mediated dependencies between concepts to be related with Husserl's (1967) basic eidetic laws of passive genesis, but such question may be considered.

Our experimental study worked with verbal stimuli, and respondents also recalled verbal associations to the words *world* and *self*. The choice of this research design is not surprising when working in the field of philosophical psychology. Heidegger (1985) pointed out that our practical encounter with the world around us is encapsulated in our language. In other words, language itself is suggested to mirror our belongingness with the world. Our results (**Figure 4**) showed that self-related associations are grouped into spheres that are related either to some aspects of everyday life (e.g., categories such as Leisure, Fun, Family and Social) or to the philosophical context of being (e.g., categories such as Value, Crisis and Existence, in the bottom part of **Figure 4**). This patterning indicates that Heidegger's (1985) practical engagement with the world is not restricted to the realm of everyday activities, but also covers engagement with the world that can be considered to be the philosophical context of being of human existence. This existential dimension of being in the world is present in the associations of our respondents, which indicate that the existential aspects of life are also an important part of verbalized Expression (*Aussage*) in Heidegger's (1985) sense.

Further, it is also possible to posit the broader philosophical question of how many categories are optimal to categorize the world? This question cannot be answered easily, but we may consider some possibilities here. Lowe (2006) posed a similar question in the framework of his four-category ontology, and according to this *squared* way of thinking he concluded, not surprisingly, with four components. In contrast, our category system, which emerged based on empirical evidence, included 18 categories for associations of the word *world* and the same number of categories for associations of the word *self* (see **Tables 1 and 2**). Each category represents a group of verbal

associations with similar semantic meanings (see **Tables 1 and 2**), and the frequencies of occurrences are different in different categories (**Figures 1 and 2**). When dealing with continuous attempts to categorize the world from the ancient times up to the present (for a review, see Hackett, 2014), some scholars have even posited the opinion that no categories exist at all. This is the case of the work *The No-Category Ontology* (Bueno et al., 2015), where instead of categories the authors introduce concepts which "can always be revised, refined and recast" (Bueno et al., 2015, p. 233). This fascinating discussion is currently very topical, and we believe that the empirically based results of the present study contribute significantly to it.

Finally, we would like to briefly outline the implication of our study in respect to the position of the concept of *self* in philosophy. Although most psychologists do not have any problems with the concept of *self*, some philosophers have questioned the existence of the *self* in terms of an essential, subjective nucleus of a person. For example, Hofstadter and Dennett (1981) rejected the existence of the *self* and regarded the *self* as a kind of illusion. This approach is very inspiring. Although the words "self" or "I" are included in most of human languages, we may adopt the illusionary nature of the underlying concept of *self* for the interpretation of our results. If we consider the concept of *self* to be an illusionary one, the *self*-related associations that have been gathered from our respondents become much more interesting in this light. If *self* is an illusion, then associations to the word *self* may be helpful in revealing the relationship of this illusionary concept to other related concepts. When looking at **Figure 2**, the word *self* is most frequently associated with concepts falling into the categories "Individual," "Trait," "Existence," "Embodiment," "Social," and "Family." This indicates that people usually understand themselves as individuals with some kind of traits. The *self* has not been consciously regarded as an illusion, because respondents frequently recalled words relating to the realm of real existence. They associated the *self* with a really existing entity embodied in a particular material form, i.e., a human body. Many associations were also targeted at the social domain of human existence, where *self* was associated with words falling into the categories "Social" and "Family." We do not aim to use our results for a discussion about the existence or non-existence of the concept of *self* in philosophy. Indeed, we admit that the *self* could be an illusionary concept, and our results (see **Figure 2**) indicated that although illusionary, the *self* was mostly associated with the conceptualization of a socially living individual that has both a material body and psychological traits. It is necessary to say that this embedding of the *self* was found in our population sample, and cross-cultural differences can be expected when conducting the same experiment in another cultural environment.

CONCLUSION

The present study reconsiders the PCT in light of the analyses of data acquired in a free association task. The results of this study develop the theoretical concept of Kelly (1955), which has often been criticized for the lack of empirical evidence supporting

the PCT. The acquired data were categorized by the dynamic classification method, creating different meaning categories for the cue words *world* and *self*. The categories are based on the acquired data, and the mentioned dynamic disposition of the method simply reflects any data content.

Subsequent analyses illustrate the sphere which was partially outlined by the PCT as dichotomies, but which to our knowledge nevertheless were not shown via the PCA or inverse covariance matrix methods. Categories are distributed in a two-dimensional range, e.g., in central and peripheral positions, semantic polarities and groups. Further, analysis of category networks detect which categories have prevailing dependencies toward other categories when associating *world* and *self*, and which categories are transmitters to other ones. The inverse covariance matrix method also provided an innovative upgrade of the association-to-association links (Nelson et al., 2004) on the level of category-to-category links.

According to the results of the conducted analyses, the present study suggests implications in three areas: (1) in the PCT, we reformulated two of the 11 ground stones which Kelly named as corollaries. After this reformulation PCT is capable to reflect that the constructing of *self* and *world* is not based only on contrasts, as also our analyses depict. Fundaments of constructs different than dichotomous fundaments are also illustrated here. Furthermore, it is suggested that one corollary should be completely removed; (2) in the area of psycholinguistics, the extension of the concept of *small world* network (Bruza et al., 2009) is suggested. Categories emerging from how people associate are linked based on their probability dependences in the present study, which enables an innovative insight into a current topic of psycholinguistics topic – the mental lexicon; (3) the results of the present study are framed in the context of Husserl's (1970, 1977) phenomenology, where human mental processes are considered as essential for the construction of the objective world. People create ideas about their existence *in* the world and belongingness *to* the world, which the free association task with the cue words *self* and *world* appropriately stimulates. In this way, the present study provides a particularized insight into the

underlying structures, which constitute the inner meaning of world's belongingness. Moreover, analysis of the category network shows that associations are mutually inter-related and that the concepts of *self* and *world* are internally connected via direct and mediated dependences, which reflects the structuring of perception and understanding of *self* and *world* in people's minds.

It is also appropriate to consider the limitations of the present study regarding culture: the sample of Czech people, mostly university students, and also the fact that all of the acquired associations stem from the Czech language repertoire. On the other hand, the Czech Republic currently represents a unique mono-cultural non-religious laboratory in the middle of Western world (see e.g., Kuška et al., 2015), so the results of the free association task could be used as control-group data in comparison with another multicultural sample. The method of data collection, their categorization and subsequent analyses is universal.

AUTHOR CONTRIBUTIONS

MK is the principal author of this article and contributed to all of its sections. RT, the second author, is the author of the main parts of the Method and Discussion sections as well a co-author of the Results and the Conclusion. AAK designed and conducted the statistical operations and is the main author of the Data analysis section of the article, and partially of the Results section. JR is the author of the Introduction and Results sections. All four co-authors critically revised the entire text and worked together on the final version of the article.

FUNDING

This work was supported by the Czech Science Foundation (GAČR), project no. 14-35577S. Statistical data analyses were supported by the Excellence Project P402/12/G097 DYME Dynamic Models in Economics of GAČR.

REFERENCES

- Aitchison, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Blackwell.
- Aylwin, S. (1977). The structure of visual and kinaesthetic imagery: a free association study. *Br. J. Psychol.* 68, 353–360. doi: 10.1111/j.2044-8295.1977.tb01600.x
- Bannister, D., and Mair, J. M. M. (1968). *The Evaluation of Personal Constructs*. London: Academic Press.
- Bruza, P. D., Kitto, K., Nelson, D., and McEvoy, C. L. (2009). Is there something quantum-like about the human mental lexicon? *J. Math. Psychol.* 53, 362–377. doi: 10.1016/j.jmp.2009.04.004
- Bueno, O., Busch, J., and Shalkowski, S. A. (2015). The no-category ontology. *Monist* 98, 233–245. doi: 10.1093/monist/onv014
- Butler, R. J. (2006). Investigating the content of core constructs. *Personal Constr. Theory Pract.* 3, 27–33.
- Carnaghi, A., Maass, A., Gresta, S., Bianchi, M., Cadinu, M., and Arcuri, L. (2008). Nomina sunt omina: on the inductive potential of nouns and adjectives in person perception. *J. Pers. Soc. Psychol.* 94, 839–859. doi: 10.1037/0022-3514.94.5.839
- Cridland, E. K., Caputi, P., Jones, S. C., and Magee, C. A. (2014). Understanding high-functioning autism during adolescence: a personal construct theory approach. *J. Intellect. Dev. Disabil.* 39, 108–118. doi: 10.3109/13668250.2013.870331
- Hofstadter, D. R., and Dennett, C. (1981). *The Mind's I: Fantasies and Reflections on Self and Soul*. New York, NY: Bantam Books.
- Eysenck, H. J. (1952). The effects of psychotherapy: an evaluation. *J. Consult. Psychol.* 16, 319–324. doi: 10.1037/h0063633
- Eysenck, H. J. (1968). Personal construct theory. The evaluation of personal constructs. *Nature* 220:99. doi: 10.1038/220099a0
- Fransella, F. (2005). *The Essential Practitioners' Handbook of Personal Construct Psychology*. Chichester: John Wiley & Sons.
- Geeraerts, D. (2010). *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Hackett, P. M. W. (2014). *Facet theory and the Mapping Sentence: evolving Philosophy, Use and Application*. New York, NY: Palgrave Macmillan. doi: 10.1057/9781137345929

- Heidegger, M. (1985). *History of the Concept of Time: Prolegomena*. Bloomington, IN: Indiana University Press.
- Home, R., Bauer, N., and Hunziker, M. (2010). Cultural and biological determinants in the evaluation of urban green spaces. *Environ. Behav.* 42, 494–523. doi: 10.1177/0013916509338147
- Horley, J., and McPhail, I. (2009). “What’s in a name? Interpreting terrorism from the perspective of personal construct theory,” in *Engaging Terror: A Critical and Interdisciplinary Approach*, eds M. Vardalos, G. K. Letts, H. M. Teixeira, A. Karzai, and J. Haig (Boca Raton, FL: BrownWalker), 118–127.
- Husserl, E. (1967). *Cartesian Meditations: An Introduction to Phenomenology*. The Hague: Martinus Nijhoff.
- Husserl, E. (1970). *Logical Investigations*. New York, NY: Humanities Press.
- Husserl, E. (1977). *Phenomenological Psychology: Lectures, Summer Semester 1925*. The Hague: Martinus Nijhoff.
- Jackson, A. F., and Bolger, D. J. (2014). Using a high-dimensional graph of semantic space to model relationships among words. *Front. Psychol.* 5:385. doi: 10.3389/fpsyg.2014.00385
- Kelly, G. A. (1955). *The Psychology of Personal Constructs*. New York, NY: Norton.
- Kelly, G. A. (1991). *The Psychology of Personal Constructs: A Theory of Personality* Vol. 1. London: Routledge.
- Kelly, G. A. (2003). “A brief introduction to personal construct theory,” in *International Handbook of Personal Construct Psychology*, ed. F. Fransella (Chichester: John Wiley & Sons).
- Kline, P. (2013). *Handbook of Psychological Testing*. London: Routledge.
- Kuška, M., Trnka, R., Tavel, P., Constantino, M. J., Angus, L., and Moertl, K. (2015). The role of cultural beliefs and expectations in the treatment process: clients’ reflections following individual psychotherapy. *Sex. Relation. Ther.* 1, 1–12. doi: 10.1080/14681994.2014.1001354
- Lowe, E. J. (2006). *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press.
- Nelson, D. L., McEnvoy, C. L., and Schreiber, T. A. (2004). University of South Florida free association, rhyme, and word fragment norms. *Behav. Res. Methods Instrum. Comput.* 36, 402–407. doi: 10.3758/BF03195588
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
- Riemann, R. (1990). The bipolarity of personal constructs. *Int. J. Pers. Constr. Psychol.* 3, 149–165. doi: 10.1080/10720539008412806
- Stacy, A. W., Ames, S. L., and Grenard, J. L. (2006). “Word association tests of associative memory and implicit processes: theoretical and assessment issues,” in *Handbook of Implicit Cognition and Addiction*, eds R. W. Wiers and A. W. Stacy (Thousand Oaks, CA: Sage), 75–90.
- Trnka, R., Lačev, A., Balcar, K., Kuška, M., and Tavel, P. (2016). Modeling semantic emotion space using a 3D hypercube-projection: an innovative analytical approach for the psychology of emotions. *Front. Psychol.* 7:522. doi: 10.3389/fpsyg.2016.00522
- Überla, K. (1968). *Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler*. Heidelberg: Springer.
- Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. New York, NY: Wiley.
- Winter, D. A. (2012). Still radical after all these years: George Kelly’s the psychology of personal constructs. *Clin. Child Psychol. Psychiatry* 18, 276–283. doi: 10.1177/1359104512454264
- Woodworth, R. S. (1921). *Psychology: A Study of Mental Life*. New York, NY: Henry Holt.

Conflict of Interest Statement: 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.

Copyright © 2016 Kuška, Trnka, Kuběna and Růžicka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Going Beyond the Data as the Patching (Sheaving) of Local Knowledge

Steven Phillips*

Mathematical Neuroinformatics Group, Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

OPEN ACCESS

Edited by:

Lieven Decock,
VU University Amsterdam,
Netherlands

Reviewed by:

Tomer Fekete,
Ben-Gurion University of the Negev,
Israel

Yasmina Jraissati,
American University of Beirut,
Lebanon

*Correspondence:

Steven Phillips
steve@ni.aist.go.jp

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 June 2018

Accepted: 19 September 2018

Published: 09 October 2018

Citation:

Phillips S (2018) Going Beyond the
Data as the Patching (Sheaving) of
Local Knowledge.
Front. Psychol. 9:1926.
doi: 10.3389/fpsyg.2018.01926

Consistently predicting outcomes in novel situations is colloquially called “going beyond the data,” or “generalization.” Going beyond the data features in spatial and non-spatial cognition, raising the question of whether such features have a common basis—a kind of systematicity of generalization. Here, we conceptualize this ability as the patching of local knowledge to obtain non-local (global) information. Tracking the passage from local to global properties is the purview of sheaf theory, a branch of mathematics at the nexus of algebra and geometry/topology. Two cognitive domains are examined: (1) learning cue-target patterns that conform to an underlying algebraic rule, and (2) visual attention requiring the integration of space-based feature maps. In both cases, going beyond the data is obtained from a (universal) sheaf theory construction called “sheaving,” i.e., the “patching” of local data attached to a topological space to obtain a representation considered as a globally coherent cognitive map. These results are discussed in the context of a previous (category theory) explanation for systematicity, vis-a-vis, categorical universal constructions, along with other cognitive domains where going beyond the data is apparent. Analogous to higher-order function (i.e., a function that takes/returns a function), going beyond the data as a higher-order systematicity property is explained by sheaving, a higher-order (categorical) universal construction.

Keywords: learning, generalization, sheaf theory, sheaf, sheaving, category theory, universal

1. INTRODUCTION

A ubiquitous cognitive ability is the capacity to “go beyond the data.” That is, to put it broadly, an ability to successfully respond to stimuli not previously encountered. Such a characterization encompasses a wide variety of situations from perception-based classification to logic-like reasoning. For example, given feedback on the edibility of a particular kind of fruit, one knows when the fruit can be eaten next time it comes into season. Or, having been (repeatedly) rewarded for choosing stimulus *A* over stimulus *B* and *B* over stimulus *C*, one correctly predicts that choosing *A* over *C* will also elicit a reward. In general, a capacity to go beyond the data is referred to as *generalization*. And, this ability is typically expressed as correct responses to novel inputs given some knowledge about other input-output (cue-target) examples.

This broad view of generalization affords an instructive comparison/contrast of two distinctive views of cognition, to wit, *classical* (symbolic) and *connectionist* (subsymbolic/vectorial). The relative merits of these two views (Fodor and Pylyshyn, 1988) have been extensively debated in

the literature (see Calvo and Symons, 2014, for a cross-section of arguments). Our interest, here, is with several key aspects of the debate that motivate and help illustrate a different conception of generalization to follow.

The classical view is that our ability to reason about the world is founded upon a compositional syntax and semantics: the world is interpreted through a *language of thought* (Fodor, 1975). A language of thought is a system of representations—complex entities are modeled by corresponding compositional representations so that the semantic relationships between the constituent entities are reflected in the syntactic relationships between the corresponding constituent representations—and processes that are compatible with the way such compositional representations are constructed. So, for instance, on seeing that John is standing to the left of Mary, there is a symbol representing John juxtaposed with a symbol representing Mary in a way that captures the relative spatial locations of John and Mary. What matters to the classical theory is not the particular syntactic relationship, but that the relationship employed is used consistently in all such situations. In this way, a classical cognitive system with the capacity to juxtapose all such relevant combinations of symbols is supposed to explain the *productivity* and *systematicity* properties of language (Chomsky, 1980) and thought (Fodor and Pylyshyn, 1988), more generally.

Productivity and systematicity can be seen as forms of generalization in the sense just introduced. Productivity, as the term suggests, is characterized as having a cognitive capacity that is more than the sum of its parts. For instance, suppose in the course of understanding the meaning of “to the left of” that upon being told, “John is standing to the left of Mary,” “Mary is standing to the left of Tom,” and “John is standing to the left of Tom” when indeed one sees that John is to the left of Mary, and so on, that symbols are recruited to represent John, Mary and Tom. Suppose, further, that a (product) rule is exercised to combine that set of symbols into a set of symbol pairs: e.g., (John, Mary) in corresponding (left, right) order, together with a process for accessing the first symbol in each pair, thereby affording the inference that John is the person on the left when applied to the pair (John, Mary). This system exhibits productivity, therefore generalization in the aforementioned sense, because a basic capacity to represent three pairs of people produces (generalizes to) a capacity to represent all six possible pairs of people without further instruction. A similar consideration applies to systematicity: where having the capacity for one such instance implies having the capacity for another (structurally-related) instance, via application of the same combinatorial process (Fodor and Pylyshyn, 1988; Aizawa, 2003).

A connectionist view, which eschews symbolic representations and processes, is that our ability to reason about the world is founded upon vector (coordinate) based representations and processes, realized as networks of neuron-inspired computational units (Rumelhart et al., 1986). A connectionist model employs vectorial representations—complex entities are modeled by corresponding vectors so that the semantic relationships between the constituent entities are reflected in the spatial (geometrical) relationships between the corresponding constituent representations—and functions that

are compatible with the way such vectorial representations are constructed. In the linear case, where the computational units involve only linear functions, connectionist models can provide analogous accounts of productivity and systematicity via linear algebra (Smolensky, 1990). In the non-linear case, where units involve non-linear functions, productivity and systematicity, and other forms of generalization, obtain from judicious choices of learning methods and non-linear functions (see e.g., Hadley, 1994; Frank et al., 2009, among many others).

Although classical and connectionist approaches can demonstrate various generalization properties, they both fall short of an important theoretical challenge. That challenge is to explain *why*, not just how properties such as systematicity derive from the core principles of the theory. This challenge was the one originally raised against connectionist theories (Fodor and Pylyshyn, 1988), and later shown to be problematic for classical theories too (Aizawa, 2003). The essence of the problem is that the core principles admit systems that do *and* systems that do not exhibit systematicity. In both cases, the core theoretical claims do not derive the systematicity properties without tailoring auxiliary assumptions to fit the data whenever such properties are evident. Such assumptions are characteristically *ad hoc* in being unconnected to the core principles of the theory, motivated solely to fit the data, and cannot be confirmed independently of confirming the theory—accordingly, classical and connectionist theories fail to *fully* explain such properties (Aizawa, 2003).

To address this challenge, a *category theory* (Eilenberg and Mac Lane, 1945; Mac Lane, 1998) approach was proposed whereby systematicity properties derive from *universal (categorical) constructions* (Phillips and Wilson, 2010). For example, of the many possible ways of combining symbols or vectors to represent pairs there is only “one” way (see remark 8 in **Appendix** of the Supplementary Material) to combine them so that the constituents are uniquely accessible in every possible case, called the *categorical product*. Various scenarios for systematicity were explained in terms of appropriate universal constructions (see Phillips and Wilson, 2016b, for an overview). A summary of the systematicity challenge, which motivates the categorical theory approach is given in the last section of the **Appendix** (Supplementary Material).

An explanation for systematicity, however, raises to a wider question, Why do people fail to exhibit systematicity in some situations? In particular, failure to apply certain rules of inference (*modus ponens* and *modus tollens*) in the relevant situations at least calls into question the classical account of systematicity (van Gelder and Niklasson, 1994). Other forms of fallacious reasoning, such as the *conjunction/disjunction fallacy* (Tversky and Kahneman, 1983) and *pseudo-transitive inference* (Goodwin and Johnson-Laird, 2008), raise a similar challenge. A general framework within which such questions and challenges may be addressed is called *dual-process* (see Evans, 2003, for a review).

Dual-process accounts of cognition assume two modes of thinking, generically labeled *Type 1* and *Type 2*, which are typically characterized as fast, reflexive, associative and relatively effortless—Type 1—vs. slow, reflective, rule-based and relatively effortful—Type 2 (Kahneman, 2011; Evans and Stanovich, 2013). The basic idea is that the two systems trade off complementary

properties so that, for example, under time pressure a faster Type 1 process may supersede a slower Type 2 process yielding an incorrect response (Kahneman, 2011; Evans and Stanovich, 2013)—the two types of processes trade the benefit of speed for the cost of accuracy.

Along similar lines, a trade-off was hypothesized in regard to systematicity: for the categorical account, lack of systematicity is due to the relative cost/benefit of constructing the appropriate universal morphism (Phillips et al., 2016). This hypothesis was tested in a stimulus-response learning experiment, where the maps to be learned were *products* of cue-target maps. The supposed trade-off involved learning a single (associative) route of n^2 mappings vs. a pair of routes (via a product-rule) of $2n$ mappings—increased memory vs. decreased attention. Two groups of participants were administered the task. The *ascend* group were trained and tested on four different cue-target maps in ascending order of map size (i.e., from three-by-three to six-by-six possible cue-target associations). This group showed generalization (correct responses) to novel stimuli in the testing set only when the number of cue-target pairs to be learned was large, indicating that they did not construct the universal morphism (product map) for small maps, even though there were sufficient training examples to induce the construction. The *descend* group were trained and tested in descending order of map size. This group showed generalization to the testing set at all sizes, indicating systematic induction of the product map. Together, these results support a cost/benefit explanation (Phillips et al., 2016).

The cost/benefit explanation, as it pertains to the experiment, raises two closely related questions: (1) what determines the choice of (associative vs. [product] rule-based) learning route, and (2) in the case of the rule-based learning route, how/why are universal morphisms systematically constructed? This paper is primarily concerned with the second question: under the assumption that participants are driven toward the rule-based route, how/why are universal morphisms constructed? We return to the broader question of how/why participants are driven to this learning route, i.e., the interaction between cost/benefit and the construction of universal morphisms in the Discussion. To the second question, then, the consistent (systematic) transition from no systematicity (no universal construction) to systematicity (universal construction) itself suggests another form of universal construction. These considerations, which constitute the starting point for the current work, lead naturally to another (closely related) branch of mathematics, called *sheaf theory* (Hartshorne, 1977; Mac Lane and Moerdijk, 1992), applied here as a basis for generalization.

The import of sheaf theory to cognitive science may seem obscure. So, a preview of the sheaf theory approach is provided in the remainder of this introduction before delving deeper into the conceptual details and cognitive applications (subsequent main text), and supporting formal theory (**Appendix** in Supplementary Material).

1.1. Preview: Generalization as Patching (Sheaving)

A capacity to generalize beyond the given instances connotes a property that is (re)constructed from local information.

Conceptually, at least, this situation is akin to tracking the passage from a local to a global property, which is the purview of sheaf theory. This way of looking at generalization renders the essential ingredients (axioms) of sheaf theory as a formalization of some classical and connectionist concepts already introduced. In this light, the path from sheaf theory to cognition is less abstruse. Indeed, sheaf theory is where algebra meets geometry/topology. If one regards classical and connectionist approaches as complementary, which some researchers do (e.g., Holyoak and Hummel, 2000; Clark et al., 2008)—a language of thought (Fodor and Pylyshyn, 1988) on one hand and a geometry of thought (Gardenfors, 2000) on the other, then sheaf theory alludes to a natural integration of the two.

There are three fundamental aspects to sheaf theory that we will interpret in terms of cognitive representations and processes: (1) *presheaf* (*sheaf*), a basic element of sheaf theory, which we will regard as a (coherent) cognitive map or representation, (2) *sheaving*, the (universal) process of constructing a sheaf from a presheaf, which we will interpret as a form of systematic generalization, and (3) *sheaf morphism*, regarded here as a kind of inference, i.e., a cognitive process acting on a cognitive representation.

A sheaf is like a work of art, and sheaf operations are like the artistic process, in the following sense. To create a portrait, an artist applies paint to canvas. The canvas is a topological space and the paint is the data attached to that space. As a work in progress, there are unpainted regions on the canvas, or sections of the portrait that don't quite match. In unfinished form, the portrait is a presheaf. Paint is added to the vacant regions, or laid over existing sections to obtain the finished form. This patching process is likened to sheaving, and the finished form to a sheaf. The finished portrait may be further altered, e.g., by changing tone to affect mood, thus creating a new portrait, and this process is likened to a sheaf morphism.

This artistic rendering of sheaves has analogs in classical and connectionist theory. The canvas is a representational space in which symbolic, or vectorial representations are constructed, combined (patched), or transformed. As we shall see, sheaf theory provides a formal basis for such processes and, in particular, generalization as the patching of local knowledge to obtain non-local (global) information. We present the basic sheaf theory and the sheaving construction considered as a “universal” basis for generalization (section 2). Then we examine sheaving in two cognitive domains (section 3): cue-target learning, involving the product of two cue-target maps, and visual search involving the integration of two visuospatial maps. In terms of sheaves, the first domain is a special case of the second domain. These results are discussed in the context of a previous (category theory) explanation for systematicity, vis-a-vis, categorical universal constructions, along with other cognitive domains where going beyond the data is apparent (section 4). Supporting technical material is provided in the **Appendix** (Supplementary Material).

2. SHEAVES AND SHEAVING

As previewed in the Introduction, sheaf theory concerns the passage from local to global properties, which we interpret as generalization in the context of cognition. This section

TABLE 1 | Corresponding set/relational database and category/sheaf theory concepts.

Set theory	Relational database theory
Element, set	Column name, header
(assignment) function	(data) table
(higher-order) function	(table) transformation
optimal function	natural join, renormalization
Category theory	Sheaf theory
Object/morphism, category	Open set/inclusion, topology
(contravariant) functor	presheaf/sheaf
natural transformation	presheaf/sheaf morphism
universal morphism	pullback, sheaving

provides a conceptualization of the formal details. Although the presentation in this section is primarily intuitive, some notation is included to facilitate links to the formal theory. Sheaf theory is initially given in terms of sets and functions. However, a category theory view is particularly relevant here, because of our interest in universal constructions as an explanatory basis for systematicity (generalization). In short, there are two levels of universality: one at the level of sheaf, which is defined via a universal construction, and the other at the level of collection of sheaves that pertains to sheaving, which is another kind of universal construction. Accordingly, sheaving pertains to a kind of second-order systematicity (see Chomsky, 1980; Aizawa, 2003; Phillips and Wilson, 2016a), alluded to earlier. Moreover, a category theory approach affords wider applicability to domains involving more structure than just sets. A guide to formal concepts for the various theoretical views and their relationships is given in **Table 1**.

Conceptually, one can think of sheaving as a process of obtaining a coherent “map” or representation of a complex situation that is a sheaf. Hence, the state of affairs *before* having a sheaf is called a *presheaf*. Sheaving is a (universal) way of going from a presheaf to a sheaf. An immediately intuitive example is navigating a city using a street directory. Each page of the directory contains a map of a local area. To visit a distant part of town, pages mapping contiguous areas must be “glued” together along common landmarks to yield a map that includes both the current location and the destination. Gluing all such pages is akin to constructing a sheaf, and the construction process is akin to sheaving.

The street directory example is intended to bootstrap some basic intuitions about presheaves and sheaves. A presheaf, or a sheaf is an assignment of data (sets of elements) to (regions of) a topological space, where a sheaf is required to satisfy some additional coherence conditions. In the context of the street directory example, we have the following interpretation.

- A *topological space* is a set of elements together with a collection of its subsets, which indicates the relative proximity of those elements (definition 6 in **Appendix** of the Supplementary Material). The collection (set) of subsets is called the *topology* of the space, and the subsets are the *open sets* of the topology. So, the pages of the street directory correspond to open sets.

- A *presheaf* is an assignment of data to the open sets of the topological space. The assignment is given by a function that sends each open set to some set (definition 16 in **Appendix** of the Supplementary Material). The data in the current example are the markings that constitute the street map *on* each page of the directory. The assignment of data to the open sets is required to satisfy a certain *restriction* condition, which says that the data assigned to an open set $V \subseteq U$ is the data assigned to open set U restricted to V (definition 4 in **Appendix** of the Supplementary Material). For instance, this condition says that the markings on two adjacent pages treated as a single page restrict to the markings on the individual pages. For a presheaf, in general, the restrictions of data to the intersection of open sets need not agree. This situation occurs when, for example, the map for an urban area may contain more detailed information than the map for an adjacent nature reserve so that the two sets of landmarks shown for the overlapping area are not the same set.
- A *sheaf* is a presheaf such that the data for overlapping open sets is the same (definition 17 in **Appendix** of the Supplementary Material). There are two conditions for a presheaf to be a sheaf (definition 17 in **Appendix** of the Supplementary Material). The requirement that the data agree on overlaps is called the *gluing* condition; the requirement that the gluing be unique is called the *locality* condition, which essentially says that there is no ambiguity in the way local information is patched together.
- *Sheaving* (definition 18 in **Appendix** of the Supplementary Material) is a universal way (theorem 1 in **Appendix** of the Supplementary Material) of constructing the “best” possible sheaf from a given presheaf (example 9 in **Appendix** of the Supplementary Material).

The foregoing illustration, though intuitive, glosses over important details that may leave one questioning the motivation for a sheaf theory approach. Firstly, why are we concerned with the more abstract notion of topological space, rather than the more concrete notion of coordinate space used, for example, in connectionist models where the vectors take on real numbers? Secondly, how are data attached to a topological space? Thirdly, in what sense does sheaving return the “best” possible sheaf for a given presheaf? Each question is addressed in the next three sections, in turn.

2.1. A Topological View of Space

The street map example may leave one wondering about the need to work with a topological space. A topological approach is appropriate when there is no suitable notion of “distance,” between representations, as required by a *metric space*. The extra abstraction also affords a parsimonious treatment of symbolic and numeric (coordinate) representations, as both sets of symbols and sets of numbers can be given a topology.

To illustrate how symbols have a (topological) order without defining a distance measure, suppose we have a set of abstract symbols, $X = \{A, B, C\}$. A topology on X is a collection T of subsets of X that, at least, includes the empty set, \emptyset , and X . These subsets are designated as the *open sets* of T and indicate the relative proximity of the elements in X . For example, suppose

$T = \{\emptyset, \{C\}, \{B, C\}, \{A, B, C\}\}$. This topology is a *specialization order topology* corresponding to the order $A \leq B \leq C$, which says that B is closer to C than A .

Every set X can be given two extreme topologies (example 1 in **Appendix** of the Supplementary Material), which have associated orders. One extreme is called the *indiscrete topology*, which contains just the empty set and X . The other extreme is called the *discrete topology*, which contains every subset of X . The (pre)order associated with the indiscrete topology on $\{A, B\}$ has the order relations $A \leq B$ and $B \leq A$, and the order associated with the discrete topology has just the order relations $A \leq A$ and $B \leq B$. Both can be interpreted as reflecting “minimal” information about the proximity of A and B .

Open sets are fundamental to the topological notion of space and continuity. Not only do they indicate proximity, but also how the regions of space relate to each other via inclusion. If two regions are open sets, then their intersection and union are also open sets (regions) of that space. As we shall see, next, intersections pertain to gluing and unions pertain to coverage. We have already introduced the importance of gluing, the importance of coverage is similarly efficacious—we require as many pages as needed to cover all city regions. Likewise, a representational space should cover the things that need to be represented. The corresponding concept in topology is *open cover* (definition 8 in **Appendix** of the Supplementary Material).

Perhaps less obvious is the role of open sets to the property of being continuous. A function between topological spaces is called *continuous* if the *preimage* of an open set is an open set (definition 7 in **Appendix** of the Supplementary Material)—a continuous transformation obtains closely related things from closely related things. Notice that this topological definition admits continuous functions over “discrete” symbols. Being continuous is a property of a function, not a space. Connectionist representations are sometimes regarded as “continuous” and symbolic representations as “discrete.” However, this difference is more akin to the difference between countable vs. uncountable sets: e.g., the set of natural numbers vs. the set of real numbers.

2.2. A Relational View of Data

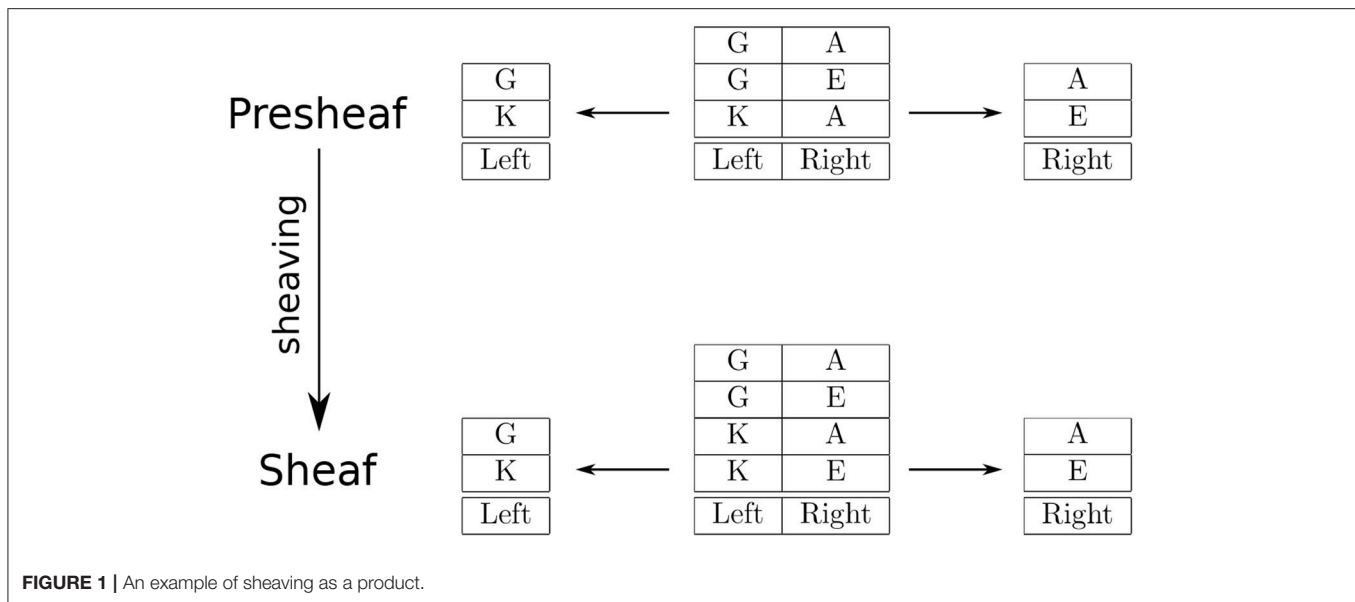
The foregoing conception of representational space lays the groundwork, as it were, for a parsimonious treatment of representation as data attached to a (topological) space. A crucial observation is that relational databases can be viewed as presheaves, or sheaves (Abramsky and Brandenburger, 2011) (Note that presheaf/sheaves are more general constructions than relational tables, because every point/column, or combination of points/columns need not constitute an open set of the topology). A relational database table consists of a header, listing the names of each column of a table, and rows that contain the data for that table. A collection of tables constitutes a relational database, and the headers constitute the relational schema. In terms of sheaf theory, the schema corresponds to a topological space, where each header is an open set, and the rows of a table correspond to the data attached to an open set. Thus, a relational database corresponds to a presheaf, or a sheaf when the tables can be glued together to give another table that is also part of the database. The relational database operation that realizes

gluing is called the *natural join* (example 6 in **Appendix** of the Supplementary Material). A natural join, described in detail next, combines two tables into a single table whose rows are just those constructed by combining (joining) the rows from each table that have the same value at the columns in common. When there are no columns with the same name, i.e., the intersection of the table headers is the empty set, the natural join reduces to the (Cartesian) product of the two tables. The next section illustrates this situation first, and the section that follows illustrates the case where the two tables have common columns. For reasons that will become clearer later, the second situation is called a *constrained product*. From a relational database view, sheaving involves constructing tables not already in the database by gluing together existing tables. This process is also described next, and applied to cognition in section 3.

2.2.1. Gluing as a Product

To illustrate gluing as a product, suppose we are given a pair of objects (characters/letters) in the visual field of view. The knowledge that the pair of characters (G, A) differs from the pair (A, G) is captured by recognizing the relative locations of each object: e.g., the first character is located at the left position, and the second character is located at the right position. Suppose the pairs (G, E) and (K, A) were also presented on separate occasions. This information is recorded in a relational database table that has two columns, named Left and Right, and three rows containing the three pairs of characters at the corresponding positions. Each row corresponds to a character pair. In addition, there is a one-column table headed Left and a one-column table headed Right, recording the locations of each character individually. The location names and headers constitute a (discrete) topological space: (Location, Proximity), where Location = {Left, Right} is the set of locations, and Proximity = $\{\emptyset, \{\text{Left}\}, \{\text{Right}\}, \{\text{Left}, \text{Right}\}\}$ is the discrete topology on that set. This situation corresponds to a presheaf, where the rows are the data attached to the topological space, see **Figure 1** (top row). (The data attached to the empty set is the singleton set $\{*\}$, i.e., the one-element set containing an element whose name is unimportant. The corresponding table, not shown, is the table with the empty header and one row containing the unnamed element.) For instance, the pair (G, A) appears as a row of the two-column table, and the individual characters as rows of the corresponding one-column tables. From a sheaf theory perspective, the topological space provides the ground on which the data are attached. Accordingly, the tables are shown with the “header” at the foot of the table—footer.

Relational databases come equipped with operations for extracting information from tables. One basic operation is called *projection*, which returns all the values at the named columns for all the rows of the specified table. For instance, the characters located at the left position are obtained by a projection onto the Left column of the two-column table (**Figure 1**, upper left arrow), and likewise for the characters located at the right position (**Figure 1**, upper right arrow). In sheaf theory terms, these projections are given by restrictions for the presheaf. Recall that a presheaf is an assignment of open sets to sets of elements that preserves inclusions as restrictions: for each inclusion in the



underlying topological space there is a corresponding restriction map. So, in this example, the restriction corresponding to the inclusion $\{\text{Left}\} \subseteq \{\text{Left}, \text{Right}\}$ is the projection onto the Left column; likewise, the restriction corresponding to $\{\text{Right}\} \subseteq \{\text{Left}, \text{Right}\}$ is the projection onto the Right column (The empty set is included in every open set, so the corresponding restrictions send every row to the only element in the singleton set).

This relational database specifies a presheaf, but not a sheaf, because the two-column table cannot be (re)constructed as the gluing (product) of the one-column tables. Specifically, the product of the one-column tables results in the row that is the pair (K, E), which is not contained in the two-column table. To be a sheaf, the gluing condition essentially says that there must be a row in the two-column table that contains (restricts to) a given pair of rows from the one-column tables, which is not the case for the pair (K, E). Thus, this presheaf is not a sheaf.

The sheaving process turns a presheaf into a sheaf by gluing along overlapping regions. In terms of relational database operators, gluing is the natural join. When the overlap is the empty set, gluing is essentially the product of tables. In this example, the product is all pairwise combinations of rows from the one-column tables. Hence, sheaving adds the (K, E) pair to the two-column table. Thus, the updated relational database corresponds to a sheaf, see **Figure 1** (bottom row). The sheaving construction is a map from the top row to the bottom row (**Figure 1**, left vertical arrow).

2.2.2. Gluing as a Constrained Product

The essential difference between gluing as a product and gluing as a constrained product is that the intersection of the underlying open sets, to which the data are attached, is not the empty set. This situation often occurs in relational databases, e.g., where personal information about employees is stored in one table and work-related information in another table, and taking the natural join on the common employee-identifier column links the two kinds of information.

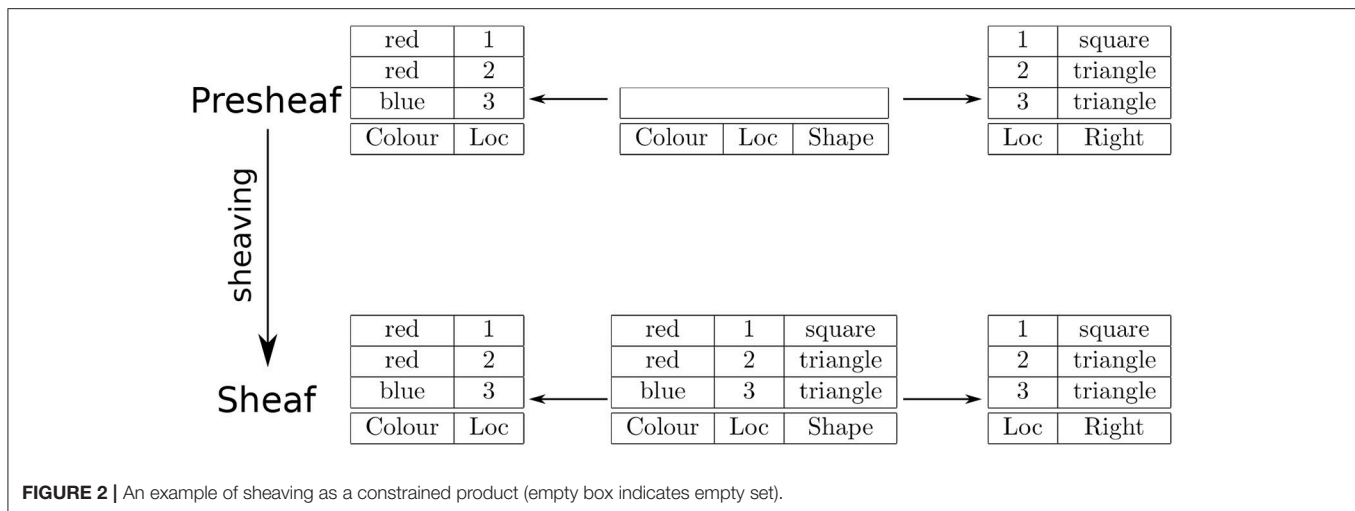
Visual cognition can be considered analogously where object features (e.g., location, color, or shape) are stored in separate tables that can be joined to recover information about feature conjunctions to identify objects, e.g., that the displayed objects are red square and green triangle, not red triangle and green square. For this situation, suppose that objects are indexed by location, and color and shape information are recorded in separate two-column tables with headers (Location, Color) and (Location, Shape), respectively. Here, we have a set of feature dimensions, $\text{Feature} = \{\text{Location}, \text{Colour}, \text{Shape}\}$, and a topology, $\text{Bind} = \{\emptyset, \{\text{Location}\}, \{\text{Location}, \text{Colour}\}, \{\text{Location}, \text{Shape}\}, \text{Feature}\}$. This topological space associates color and shape more closely to location than each other, which is interpreted as saying that color and shape feature maps are more basic than color-shape conjunction maps.

In this situation, sheaving recovers the binding of color and shape as the natural join of Location-Color and Location-Shape tables, which results in the Location-Color-Shape table corresponding to a color-shape conjunction map. A psychologically compatible interpretation of this situation is binding-by-location (Treisman, 1996). An example is shown in **Figure 2**. The natural join in this case is constrained to return only those rows that agree on location, not all combinations of rows, as in the previous example. Hence, this case is called a constrained product.

Note that for the presheaf shown in **Figure 2**, the empty box indicates that the data attached to the open set is the empty set. Hence, all restrictions from this set are empty maps.

2.3. A Category Theory View of Sheaves and Sheaving (Universality/Systematicity)

Up to this point, we have presented the basic ideas of sheaves and sheaving in terms of sets and functions. This approach is easier to grasp, but obscures the importance of universal construction and its role in an explanation for systematicity and



productivity. So, in this section, we present the category theory view of sheaves. The core concept that links sheaves, systematicity and generalization is *universal morphism* (definition 15 in **Appendix** of the Supplementary Material). This concept depends on the concepts of *category* (definition 9 in **Appendix** of the Supplementary Material) and *functor* (definition 13 in **Appendix** of the Supplementary Material), and is closely related to the concept of *natural transformation* (definition 14 in **Appendix** of the Supplementary Material). For a quick intuition, one can think of a category as a set with relations (*morphisms*) between its elements (*objects*), a functor as a function between categories that “preserves” those relations, and a natural transformation as a kind of higher-order function (i.e., a function that takes/returns a function, see remark 5 in **Appendix** of the Supplementary Material).

In the context of category theory, a topological space is a category with open sets for objects and inclusions for morphisms (example 2 in **Appendix** of the Supplementary Material). A presheaf (hence, a sheaf) is a functor from a topological space to the category set, which consists of sets for objects and functions for morphisms. This functor sends each open set to the data attached to that set, and each inclusion to the corresponding restriction. Since sheaves are functors, maps between sheaves are maps between functors, i.e., natural transformations. Thus, sheaving pertains to a particular universal natural transformation, i.e., a second-order universal morphism.

The category theory concept of universal morphism is central to an explanation of systematicity (Phillips and Wilson, 2010). Conceptually, a universal morphism is the “best” possible construction (definition 15 in **Appendix** of the Supplementary Material). We have already seen two examples: product and constrained product. In general, the categorical product of two objects A and B is the best possible way of constructing an object that affords the recovery of A and B . In the category set, the product is the Cartesian product $A \times B$ together with two functions (projections) that retrieve the first and second elements from each pair (example 2 in **Appendix** of the Supplementary Material). In regard to tables, the Cartesian product is just all pairwise combinations of rows from each table, which are retrieved by the (relational) projection operations. Similarly,

the constrained product is a universal construction: the best possible way of combining two tables so that they agree on overlapping columns, which is just the natural join. The product is a special case of the constrained product in that the agreement is automatic. In category theory, the constrained product (natural join) is an instance of the universal construction, called *pullback* (definition 12 in **Appendix** of the Supplementary Material).

The relevance of these concepts to sheaves and sheaving is two-fold. Firstly, a presheaf must satisfy the gluing (and the locality) condition to be a sheaf. From a category theory perspective, the gluing condition is given by products (pullbacks). Thus, to be a sheaf, a presheaf must satisfy a certain universality condition. Secondly, sheaving also pertains to a universal morphism in the context of a category of presheaves and presheaf morphisms. In this sense, sheaving obtains the best possible sheaf for the given presheaf. We have already explained that systematicity results from universal constructions (Phillips and Wilson, 2010). Thus, sheaving is a universal form of generalization. Since the construction returns a sheaf, which itself is a form of universal construction, sheaving pertains to a kind of second-order systematicity.

3. GOING BEYOND THE DATA: SHEAVING IN COGNITION

The sheaf theory constructions just presented are applied to cognitive domains.

3.1. Cue-Target Learning: Product

In this section, we show why generalization is afforded by sheaving for a task requiring participants to learn a set of cue-target mappings that is the product of two sets of cue-target mappings (Phillips et al., 2016). More details of the sheaf theory basis for generalization in this task are given in the “Cue-target (product) task: generalization as sheaving” section of the **Appendix** in Supplementary Material.

The task was to learn cue-target maps where the cues were pairs of characters and the targets were colored shapes, e.g., $(G, K) \mapsto (\text{red}, \text{square})$, $(G, P) \mapsto (\text{red}, \text{triangle})$, and so

on. In the product condition, the map was the product of a map from characters to colors and a map from characters to shapes, e.g., $G \mapsto \text{red}$ and $K \mapsto \text{square}$, etc. The motivation for this task was to test the hypothesis that systematicity, or failure to exhibit systematicity is due to a cost/benefit trade-off: for a small number of mappings participants were expected to learn the training set without the overhead of inducing the universal (product) construction and thereby not demonstrate generalization to novel pairs (testing set); for a larger number of mappings, where the demand on learning each pair separately becomes excessive, participants were expected to induce the product construction and thereby demonstrate generalization (systematicity). Experimental results supported these predictions (Phillips et al., 2016).

The cue-target task investigated conditions that elicit universal constructions, hence systematicity. Here, we are interested in why such constructions are generated. According to the sheaf theory account, the relevant universal construction is a sheaf, which is obtained from another universal construction, sheaving. We have already shown how a product results from the sheaving process, in the previous section. Here, we show that the product map is a result of a sheaf morphism. From the relational database view of sheaves, a (pre)sheaf morphism is a map between relational databases. Sheaving as a basis for generalization is shown in **Figure 3**, where the sheaving constructions are the horizontal arrows, the presheaf morphism obtained from the training set is the left vertical arrow, and the sheaf morphism obtained from sheaving is the right vertical arrow, which affords generalization on the testing set. This arrangement is an instance of the commutative diagram (square) for a natural transformation (diagram 4 in **Appendix** of the Supplementary Material). Note that the morphism relating the test cues back to the training cues is given by the fact that sheaving involves an *adjoint functor* (see remark 16 in **Appendix** of the Supplementary Material). In psychological terms, participants recognize the importance of decomposing a pair of characters into their component characters: responses to novel pairs of characters can be determined by the responses to the individual characters as they appeared in other pairs during training.

The product task revealed that participants failed to exhibit systematicity for small maps even though the training set contained sufficient information to specify the underlying product and participants correctly learned the cue-target mappings for that set. From a sheaf theory perspective, this failure to demonstrate systematicity results from failure to identify the appropriate underlying topology. In this case, we regard the training set as a presheaf on an indiscrete topological space, in contrast to a discrete topological space. Recall, that an indiscrete topology on a set X consists of just the empty set and X as the open sets. A presheaf on an indiscrete topological space is trivially a sheaf: the gluing condition is automatically satisfied, because there is only one nonempty open set. Sheaving is just the identity transformation in this case, so no new rows are added to the table, hence participants do not go beyond the training data.

A psychological interpretation is that participants fail to recognize/represent the appropriate order relationship between the points of the space that correspond to the dimensions of the

task, which impacts upon generalization. Recall (section 2.1) that a two-point space with the indiscrete topology corresponds to the preordered set: e.g., $A \leq B$ and $B \leq A$, where A and B are the two points. The two points (dimensions) are equivalent, effectively regarded as a single point, hence sheaving has no effect in terms of generalization. By contrast, the discrete topology corresponds to the discrete ordered set: $A \leq A$ and $B \leq B$, effectively regarding the two dimensions as independent, which affords generalization via sheaving. This difference can be interpreted as attentional load: spatial attention to stimuli as data attached to one vs. two locations.

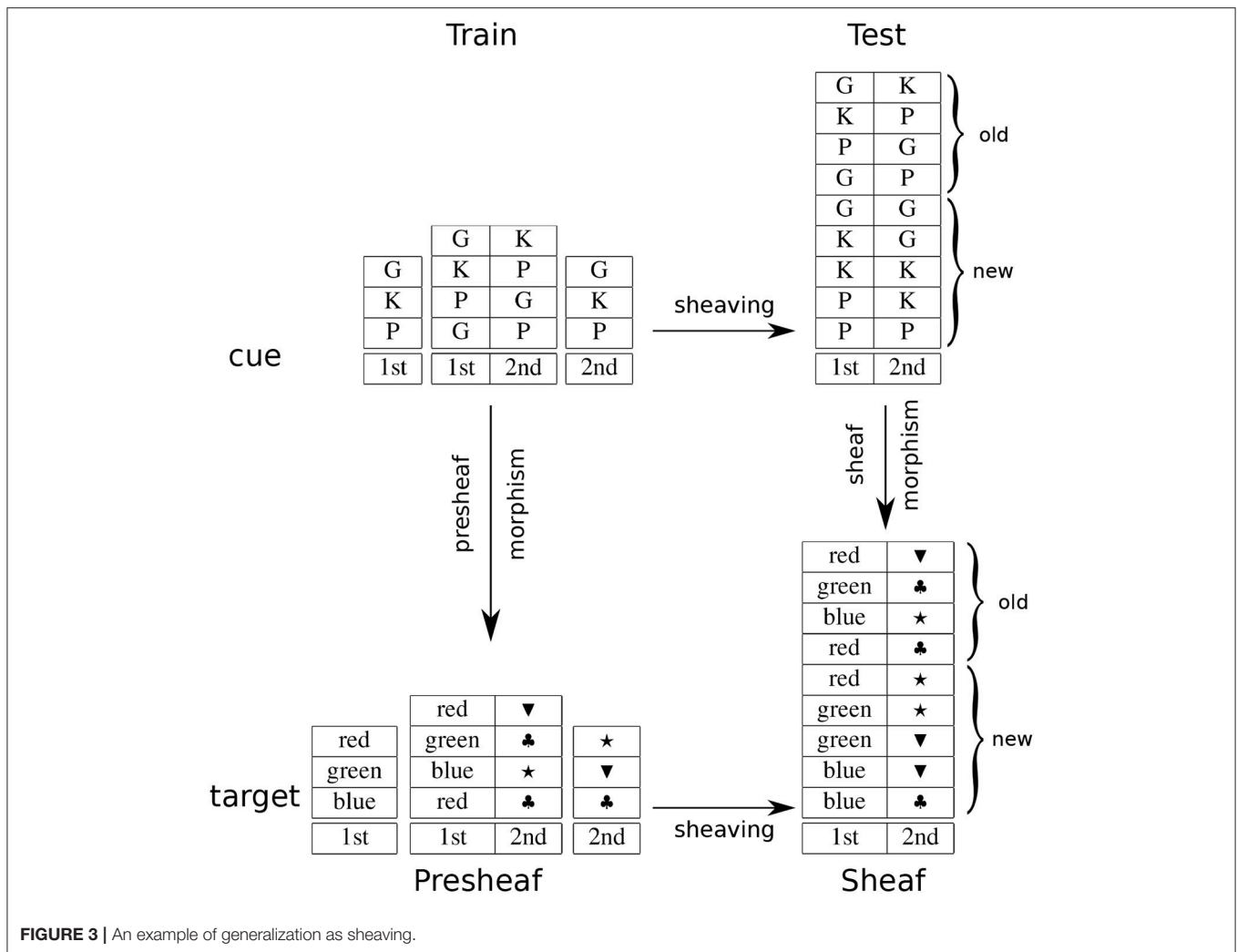
Analysis of response data based on participant self-reports (Phillips et al., 2016) lends support to this interpretation. Upon completion of the experiment, participants were asked to report on how they performed the task. Participants were then divided into two groups indicating whether or not they were aware of the product structure of the mapping task. The aware group showed the same effects as observed in the original analysis. By contrast, the unaware group were not significantly above chance level performance in all conditions.

3.2. Visual Search: Constrained Product

In a visual search task, participants are required to locate an object, designated as the target of search, in a display also containing nontargets. Typically, the target is uniquely identifiable by one or more features, e.g., color, shape, or orientation. The time to locate the target as a function of the number of objects in the search field is called the *search slope*. Search slope is typically shallower when targets can be identified by a single feature than when targets are identifiable by a conjunction of two or more features, e.g., color and shape (see Wolfe, 2003, for a review). Such behavioral differences led to the well-known and influential Feature Integration Theory of visual attention (Treisman and Gelade, 1980; Treisman and Sato, 1990), see Humphreys (2016) for a recent review. Although search slope may not be indicative of feature (shallow) vs. conjunctive (steep) search (Wolfe, 1998), recordings of monkey cortical activity support a feature vs. conjunction mode of attention (Buschman and Miller, 2007).

There is an obvious cost/benefit trade-off associated with having a features-detection system on one hand and a conjunctive construction system on the other. Dedicated feature units afford rapid response, but a unit is needed for each possible feature; (re)constructing conjunctions of features with dynamically reconfigurable units (i.e., units that can represent more than one conjunction) requires fewer units, but more time to detection. This trade-off is analogous to a trade-off in relational database design in that large tables are typically (re)constructed from smaller tables to save space, as well as to maintain data integrity (Halpin, 1995), but at the expense of longer query times. Construction of feature conjunctions is formally a natural join (Phillips et al., 2012, Text S3), hence it involves a sheaving process.

There is also a systematicity property in regard to feature binding: if one has the capacity to bind say features red and square, and features green and triangle, then one also has the capacity to bind features red and triangle, and green and square,



regardless of whether one has seen that exact combination of features before. This property raises the familiar challenge of explaining why such a property exists: Why does having the capacity to bind, say, red with triangle and blue with square imply having the capacity to bind red with square and blue with triangle, assuming the capacity to recognize red, blue, triangle and square? The sheaf theory explanation is that systematicity of conjunctive features follows from a categorical universal construction, sheaving, which involves the natural join of feature maps, as explained in section 2.2.2. The universal morphism explanation for the systematicity of binding as a constrained product parallels the universal morphism explanation for cue-target pairs as products given in the previous section: the conditions for being a universal morphism (pullback) imply just those combinations.

Sheaving comports with the primacy of location-based feature maps (Riesenhuber and Poggio, 2004). In terms of the underlying topological space, the color and shape feature dimensions are closer to the location dimension than the color and shape dimensions are to each other. Accordingly, color-location and shape-location information are computed before

color-shape-location information, which is typically expressed as faster response times (shallower search slopes) for feature than conjunction search (Treisman and Gelade, 1980; Wolfe, 2003).

The importance of the topology is reflected in the implications for binding. Dimensions are typically regarded as orthogonal and independent (as in the cue-target example of the previous section), which corresponds to a discrete topology. However, the discrete topology generates all possible conjunctions of features, not just those bindings present in the field of view.

Note that for ease of exposition, location is identified by a label/symbol for the example shown in **Figure 2**. However, location can also be modeled as a topological space, e.g., the product of topological spaces modeling the horizontal and vertical axes of two-dimensional display screen. Indeed, a parsimonious treatment of symbolic and spatial forms was one of the motivations for taking a sheaf theory approach, as foreshadowed in the Introduction.

In this example, we concerned ourselves with just the construction of representations for conjunctions of features, not with the process of searching for the target given those representations. A category theory approach to visual search

has been discussed elsewhere (Phillips and Takeda, 2017). The theory employed there introduces another form of pullback, involving a change of base, that is beyond the scope of our current concern. Also, we have not considered the learning/development of conjunction search: e.g., young children are less efficient at conjunction search than older children and adults (Merrill and Lookadoo, 2004). Here and in the previous example, we concerned ourselves with representations and processes that pertain to a single topological space. More general situations that require changing the topological space are discussed in the next section.

4. DISCUSSION

Our main purpose in this paper has been to (re)conceptualize generalization as sheaving: a process of “putting two and two together to make five,” so to speak. In the service of understanding cognition, sheaf theory appears to be a relatively unexplored area of mathematics—see, e.g., Goguen (1992) and Malcolm (2009) for applications to the related area of distributed systems, and Goguen (2018) for a discussion in regard to information integration. In this section, we discuss the prospects of a sheaf theory approach to learning and generalization, generally.

From a sheaf theory perspective, going beyond the data is about patching (or, gluing) local information to obtain new knowledge. The core property that affords sheaving is the ability to form the product of pieces of local knowledge constrained by their common source. So, by this account, sheaving should be evident in other cognitive abilities where products play a key role. Cognitive abilities such as matrix reasoning and transitive inference come to mind. A matrix reasoning task typically consists of a matrix of items (e.g., colored shapes) with the goal of identifying the item that goes in the empty cell location (e.g., *Raven's Progressive Matrices* Raven et al., 1998). For a relatively simple example, suppose that the rows are identified with colors: red, green and blue, and the columns are identified with shapes: circle, triangle and square, in those orders. The target that goes in the cell located at the third row and column is a blue square. This situation is similar to the cue-target learning task in section 3.1, as both involve a product of two dimensions. The topology consists of the two dimensions as open sets, and sheaving obtains the target by the product of the shape and color features attached to their respective dimensions.

More complex examples of matrix reasoning involve relations between the items within rows or columns. Models have been developed to account for simple and complex forms of matrix reasoning (Carpenter et al., 1990; Lovett et al., 2010). From our viewpoint, these situations involve data that have more internal structure than sets. The category theory approach to sheaves extends naturally to such cases as functors from a topological space to some other kind of category that has products, e.g., a category whose objects are groups, or rings (i.e., sets with one, or two internal operations). A challenge for the sheaf theory approach is to model both simple and complex forms of matrix reasoning.

Another cognitive ability pertaining to constrained products is *transitive inference*. Transitive inference has the form, if A is

R -related to B and B is R -related to C , then A is R -related to C , where the relation R has the transitivity property. For example, if John is shorter than Mary and Mary is shorter than Tom, then John is shorter than Tom. In this situation, the premises are given by the order topology: $P = \{\emptyset, \{P_2\}, \{P_1, P_2\}\}$ and $Q = \{\emptyset, \{Q_2\}, \{Q_1, Q_2\}\}$, where P and Q are the order topologies for the premises John is shorter than Mary and Mary is shorter than Tom, respectively. A capacity for transitive inference is regarded as crucially depending on an ability to integrate the premises into an ordered triple (Maybery et al., 1986; Andrews and Halford, 1998). In topological terms, integration corresponds to attaching data to a topology that encodes the three-term order, e.g., $T = \{\emptyset, \{T_3\}, \{T_2, T_3\}, \{T_1, T_2, T_3\}\}$. Modeling this situation requires methods for changing the topological space. Here, category theory is again useful as there are two functors for changing the topology of a sheaf (presheaf): the *direct image functor* and the *inverse image functor* (Hartshorne, 1977; Mac Lane and Moerdijk, 1992). Another challenge, then is to model various aspects of transitive inference, including *pseudo-transitive inference* (Goodwin and Johnson-Laird, 2008), where the elements of the premises are locally, but not globally ordered.

A sheaf theory approach may also have something to say about the development of transitive inference and other reasoning tasks in terms of the development of the underlying topological space. Young children (below about 5 years of age) repeatedly have been shown to lack a capacity for transitive inference and a range of other reasoning tasks (Halford, 1984; Andrews and Halford, 1998, 2002). Some have argued that such capacities turn on the development of relational information processes (Halford et al., 1998, 2014; Penn et al., 2008), which has also been given a category theory account (Phillips et al., 2009). The category theory perspective attributed the difference to a capacity for products, including constrained products (pullbacks). We have already seen how these constructions are related to presheaves and sheaves, and the underlying topology. The sheaf theory approach presented here provides another related perspective on the development and evolution of intelligence, i.e., as a capacity to represent space. In particular, we noted that every set has two extreme topologies: indiscrete and discrete. For the collection of topological spaces on a given set, the indiscrete and discrete spaces are, respectively, the *coarsest* and *finest* topologies that can be given for that set, which are themselves instances of particular universal constructions. The relative coarseness/fineness of the underlying space alludes to developing progressively coarser/finer capacity to make spatial distinctions. For example, young children represent changes in shape differently than adults (Abecassis et al., 2001). The progression from holistic to category (class) based processes has been modeled computationally as learning/development via “intersection discovery” (Doumas and Hummel, 2010), by a symbolic connectionist model (DORA; Doumas et al., 2008). In our sheaf theory view, intersection discovery connotes development of a topological space.

The process of intersection discovery in DORA raises the possibility of developing a neural semantics for our category/sheaf theoretic approach to systematicity and generalization, as the pullback is a kind of intersection: in

the category of sets and inclusions the pullback is just set intersection; in the category of sets and functions the pullback is the set of points that intersect (agree) on their images. DORA uses the role-filler binding method of the LISA model (Hummel and Holyoak, 2003) to induce relational representations via the interaction between proposition units representing relations, role-filler units representing the binding of values to relational roles, and feature units representing features of the related fillers (values)—role-filler units that coactivate the same feature units tend to be bound together by units representing a common relation. Conceptually, this arrangement is akin to a pullback of functions $f: A \rightarrow C$ and $g: B \rightarrow C$, where the feature units correspond to the constraining object C , the interaction between role-filler and feature units to f and g , and the pullback object $A \times_C B$ to the units representing the relation. The dynamics of the DORA model are more complex than projections. So, the extent of a formal connection is not yet known. Developing a neural model for the theory would provide a basis for cost in terms of the neural resources needed to realize a universal construction.

Whether similar considerations apply to the development of conjunction search is a topic for future work. A capacity to represent conjunctions is just one aspect of visual attention, and there are multiple possible reasons for a change in search efficiency with age (see Merrill and Lookadoo, 2004, for a discussion). Here, we simply note that the pullback of two morphisms $f: A \rightarrow C$ and $g: B \rightarrow C$ is constrained by C . Thus, changing C (which means changing f and g) can change the number of elements constructed by the pullback, hence the number of elements selected for search, and thereby search efficiency.

Other potential applications are probability judgements that violate classical probability laws, e.g., *conjunction fallacy* (Tversky and Kahneman, 1983). In this situation, people judge the conjunction of two events A and B as more likely than either event A or event B : e.g., $P(A \wedge B) \geq P(A)$, which violates the classical probability law, $P(A \wedge B) \leq P(A)$. *Quantum probability theory* was introduced to explain such fallacies (see Busemeyer and Bruza, 2012, for an overview of theory and example applications). An important feature of this theory is contextuality where the act of measuring affects the outcome. The conditions for having quantum-like contextuality effects are closely related to the conditions for being a presheaf, but not a sheaf (Abramsky and Brandenburger, 2011). In these situations, the points of the topological space are measures and the values (data attached to the space) our outcomes, or outcome probabilities. The close connection between presheaves/sheaves and contextuality suggests that sheaf theory can also be applied to address contextuality effects in cognition.

Presheaves involve three kinds of morphisms, in addition to inclusions and restrictions: (1) morphisms from the topological space to the data, i.e., presheaves/sheaves, (2) morphisms between presheaves, i.e., presheaf morphisms, and (3) morphisms from sheaves on one topological space to sheaves on another topological space. We have primarily concerned ourselves with the second kind, in the form of sheaving, with regard to the

generalization aspect of learning. However, for a more complete picture, we also need to consider how the first and third kinds of morphism pertain to other aspects of learning.

The first kind of morphism is important with regard to training and the partial state of knowledge acquisition. In particular, one difficulty with a category theory approach to cognition is how to model partial knowledge (Navarrete and Dartnell, 2017). With sheaf theory, partial knowledge can be related to the data attached to open sets and their restrictions. For example, the presheaf in **Figure 2** has the empty set as the data attached to the open set, Feature, hence the associated restriction maps are empty maps. This situation reflects a temporary state of having partial (no) knowledge about, or representation of color-shape binding. In the context of learning, partial acquisition of knowledge can be modeled as a subset of the sections (rows) attached to an open set. How data get attached to open sets as a result of learning is a topic of future work.

The third kind of morphism is important in regard to explaining the transition from non-systematicity to systematicity. As mentioned in section 2.2.1, failure to generalize for small tasks can be attributed to having an indiscrete topological space. However, this account raises the question of why/how participants recognize the need to (re)represent space as a discrete topology to afford generalization. The cost/benefit hypothesis (Phillips et al., 2016) may help here, because the discrete topology (generally) contains more open sets than the indiscrete topology on the same set, hence requires more resources to represent. How cost/benefit interacts with learning in a sheaf theory setting is also a topic for further research.

Throughout this paper, we have focussed only on interpreting the gluing condition for a sheaf as a formal basis for the ability to go beyond the data. However, a presheaf must also satisfy the locality condition to be a sheaf. The locality condition says that gluing must be unique. Sheaving in this situation essentially identifies the alternatives as being the “same” data up to an equivalence. A detailed exposition will take us too far afield, however, this situation is like treating different instances of an object as the same object up to some equivalence relation. This situation typically does not arise in a relational database, because the relational database schema is essentially treated as a discrete topological space, in which case all rows must be unique, i.e., the locality condition is automatically satisfied. In a cognitive context, the locality condition may also have interpretations in terms of treating two distinct entities as the same thing: generalization on the basis of object class, rather than object instance.

Some authors have argued that the architecture of cognition (i.e., the basic processes and ways of combining such processes to afford cognition) is a “kludge” of disparate abilities that are somehow patched together to give the illusion of a well-organized system (Clark, 1987; Marcus, 2008). Be that as it may, viewing cognitive architecture as a hodgepodge of subsystems begs the question of why the system does actually work coherently, for the most part. The sheaf theory view presented here says that patching is a universal construction: an optimal solution to reconciling differences between subsystems put together as a kludge.

A sheaf can be likened to a kind of analogy in that the relations (inclusions) in the source domain (topological space) are mapped to relations (restrictions) in the target domain (data attached to the space), cf. *structure mapping theory* (Gentner, 1983). Category theory has been used as a basis for children's difficulty with understanding and exploiting the common relations in a reasoning problem (Halford and Wilson, 1980), and as an approach to analogy (Navarrete and Dartnell, 2017). However, a sheaf is a *contravariant* functor: the directions of arrows in the source are reversed in the target, which may strike some people as puzzling, given that analogy is typically conceptualized as a *covariant* mapping: the directions of arrows in the source and target are the same. One can conceptualize the role of contravariance in sheaf theory as persistence. Topological spaces can be built up by taking intersections and unions of the open sets in a basis set. Inclusions order open sets by size from small to large. A global property can be regarded as a property that *persists* over all the open sets—a property that is systematic as opposed to idiosyncratic (specific) to just some open sets—as we zoom in on smaller regions of space.

An important topic for further work is to explain how the cost/benefit proposal is supposed to interact with the construction of universal morphisms, as mentioned in the Introduction. An apparently straightforward approach would be to assign a cost to the alternative routes (see Phillips and Wilson, 2016b, for a discussion). However, this approach requires independent justification for the costs assigned, lest the cost/benefit principle becomes another *ad hoc* assumption, i.e., an assumption motivated solely to fit the data (Aizawa, 2003). Independent motivation may come in the form of empirical measures of the cost of each supposed alternative. For such purposes, a split-screen paradigm was developed to examine cost in the context of feature vs. conjunction visual search (Phillips et al., 2017). In this paradigm, participants could search for the target object in either the left or right

visual field, which corresponded to feature or conjunction search. Search time when only one field was presented provided independent measures of the baseline costs of feature and conjunction search, which were then used to assess whether participants chose the alternative of least cost when both alternatives were presented at the same time. Analysis indicated that the choice of search field depended not only on the relative costs of the alternatives, but also on the cost of that assessment (Phillips et al., 2017). In this way, a categorical account of least cost may provide a principled explanation for the interaction between cost/benefit and universal construction, and its implications for systematicity and generalization.

Going beyond the data is a ubiquitous cognitive capacity in need of a theoretical explanation to motivate modeling as more than just an exercise in data fitting. The theoretical picture painted here is a view beyond local perception of the world. This sheaf theory approach formalizes our propensity to connect the dots. After all, that's what people do.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by a Japanese Society for the Promotion of Science Grant-in-aid (16KT0025).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01926/full#supplementary-material>

REFERENCES

- Abecassis, M., Sera, M. D., Yonas, A., and Schwade, J. (2001). What's in a shape: children represent shape variability differently than adults when naming objects. *J. Exp. Child Psychol.* 78, 303–326. doi: 10.1006/jecp.2000.2573
- Abramsky, S., and Brandenburger, A. (2011). The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.* 13:113036. doi: 10.1088/1367-2630/13/11/113036
- Aizawa, K. (2003). *The Systematicity Arguments*. Studies in Mind and Brain. New York, NY: Kluwer Academic.
- Andrews, G., and Halford, G. S. (1998). Children's ability to make transitive inferences: the importance of premise integration and structural complexity. *Cogn. Dev.* 13, 479–513. doi: 10.1016/S0885-2014(98)90004-1
- Andrews, G., and Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cogn. Psychol.* 45, 153–219. doi: 10.1016/S0010-0285(02)00002-6
- Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071
- Busemeyer, J. R., and Bruza, P. D. (2012). *Quantum-Inspired Models of Concept Combination*. Cambridge: Cambridge University Press.
- Calvo, P., and Symons, J. (eds.) (2014). *The Architecture of Cognition: Rethinking Fodor and Pylyshyn's Systematicity Challenge*. Cambridge, MA: MIT Press.
- Carpenter, P. A., Just, M. A., and Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices test. *Psychol. Rev.* 97, 404–431. doi: 10.1037/0033-295X.97.3.404
- Chomsky, N. (1980). *Rules and Representations*. New York, NY: Columbia University Press.
- Clark, A. (1987). The kludge in the machine. *Mind Lang.* 2, 277–300. doi: 10.1111/j.1468-0017.1987.tb00123.x
- Clark, S., Coecke, B., and Sadrzadeh, M. (2008). “A compositional distributional model of meaning,” in *Proceedings of the Second Symposium on Quantum Interaction*, eds P. Bruza, W. Lawless, K. van Rijsbergen, D. Sofge, B. Coecke, and S. Clark (Oxford: College Publications), 133–140.
- Doumas, L. A., and Hummel, J. E. (2010). A computational account of the development of the generalization of shape information. *Cogn. Sci.* 34, 698–712. doi: 10.1111/j.1551-6709.2010.01103.x
- Doumas, L. A., Hummel, J. E., and Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychol. Rev.* 115, 1–43. doi: 10.1037/0033-295X.115.1.1
- Eilenberg, S., and Mac Lane, S. (1945). General theory of natural equivalences. *Trans. Am. Math. Soc.* 58, 231–294. doi: 10.1090/S0002-9947-1945-0013131-6
- Evans, J. S. (2003). In two minds: dual-process accounts of reasoning. *Trends Cogn. Sci.* 7, 454–459. doi: 10.1016/j.tics.2003.08.012
- Evans, J. S., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241. doi: 10.1177/1745691612460685

- Fodor, J. A. (1975). *The Language of Thought*. Language and Thought. New York, NY: Crowell.
- Fodor, J. A., and Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71. doi: 10.1016/0010-0277(88)90031-5
- Frank, S. L., Haselager, W. F., and van Rooij, I. (2009). Connectionist semantic systematicity. *Cognition* 110, 358–379. doi: 10.1016/j.cognition.2008.11.013
- Gardenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. New York, NY: MIT Press.
- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cogn. Sci.* 7, 47–59. doi: 10.1207/s15516709cog0702_3
- Goguen, J. (1992). Sheaf semantics for concurrent interacting objects. *Math. Struct. Comput. Sci.* 2, 149–191. doi: 10.1017/S0960129500001420
- Goguen, J. (2018). “Information integration in institutions,” in *Papers in Honor of Jon Barwise*, ed L. Moss (Stanford, CA: CSLI Publications).
- Goodwin, G. P., and Johnson-Laird, P. N. (2008). Transitive and pseudo-transitive inferences. *Cognition* 108, 320–352. doi: 10.1016/j.cognition.2008.02.010
- Hadley, R. F. (1994). Systematicity in connectionist language learning. *Mind Lang.* 9, 247–272. doi: 10.1111/j.1468-0017.1994.tb00225.x
- Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cogn. Psychol.* 16, 65–93.
- Halford, G. S., and Wilson, W. H. (1980). A category theory approach to cognitive development. *Cogn. Psychol.* 12, 356–411. doi: 10.1016/0010-0285(80)90014-6
- Halford, G. S., Wilson, W. H., Andrews, G., and Phillips, S. (2014). *Categorizing Cognition: Toward Conceptual Coherence in the Foundations of Psychology*. Cambridge, MA: MIT Press.
- Halford, G. S., Wilson, W. H., and Phillips, S. (1998). Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behav. Brain Sci.* 21, 803–831. doi: 10.1017/S0140525X98001769
- Halpin, T. A. (1995). *Conceptual Schema and Relational Database Design, 2nd Edn.* Sydney, NSW: Prentice Hall.
- Hartshorne, R. (1977). *Algebraic Geometry*, vol. 52 of *Graduate Texts in Mathematics*. New York, NY: Springer-Verlag.
- Holyoak, K. J., and Hummel, J. E. (2000). “The proper treatment of symbols in a connectionist architecture,” in *Cognitive Dynamics: Conceptual Change in Humans and Machines*, eds E. Deitrich and A. Markman (Mahwah, NJ: Erlbaum), 229–263.
- Hummel, J. E., and Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychol. Rev.* 110, 220–264. doi: 10.1037/0033-295X.110.2.220
- Humphreys, G. W. (2016). Feature confirmation in object perception: feature integration theory 26 years on from the Treisman Bartlett lecture. *Q. J. Exp. Psychol.* 69, 1910–1940. doi: 10.1080/17470218.2014.988736
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.
- Lovett, A., Forbus, K., and Usher, J. (2010). “A structure-mapping model of Raven’s progressive matrices,” in *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, eds S. Ohlsson and R. Catrambone (Austin, TX: Cognitive Science Society), 2761–2766.
- Mac Lane, S. (1998). *Categories for the Working Mathematician, 2nd Edn.* Graduate Texts in Mathematics. New York, NY: Springer.
- Mac Lane, S., and Moerdijk, I. (1992). *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. New York, NY: Springer.
- Malcolm, G. (2009). Sheaves, objects, and distributed systems. *Electr. Notes Theor. Comput. Sci.* 225, 3–19. doi: 10.1016/j.entcs.2008.12.063
- Marcus, G. F. (2008). *Kludge: The Haphazard Construction of the Human Mind*. New York, NY: Houghton Mifflin Company.
- Maybery, M. T., Bain, J. D., and Halford, G. S. (1986). Information processing demands of transitive inference. *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 600–613. doi: 10.1037/0278-7393.12.4.600
- McLaughlin, B. P. (2009). Systematicity redux. *Synthese* 170, 251–274. doi: 10.1007/s11229-009-9582-0
- Merrill, E. C., and Lookadoo, R. (2004). Selective search for conjunctively defined targets by children and young adults. *Cogn. Psychol.* 89, 72–90. doi: 10.1016/j.jecp.2004.04.001
- Navarrete, J. A., and Dartnell, P. (2017). Towards a category theory approach to analogy: analyzing re-representation and acquisition of numerical knowledge. *PLoS Comput. Biol.* 13:e1005683. doi: 10.1371/journal.pcbi.1005683
- Penn, D. C., Holyoak, K. J., and Povinelli, D. J. (2008). Darwin’s mistake: explaining the discontinuity between human and nonhuman minds. *Behav. Brain Sci.* 31, 109–178. doi: 10.1017/S0140525X08003543
- Phillips, S. (2018). “What underlies dual-process cognition? adjoint and representable functors,” in *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, eds C. Kalish, M. Rau, J. Zhu, and T. T. Rogers (Austin, TX: Cognitive Science Society), 2250–2255.
- Phillips, S., and Takeda, Y. (2017). Mathematical fixation: search viewed through a cognitive lens. *Behav. Brain Sci.* 40, 40–41. doi: 10.1017/S0140525X16000224
- Phillips, S., Takeda, Y., and Singh, A. (2012). Visual feature integration indicated by phase-locked frontal-parietal EEG signals. *PLoS ONE* 7:e32502. doi: 10.1371/journal.pone.0032502
- Phillips, S., Takeda, Y., and Sugimoto, F. (2016). Why are there failures of systematicity? The empirical costs and benefits of inducing universal constructions. *Front. Psychol.* 7:1310. doi: 10.3389/fpsyg.2016.01310
- Phillips, S., Takeda, Y., and Sugimoto, F. (2017). Dual-routes and the cost of determining least-cost. *Front. Psychol.* 8:1943. doi: 10.3389/fpsyg.2017.01943
- Phillips, S., and Wilson, W. H. (2010). Categorical compositionality: a category theory explanation for the systematicity of human cognition. *PLoS Comput. Biol.* 6:e1000858. doi: 10.1371/journal.pcbi.1000858
- Phillips, S., and Wilson, W. H. (2016a). Second-order systematicity of associative learning: a paradox for classical compositionality and a coalgebraic resolution. *PLoS ONE* 11:e0160619. doi: 10.1371/journal.pone.0160619
- Phillips, S., and Wilson, W. H. (2016b). Systematicity and a categorical theory of cognitive architecture: universal construction in context. *Front. Psychol.* 7:1139. doi: 10.3389/fpsyg.2016.01139
- Phillips, S., Wilson, W. H., and Halford, G. S. (2009). What do transitive inference and class inclusion have in common? Categorical (co)products and cognitive development. *PLoS Comput. Biol.* 5:e1000599. doi: 10.1371/journal.pcbi.1000599
- Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven’s Progressive Matrices and Vocabulary Scales: General Overview*. San Antonio, TX: NCS Pearson.
- Riesenhuber, M., and Poggio, T. (2004). *How the Visual Cortex Recognizes Objects: The Tale of the Standard Model*, Chap. 111. Vol. 2. Cambridge, MA: MIT Press.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (eds.) (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1 of Computational Models of Cognition and Perception* Cambridge, MA: MIT Press.
- Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* 46, 159–216. doi: 10.1016/0004-3702(90)90007-M
- Treisman, A., and Sato, S. (1990). Conjunction search revisited. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 459–478. doi: 10.1037/0096-1523.16.3.459
- Treisman, A. M. (1996). The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178. doi: 10.1016/S0959-4388(96)80070-5
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–113. doi: 10.1016/0010-0285(80)90005-5
- Tversky, A., and Kahneman, D. (1983). Extensional versus intuitive reasoning: the conjunctive fallacy in probability judgment. *Psychol. Rev.* 90, 293–315. doi: 10.1037/0033-295X.90.4.293
- van Gelder, T., and Niklasson, L. (1994). “Classicism and cognitive architecture,” in *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, eds A. Ram and K. Eiselt (Atlanta, GA: Lawrence Erlbaum), 905–909.
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychol. Sci.* 9, 33–39.
- Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends Cogn. Sci.* 7, 70–76. doi: 10.1016/S1364-6613(02)00024-4

Conflict of Interest Statement: 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.

Copyright © 2018 Phillips. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Language or motor: reviewing categorical etiologies of speech sound disorders

Kelly Farquharson *

Department of Communication Sciences and Disorders, Emerson College, Boston, MA, USA

Keywords: speech sound disorders, speech development, motor ability, language development, speech language therapy

OPEN ACCESS

Edited by:

Rick Dale,
University of California, Merced, USA

Reviewed by:

Nikita Kuznetsov,
University of North Carolina
Greensboro, USA

*Correspondence:

Kelly Farquharson
kelly_farquharson@emerson.edu

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 13 August 2015

Accepted: 23 October 2015

Published: 04 November 2015

Citation:

Farquharson K (2015) Language or
motor: reviewing categorical etiologies
of speech sound disorders.
Front. Psychol. 6:1708.
doi: 10.3389/fpsyg.2015.01708

Children with speech sound disorders (SSDs) exhibit marked weakness with accurate production of age-appropriate speech sounds (Lewis et al., 2006). For some of these children, the etiology of the SSD is clear (e.g., cleft palate, a genetic syndrome, or hearing loss). For others, the cause of their aberrant speech development is unknown; this type of SSD is “functional.” Functional SSDs may eventually remediate after a course of treatment, but may also persist into adolescence or even adulthood (Felsenfeld et al., 1994). Regardless of the outcome, the underlying construct that contributes to this disorder remains elusive. The extant literature is comprised of two primary categorical constructs used to explain functional speech sound disorders: language-based deficits and motor-based deficits. Undeniably, all speech productions are both linguistic (speech sounds, meanings of words, syntax of context, etc.) and motoric (the muscle movement of the speech articulators—lips, tongue, jaw, soft palate, etc.) in nature. However, there are certainly competing theories that suggest language (Raitano et al., 2004; Sutherland and Gillon, 2005; Lewis et al., 2006; Preston and Edwards, 2007; Anthony et al., 2011) or motor (Webster et al., 2005; Newmeyer et al., 2007; Peter and Stoel-Gammon, 2008; Visscher et al., 2010; Redle et al., 2015) to be the predominant causal mechanism for persistent deficits in speech production abilities. I argue that the relation between motoric and linguistic ability is likely complimentary, rather than starkly categorical.

Although on some levels, the distinction between these two constructs may seem trivial, it is clinically prudent to consider. Presently, many school districts across the country deny services to children who have “just” articulation (i.e., motor-based) impairments. Interestingly, even children with “just” an articulation impairment have been reported to experience academic difficulties, even once the child has remediated the speech production error (Raitano et al., 2004; Farquharson, 2012, 2015). Specifically, difficulties with reading, spelling, and phonological awareness persist often throughout schooling. Studies have supported that adults with a history of speech sound disorders have more often repeated a grade in school than adults with no history. Interestingly, some children with SSD who have spelling difficulties exhibit similar error patterns within their spelling as they do in their speech (e.g., substitution of a particular phoneme, such that a word like “rain” may be spelled “wain”). Such reports would suggest that speech sound disorders are not strictly motoric in nature. However, as a field, we remain unclear on the extent to which motoric deficits contribute to SSDs and the relationship between language ability and motoric ability. That is, although the speech articulators are not independently achieving correct placement for age-appropriate speech sounds, it is often the case that the child is *able* to correctly move the articulators, but *does not* do so in connected speech. Some research indicates that this discrepancy is related to phonological representations, or the process by which linguistic/phonological information is stored within memory.

Phonological representations may be difficult to access for children with speech sound disorders due to underlying linguistic or cognitive deficits (Larrivee and Catts, 1999; Sutherland and Gillon, 2005; Farquharson, 2012, 2015). The development of phonological representations

requires specification of phonological details as well as organization of the segments of a word (Swan and Goswami, 1997). For children who have phonological weaknesses, such as those with speech sound disorders and/or dyslexia these representations do not develop properly. As a result, activities that require repeated access to these representations—reading, speaking, spelling—are difficult or impossible (Sutherland and Gillon, 2005; Preston and Edwards, 2007). Phonological forms that are more complex, have more syllables, or are less familiar will be particularly difficult. This is educationally relevant because children encounter substantial amounts of new vocabulary as they progress through school. For children with speech sound disorders, their ability to access, store, and use those words is circumscribed by their phonological deficits. However, there is a separate body of work that has provided substantial evidence that children with speech sound disorders exhibit motoric weaknesses.

Motor ability has been measured in children with speech sound disorders and has examined oral motor, fine motor, and gross motor abilities. For instance, Peter and Stoel-Gammon reported central timing deficits in children with SSD, as evidenced by weaknesses in non-word repetition, clapping imitation, and paced tapping. However, in that study, the researchers examined language skills but did not report them or use them for covariates in analyses. As such, the contribution of language, especially to non-word repetition skills, is not considered. Recently, Redle et al. (2015) reported neuroimaging and behavioral data examining the motoric abilities of children with SSDs. Their results revealed that children with SSDs exhibited weaker oral and fine motor skills compared to typically developing peers. Similar to other investigations of children with SSDs (Farquharson, 2012, 2015), Redle and colleagues found that persistent SSD group performed within the average range for language and cognitive skills, but still significantly differ from their peers. This was strong evidence to support a motoric deficit in children with an otherwise functional SSD. One caveat to this study is that the researchers gathered information regarding the children's classroom performance via parent survey. It would be interesting to gather these data directly from the classroom teacher and examine how the child is truly performing academically. It remains unclear how these

“subclinical” linguistic and cognitive deficits interact with the motoric weaknesses.

In my opinion, it is very likely that language and motor have an intricate relationship in terms of speech production. For instance, a young child who exhibits difficulty with speech sound production due to motor-based deficits may eventually persist with the speech sound production errors as a result of eventual language deficits. That is, the motor deficits may have “snow-balled” into language deficits after repeated incorrect production of meaningful linguistic units. Over time, those incorrect productions may result in incorrect phonological representations—this causes difficulties with language and literacy-based skills. Certainly, this particular scenario needs empirical support. However, from my perspective, this seems to be a logical and plausible explanation of the relation between language and motor for children with speech sound disorders.

Collectively, research supports that children with SSDs perform below their typically developing peers on measures requiring linguistic and motoric output. Thus, it is possible that the contributions of motor and language to speech production are not disparate, but are dynamically complementary (see Nip et al., 2009; Iverson, 2010, for reviews). To date, there is not one comprehensive investigation of both of these constructs within the same population of children with SSDs. Such a study would substantiate the relationship between language and motor and potentially ascertain the direction of said relationship.

In conclusion, it is evident that future work is necessary to better conceptualize the underlying mechanisms related to speech sound disorders—theoretical or otherwise. Such work will help to improve the quality of both assessment and treatment of this population of children. Further, it is hoped that this line of work will provide policy-makers and administrators with the evidence necessary to make appropriate decisions regarding service provision. It is unjust to regard any form of communication impairment as “just” a deficit that a child should deal with for life.

ACKNOWLEDGMENTS

The author is grateful to Dr. Sherine Tambyraja and Beau Bevens who both provided valuable feedback to original versions of this manuscript.

REFERENCES

- Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. L., Williams, J. M., and Zhang, Z. (2011). What factors place children with speech sound disorders at risk for reading problems? *Am. J. Speech Lang. Pathol.* 20, 146–160. doi: 10.1044/1058-0360(2011/10-0053)
- Farquharson, K. (2012). *Working Memory Processes in Children with and Without Persistent Speech Sound Disorders*. Unpublished doctoral dissertation, University of Nebraska-Lincoln, Lincoln, NE.
- Farquharson, K. (2015). After dismissal: examining the language, literacy, and cognitive skills of children with remediated speech sound disorders. *Perspect. Sch. Based Issues* 6, 50–59. doi: 10.1044/sbi16.2.50
- Felsenfeld, S., Broen, P. A., and McGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: educational and occupational results. *J. Speech Hear. Res.* 37, 1341–1353. doi: 10.1044/jshr.3706.1341
- Iverson, J. M. (2010). Developing language in a developing body: the relationship between motor development and language development. *J. Child Lang.* 37, 229–261. doi: 10.1017/S0305000909990432
- Larriee, L., and Catts, H. W. (1999). Early reading achievement in children with expressive phonological disorders. *Am. Speech Lang. Hear. Assoc.* 8, 118–128. doi: 10.1044/1058-0360.0802.118
- Lewis, B. A., Freebairn, L. A., Hansen, A. J., Stein, C. M., Shriberg, L. D., Iyengar, S. K., et al. (2006). Dimensions of early speech sound disorders: a factor analytic study. *J. Commun. Disord.* 39, 139–157. doi: 10.1016/j.jcomdis.2005.11.003
- Newmeyer, A. J., Grether, S., and Grasha, C. (2007). Fine motor function and oral-motor skills in preschool-age children with speech sound disorders. *Clin. Pediatr.* 46, 604–611. doi: 10.1177/0009922807299545

- Nip, I. S., Green, J. R., and Marx, D. B. (2009). Early speech motor development: cognitive and linguistic considerations. *J. Commun. Disord.* 42, 286–298. doi: 10.1016/j.jcomdis.2009.03.008
- Peter, B., and Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clin. Linguist. Phon.* 22, 171–198. doi: 10.1080/02699200701799825
- Preston, J. L., and Edwards, M. L. (2007). Phonological processing skills of adolescents with residual speech sound errors. *Lang. Speech Hear. Serv. Sch.* 38, 297–308. doi: 10.1044/0161-1461(2007/032)
- Raitano, N. A., Pennington, B. F., Tunick, R. A., Boada, R., and Shriberg, L. D. (2004). Pre-literacy skills of subgroups of children with speech sound disorders. *J. Child Psychol. Psychiatry* 45, 821–835. doi: 10.1111/j.1469-7610.2004.00275.x
- Redle, E., Vannest, J., Maloney, T., Tsevat, R. K., Eikenberry, S., Lewis, B., et al. (2015). Functional MRI evidence for fine motor praxis dysfunction in children with persistent speech disorders. *Brain Res.* 15, 47–56. doi: 10.1016/j.brainres.2014.11.047
- Sutherland, D., and Gillon, G. (2005). Assessment of phonological representations in children with speech impairment. *Lang. Speech Hear. Serv. Sch.* 36, 294–307. doi: 10.1044/0161-1461(2005/030)
- Swan, D., and Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *J. Exp. Child Psychol.* 66, 18–41. doi: 10.1006/jecp.1997.2375
- Visscher, C., Houwen, S., Moolenaar, B., Lyons, J., Scherder, E. J., and Hartman, E. (2010). Motor proficiency of 6- to 9-year-old children with speech and language problems. *Dev. Med. Child Neurol.* 52, e254–e258. doi: 10.1111/j.1469-8749.2010.03774.x
- Webster, R. I., Majnemer, A., Platt, R. W., and Shevell, M. I. (2005). Motor function at school age in children with a preschool diagnosis of developmental language impairment. *J. Pediatr.* 146, 80–85. doi: 10.1016/j.jpeds.2004.09.005

Conflict of Interest Statement: 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.

Copyright © 2015 Farquharson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Visualizing Similarity of Appearance by Arrangement of Cards

Nao Nakatsuji¹, Hisayasu Ihara², Takeharu Seno^{2,3} and Hiroshi Ito^{2*}

¹ Graduate School of Design, Kyushu University, Fukuoka, Japan, ² Faculty of Design, Kyushu University, Fukuoka, Japan,

³ Institute for Advanced Study, Kyushu University, Fukuoka, Japan

OPEN ACCESS

Edited by:

Paul M. W. Hackett,
University of Oxford, UK

Reviewed by:

Nikita Kuznetsov,
University of North Carolina
Greensboro, USA
Alessio Plebe,
University of Messina, Italy

*Correspondence:

Hiroshi Ito
hito@design.kyushu-u.ac.jp

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 28 November 2015

Accepted: 26 April 2016

Published: 13 May 2016

Citation:

Nakatsuji N, Ihara H, Seno T and Ito H
(2016) Visualizing Similarity of
Appearance by Arrangement of
Cards. *Front. Psychol.* 7:698.
doi: 10.3389/fpsyg.2016.00698

This study proposes a novel method to extract the configuration of the psychological space by directly measuring subjects' similarity rating without computational work. Although multidimensional scaling (MDS) is well-known as a conventional method for extracting the psychological space, the method requires many pairwise evaluations. The times taken for evaluations increase in proportion to the square of the number of objects in MDS. The proposed method asks subjects to arrange cards on a poster sheet according to the degree of similarity of the objects. To compare the performance of the proposed method with the conventional one, we developed similarity maps of typefaces through the proposed method and through non-metric MDS. We calculated the trace correlation coefficient among all combinations of the configuration for both methods to evaluate the degree of similarity in the obtained configurations. The threshold value of trace correlation coefficient for statistically discriminating similar configuration was decided based on random data. The ratio of the trace correlation coefficient exceeding the threshold value was 62.0% so that the configurations of the typefaces obtained by the proposed method closely resembled those obtained by non-metric MDS. The required duration for the proposed method was approximately one third of the non-metric MDS's duration. In addition, all distances between objects in all the data for both methods were calculated. The frequency for the short distance in the proposed method was lower than that of the non-metric MDS so that a relatively small difference was likely to be emphasized among objects in the configuration by the proposed method. The card arrangement method we here propose, thus serves as a easier and time-saving tool to obtain psychological structures in the fields related to similarity of appearance.

Keywords: similarity, appearance, multidimensional scaling, typeface, visualization

INTRODUCTION

Studies in psychology and cognitive science have attempted to structure the similarities among objects. Visualization of psychological structure has been intensively studied in the research field. The methods for visualization have been applied in practical-oriented study. For instance, Holleran (1992) visualized the relationship of similarity among 52 fonts as a map to extract common factors of fonts that people preferred. Chen (2009) also established a visualized map concerning 74 registered design patents for cars for the purpose of planning design strategies.

The geometrical approach has been employed to visualize the psychological structure. The geometric approach assumes that similarity or dissimilarity between the objects corresponds to the metric distances between the objects. Multidimensional scaling (MDS) is one of the geometric model approaches and has been commonly used for analyzing data, testing structural hypotheses, and exploring psychological structure in various field (e.g., Guttman, 1968; Borg and Groenen, 2005). Similarities between a pair of objects are transformed into a distance in a certain low-dimensional space and the objects are mapped into the space so as to satisfy each object's distance as best as possible. In general, the number of space dimensions is to be determined based on the fitting value, *Stress*, which expresses the errors between the similarity data and the distance (Kruskal, 1964a). The dimension can be selected so as to decrease the stress value.

This graphical display by MDS enables us to visually understand data structure even if there is no strong hypothesis that predicts patterns of data. For instance of analyzing data, Borg and Groenen (2005) provided a two-dimensional MDS representation regarding the correlations of crime rates over 50 U.S. states. Although it was difficult to understand the relationship behind the crime data without the MDS, it turned out that the crime data could be categorized by several items. That is, the horizontal axis and the vertical axis of the MDS representation can be interpreted as "violence vs. property" crimes and "hidden vs. street" crimes, respectively. The meaning of these axes is not obtained by MDS so we need to find rules of interpretation for describing MDS configurations by using additional knowledge.

Another object of MDS is testing structural hypotheses. For example, Levy (1983) categorized 18 types of attitudes toward political protest acts and confirmed that experimental results can reflect this organizational principle by using MDS. In this case, a three-dimensional MDS configuration was needed to explain the organizational principle appropriately.

MDS also enables us to discover the psychological structure that underlies similarity judgment. For instance, Wish (1971) collected similarity data among 12 nations from subjects and obtained the two-dimensional MDS configuration. The first axis of the two-dimensional map was interpreted as "pro-Western and pro-Communist" and the second axis as "economically developed and under-developed."

When we explore the psychological structure of a subject by MDS, the direct method is often used for collecting similarity data. Asking subjects to evaluate a numerical value for each pair of objects, such as for the 9-point scale, is one of the ways to collect similarity data directly. Using numerical value as similarity data may cause possible problems in the psychological field, because the value of the subjects' psychological distance is not interval scale but ordinal scale. To resolve this problem, non-metric MDS has been introduced to use rank orders for similarity among objects to construct the object's configuration (e.g., Kruskal, 1964a,b). There are several procedures to collect order of similarity (Borg and Groenen, 2005). One is to ask subjects to sort cards from the highest similarity pair to the lowest one. Another method is asking subjects to classify the pairs of objects into two groups according to similarity. The pairs in

each group are again classified into two groups in the same way. This procedure is repetitively performed until the subject thinks that it is no longer possible to find any differences in similarity among any pairs in the group. These methods are sometimes too time-consuming and demanding.

Although MDS is a useful tool as noted above, there are some problems in MDS techniques for exploring psychological structure within the realms of experimentation. Firstly, the subjects cannot assign meaning to MDS space. This makes it harder to understand one's psychological structure by translating the meaning of axes into a language, when there is no hypothesis or previous knowledge.

Secondly, when we try to obtain all combinations of similarity data between several objects, the subjects need to make a decision $n(n-1)/2$ times, where n is the number of the objects. If we take the subjects' workload into consideration, the number of objects should be limited to a relatively small number.

Thirdly, when the number of objects increases, it becomes difficult to find a solution for meeting the relationship between each similarity data and each distance on MDS space. In other words, the higher the number of objects is, the higher the stress value is (Spence and Ogilvie, 1973).

This study proposes a novel method to obtain the configuration of the psychological space of similarity data by experimental results without computational work. The proposed method asks subjects to arrange cards with stimuli on a poster sheet according to the degree of similarity between the cards, hereafter referred to as "card arrangement method."

We expected that the configurations of MDS and the card arrangement method would be equivalent, because the configuration obtained by the card arrangement method should be reflected by the subjects' psychological space in the same way as by MDS. We also expected that this proposed method would enable us to obtain the configuration of psychological space in a shorter time than MDS, because subjects can arrange the cards while seeing, comparing, and moving all of the cards simultaneously.

In this study, we applied the card arrangement method and non-metric MDS, hereafter referred to as "nMDS," to classify typefaces in order to clarify the common points and differences between the two configurations obtained by these methods and the advantage of the card arrangement method over MDS, including nMDS. We confirmed that each configuration of the card arrangement was approximately the same as that of nMDS. We found that the subjects exaggerated small differences between objects in the card arrangement method, which became a point of difference between the two methods. We also found that the experiment time of the card arrangement method was shorter than that of nMDS.

MATERIALS AND METHODS

Ethics Statement

Our experiments were preapproved by the Ethics Committee of Kyushu University and informed written consent was obtained from each subject prior to testing.

Subjects

Twenty volunteers participated in this experiment. They were divided into two groups comprising of ten subjects for each group, which were involved in two different experiments; namely, two conditions were employed according to a between-subjects design. The subjects were undergraduate students without expertise in the field of typeface design. They consisted of 12 males and 8 females, and their mean age was 21.05 (Min = 19, Max = 22, and $SD = 0.86$). All subjects were unaware of the exact purpose of this experiment and had normal or corrected to normal vision.

Stimuli

We chose 10 typefaces from among the roman type, the sans-serif type, and the slab-serif type, which can be easily obtained because they are installed in Mac OS X by default. (i) Garamond, (ii) Baskerville, (iii) Bell MT, and (iv) Didot fall into the roman type; (v) Futura, (vi) Gill Sans, (vii) Helvetica, and (viii) Optima fall into the sans-serif type; and (ix) Rockwell and (x) Playbill fall into the slab-serif type.

For the card arrangement method, we prepared 10 cards that each contained all lowercase and uppercase letters in alphabetical order (from “a” to “z” and from “A” to “Z”) and Arabic numerals (from “0” to “9”), as shown in **Figure 1B**. The cards were produced by PowerPoint (Microsoft). The colors of the cards and characters were white and black, respectively. The cards were rectangular in shape, measuring 5 cm in length by 15 cm in width, and each card weighed less than 1 g. The typeface size of the letters was 22 points, except for Playbill. Since height-to-width ratio of Playbill is much higher than the others, we assigned 40 points for Playbill to equalize the size of the typefaces in appearance. In order to have the subjects practice on the card arrangement method, six square-shaped cards were also prepared, measuring 15 cm in length by 15 cm in width. The six cards showed different shapes of a refrigerator; for example, one of the figures showed a two-door type refrigerator while another showed a three-door type.

For nMDS, we prepared 45 papers that each contained a different pair of typefaces. The papers were A-4 size and contained the same letters and numerals of the same size as those in the card arrangement method.

Apparatus

For the card arrangement method, we prepared a white poster sheet of 100 cm square, with a 5 cm black grid onto which subjects were to place the cards (**Figure 1A**). We also prepared a digital camera (Power Shot SX 40, Canon) and a tripod stand to record the placement of the cards on the poster sheet. Furthermore, a desk of a certain size was prepared onto which the poster, digital camera, and tripod stand were to be placed. For nMDS, we prepared a desk onto which 45 cards could be placed. Both experiments were conducted in different rooms simultaneously, and each room had a luminance of around 400 lux.

Procedure

For a practice on the card arrangement method, ten subjects were asked to place six cards with different refrigerators onto

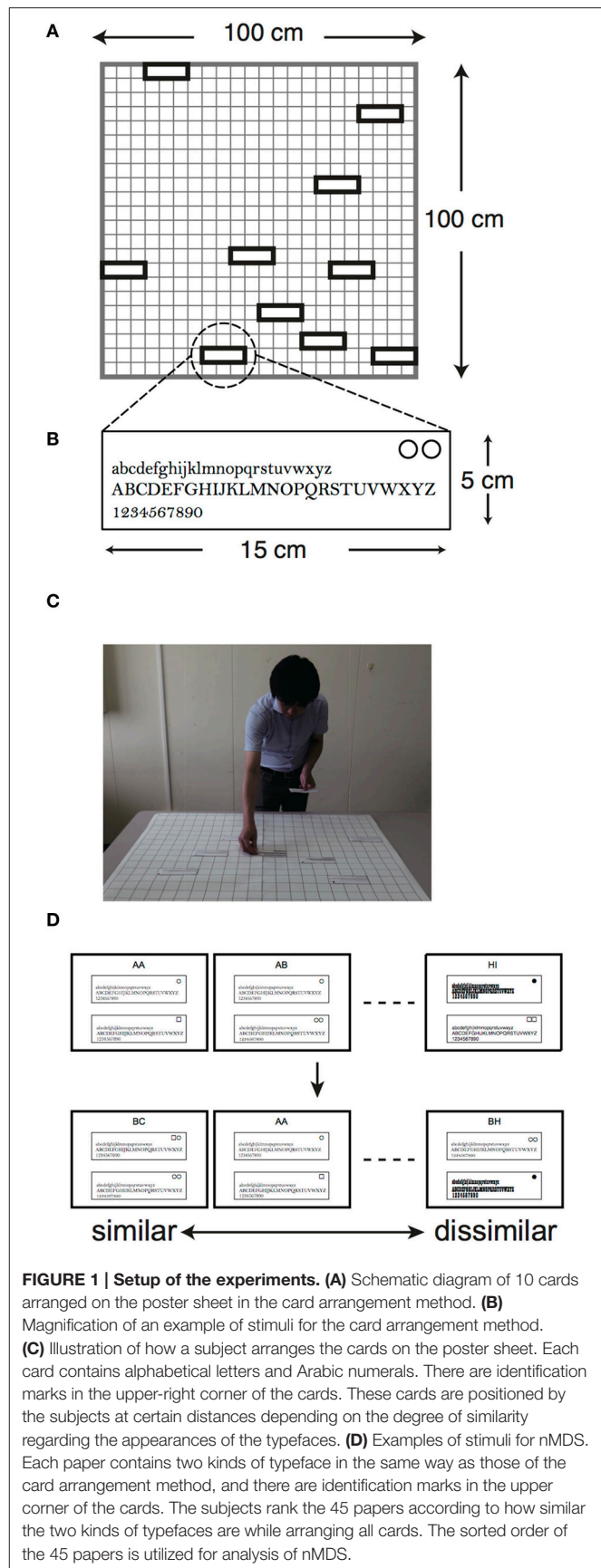


FIGURE 1 | Setup of the experiments. (A) Schematic diagram of 10 cards arranged on the poster sheet in the card arrangement method. **(B)** Magnification of an example of stimuli for the card arrangement method. **(C)** Illustration of how a subject arranges the cards on the poster sheet. Each card contains alphabetical letters and Arabic numerals. There are identification marks in the upper-right corner of the cards. These cards are positioned by the subjects at certain distances depending on the degree of similarity regarding the appearances of the typefaces. **(D)** Examples of stimuli for nMDS. Each paper contains two kinds of typeface in the same way as those of the card arrangement method, and there are identification marks in the upper corner of the cards. The subjects rank the 45 papers according to how similar the two kinds of typefaces are while arranging all cards. The sorted order of the 45 papers is utilized for analysis of nMDS.

the sheet. The subjects were given instructions on how to adjust the position of those cards with similar refrigerators to a smaller distance. We wrote down the sentences “Being similar is arranged close-by” and “Being dissimilar is arranged farther away” on the white board to avoid confusion for the subjects. The subjects were also asked to use the entire space of the sheet and to match the corners of the cards to the corners of the grid. Even though the subjects assigned meaning to two axes of a two-dimensional space once, they were allowed to change the meaning of these two axes during the experiment. After the practice was completed, the ten subjects were asked to place the 10 cards with different typefaces in the same way as described above (Figures 1A–C). After arrangement of the cards, we took photos of the placements of the cards, checked the identification mark in the upper-right corner of the cards to identify them, recorded coordinates of the cards, calculated the Euclidean distances between each card as arranged on the sheet, and then considered the obtained distances as similarity data.

For nMDS, the other ten subjects ranked the 45 papers containing pairs of typefaces in order of similarity (Figure 1D). We collected a matrix of ordinal similarity data from the results. The nMDS configuration was obtained from the matrix based on Sammon’s Non-Linear Mapping (Sammon, 1969).

Both experiments were carried out once for each of the subjects. Both the card placement in our proposed method and the evaluation of similarity in nMDS were continued until the subjects decided to finalize the process, so there was no time pressure in each condition. The subjects in both experiments were instructed to do the task as much as they like until they are fully satisfied.

The subjects evaluated their level of fatigue for the task and their satisfaction with the obtained configuration by providing ratings on two 5-point scales (1: not tired at all or not satisfied at all, 5: very tired or very satisfied). We also recorded the total time duration of the task.

Quantification of Similarity between Configurations

We used a trace correlation coefficient (Hooper, 1959) to compare the configurations of the card arrangements with nMDS. Trace correlation coefficient takes a value between 0 and 1. The closer to 1 the value is, the more similar the pattern between two configurations is. Consider the matrix of configuration $X = [x_1, x_2, \dots, x_n]$, where column vector x_i is the Cartesian coordinate produced by $[i]$ th stimuli. Let $X^{(j)}$ be the configuration obtained by the $[j]$ th subjects. The trace correlation coefficient between two of the configurations, $X^{(1)}$ and $X^{(2)}$, is defined as

$$\alpha(X^{(1)}, X^{(2)}) = \sqrt{\frac{\rho_1^2 + \rho_2^2}{2}},$$

where ρ_i is the $[i]$ th canonical correlation coefficient between $X^{(1)}$ and $X^{(2)}$. The value of trace correlation coefficient is invariant under the affine transformation of X . This property guarantees the same value of trace correlation coefficient can be obtained even if the configurations are rotated or parallel shifted. Thus

we can compare two configurations through trace correlation coefficient regardless of their directions and shifts.

Software

All statistical analyses and computations of the trace correlation coefficient were conducted using R version 3.2.1 (R Core Team, 2015) and R package “stats.”

RESULTS

Examples of Mapping

Figure 2 shows two examples of the configuration results of the card arrangement method and nMDS. As can be seen in Figure 2, Garamond, Baskerville, Bell MT, and Didot were placed at a short distance in both methods. It should be noted that the meaning of space in the two mapping methods is different. The axes shown in Figure 2B do not mean anything by the only principle of nMDS. They are supposed to be construed by additional knowledge such as serif type or non-serif type. In contrast, by using the card arrangement method, subjects can define the meaning of the axes. For instance, according to the subject whose result is shown in Figure 2A, the horizontal axis represents how bold or not bold the letters are, and the vertical axis represents how decorated or undecorated they are. As for all of the subjects, 90% of them responded that the first axis was how bold the letters are but the meaning of the second axis was not consistent among them.

Difference Level of Configuration in Space

In order to evaluate the similarity of the configuration patterns obtained by the two methods, we computed the trace correlation coefficient for each configuration pattern (see Material and Methods). We decided the threshold value of trace correlation coefficient for extracting similar pairs of configurations that is statistically significantly similar. The threshold value was defined based on random data as follows. The random configuration was generated by 10 pairs of random values following uniform distribution, $U(0, 1)$. Thereafter, we calculated the trace correlation coefficient between a pair of random configurations 10^5 times to obtain its distribution. The random trace correlation

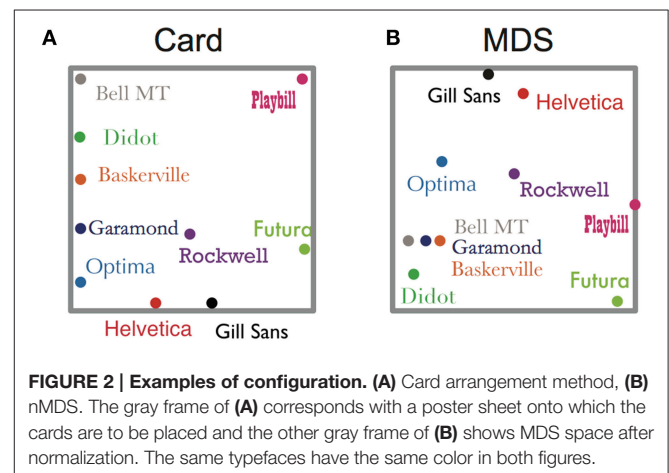


FIGURE 2 | Examples of configuration. (A) Card arrangement method, **(B)** nMDS. The gray frame of **(A)** corresponds with a poster sheet onto which the cards are to be placed and the other gray frame of **(B)** shows MDS space after normalization. The same typefaces have the same color in both figures.

coefficient that showed a 0.05 level of significance corresponded to the trace correlation coefficient of 0.658, suggesting that 5% pairs of configuration by random data can show the trace correlation coefficient of more than 0.658. Thus, we chose 0.658 as the threshold value for similarity judgment.

We calculated the trace correlation coefficient among all combinations of the configuration obtained from the card arrangement method and nMDS (**Figure 3A**). The ratio of the trace correlation coefficient exceeding the threshold was 62.0%. This result suggests that these methods can produce statistically significantly similar configurations.

We also examined variations of the subjects for each method. The trace correlation coefficient was calculated for all combinations of the configurations obtained by the proposed method and nMDS, respectively (**Figures 3B,C**). These distributions of the trace correlation coefficient have similar values for average and *SD*: Mean = 0.71, *SD* = 0.13 for the card arrangement method and Mean = 0.72, *SD* = 0.13 for nMDS, suggesting that both the methods have approximately the equivalent ability to express a psychological structure in similarity.

Moreover, we examined the average of distance between the placed cards on the sheet and extracted the top 10 pairs with a small distance in each method (**Figure 4A**). Eight pairs among

the top ten pairs are common in both methods. This also supports that there is a similarity between the configurations obtained from the card arrangement method and nMDS.

Distances among Objects Are Not Close in Configuration in the Card Arrangement Method

Although we indicated the similarity in both methods, there are some differences in detailed point. For example, the configurations shown in **Figure 2** had somewhat different features. Garamond, Didot, and Baskerville were located near each other in both the methods (**Figures 2A,B**), but these cards were arranged less closely in the card arrangement method (**Figure 2A**), compared with nMDS (**Figure 2B**). We assumed that subjects focused on the slight difference between objects and expressed their similarity judgment in a two-dimensional space, which may be the point of difference between the configurations of the two methods. In order to confirm this hypothesis, we calculated all distances between objects in all data (**Figures 4B,C**). We found the frequency of short distance to be lower in the card arrangement method compared with that in nMDS, which supports our hypothesis.

The average Euclidean distances between objects also support the hypothesis. In nMDS, Garamond, Didot, Baskerville, and Bell MT were placed close to each other (**Figure 4E**). However, in the card arrangement method, these typefaces were placed less closely (**Figure 4D**). Based on these results, we can conclude that card arrangement method tends to exaggerate the small difference between objects.

Experiment Time, Level of Fatigue, and Level of Satisfaction

The required duration of the card arrangement method was drastically shorter than that of nMDS, as we expected (**Figure 5**). The duration of the card arrangement method was approximately one third of nMDS's duration. Regarding the level of fatigue and satisfaction, no significant difference between the two methods was observed.

DISCUSSION

This study attempted to clarify common and different points between the configurations obtained by the card arrangement method and nMDS and the advantage of the proposed method. We found that there is a correlation between the configurations obtained from the proposed method and those obtained from nMDS. This result demonstrates the fact that the proposed method can be used instead of nMDS to obtain psychological structures. The required duration for the proposed method is considerably shorter than that required for nMDS, which indicates that the proposed method is a useful tool to save time. In particular, the proposed method can be utilized effectively in the psychological field, because evaluation of similarity must be obtained from subjects directly and their workload should be reduced. While this study asked subjects to sort papers from the highest similarity pair to the lowest one and collected ordinal

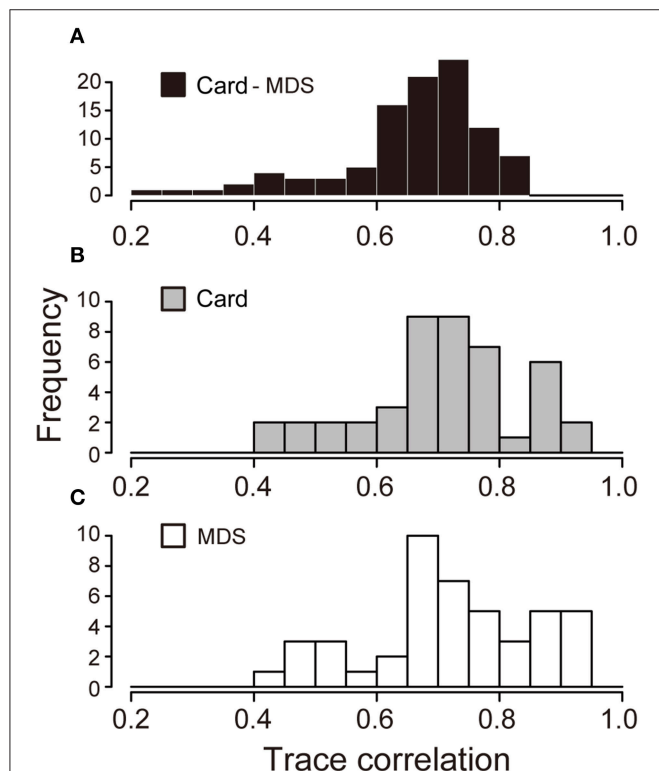


FIGURE 3 | Histogram of trace correlation coefficient among configurations. (A) The histogram of trace correlation coefficient for all combination of configurations obtained by the card arrangement method and nMDS. **(B,C)** The histogram of trace correlation coefficient among all pairs of configurations by the card arrangement method and nMDS, respectively.

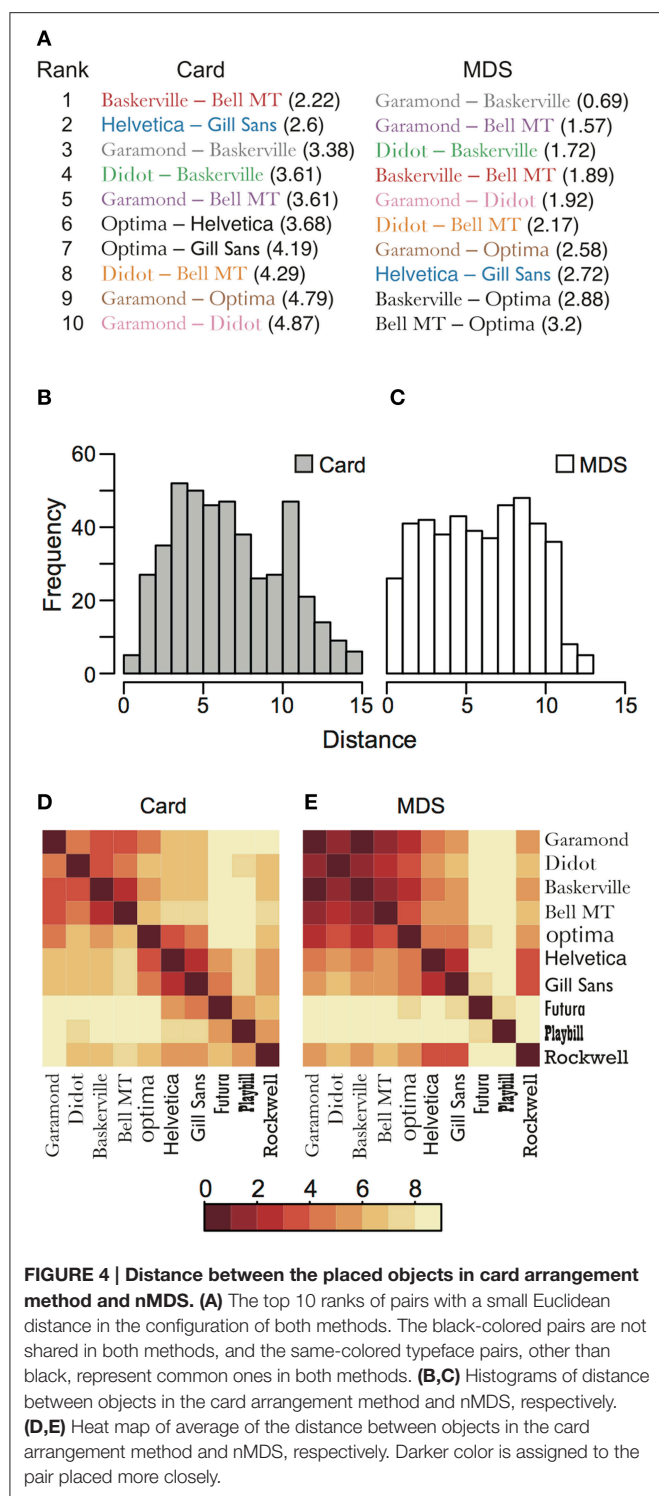
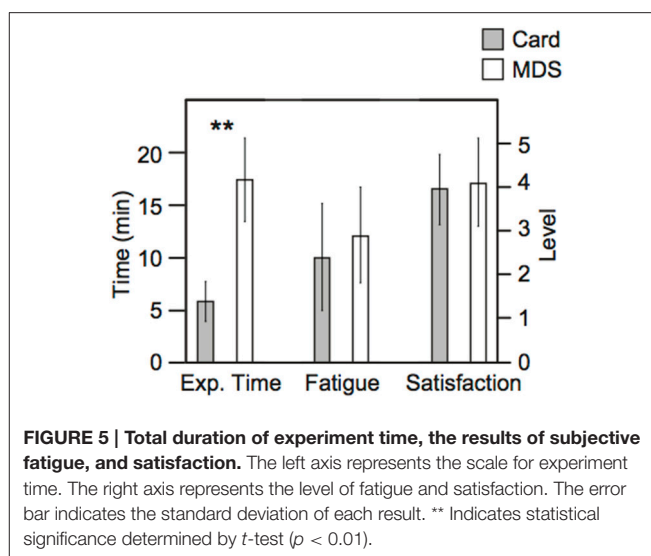


FIGURE 4 | Distance between the placed objects in card arrangement method and nMDS. (A) The top 10 ranks of pairs with a small Euclidean distance in the configuration of both methods. The black-colored pairs are not shared in both methods, and the same-colored typeface pairs, other than black, represent common ones in both methods. **(B,C)** Histograms of distance between objects in the card arrangement method and nMDS, respectively. **(D,E)** Heat map of average of the distance between objects in the card arrangement method and nMDS, respectively. Darker color is assigned to the pair placed more closely.

similarity data for nMDS, the required duration for the proposed method would be shorter than that of the other collecting method for ordinal similarity, such as forming groups corresponding to the degree of similarity. Moreover, the proposed method is highly simple and easy since it requires only cards and poster sheets while nMDS requires relatively complicated data analysis.



We understand that this proposed method could be applied to existing research related to similarity of appearance. For example, Holleran (1992) asked 50 subjects to rate the similarity between 52 typefaces and tried to map the subjects' psychological structure of similarity into two dimensions. If we apply the card arrangement method to these stimuli, we could obtain similarity data and the configuration of the stimuli in a markedly shorter time. In addition, Holleran (1992) did not specify the meaning of horizontal axis and vertical axis, but we could accomplish this by using the card arrangement method, because we could directly obtain the axes' meaning from the subjects.

The Management and Law fields also have a need for similarity of appearance analyses. Enterprises file applications and obtain design patent rights to prevent their appearance of products from imitation. They develop a design patent map from publication of the design patent not only to manage their own designs already filed but also to ensure positioning between their product design and their competitor's design. In the design patent map, each design should be arranged based on the degree of similarity among them. For example, Chen (2009) collected similarity data regarding 74 registered design patents for cars from industrial designers, established a design map by using MDS, and drafted the design map for the purpose of planning design strategies.

In addition, once the infringement lawsuit is filed, similarity of the registered design is becoming a material matter, because there is a possibility that they have to stop selling their products and pay damage. In the practice of major countries regarding design patent infringement, if the registered design and an accused design are the same or similar and give a common impression, it is judged that the accused design violates the rights of the registered design (e.g., Article 23 of the Japan Design Act¹; Egyptian Goddess v. Swisa, 2008²; Article 10 of Council Regulation (EC) No. 6/2002 of 12 December 2001 on Community Designs³). Also, when the accused design is compared to the

¹<http://www.japaneselawtranslation.go.jp/law/detail?id=44&vm=04&re=01>.

²<http://www.cafc.uscourts.gov/content/egyptian-goddess-v-swisa-inc>.

³<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32002R0006&rid=4>.

registered design, the prior art should be taken into account. For example, 10 prior designs were submitted to decide the essential part of the registered design in a certain Japanese infringement case (Hei-12 (Wa) 2240, 2001⁴). If MDS were to be applied to such a case, the experiment would take much time due to the large number of stimuli. The method proposed in this study could be applied to such a situation to save experiment time. Moreover, there are many cases where the registered design and the accused design appear very similar at first glance. In such a case, we can make good use of the proposed method, because subjects can focus on slight differences when they make similarity judgments according to this method.

The subjects can watch all the cards and evaluate similarity among them simultaneously in the proposed method, so that the results of the proposed method can be affected by context. It is known that the similarity of objects A and B are influenced by object C (Tversky, 1977). In the proposed method, subjects can see and compare all the objects while arranging cards so that all the objects can be considered when evaluating the similarity between objects A and B. On the other hand, regarding the pairwise comparison, subjects can see only A and B and evaluate the similarity of A and B so that C would not be taken into account. Thus, if we expect the results to be affected by the context effect, the proposed method should be applied. Otherwise, the pairwise comparison would be better.

The proposed method allows the subject to set the axis of the configuration pattern. In this case, the meaning of the axis can be revealed through the speech of the subjects. In contrast, there is a case where the subjects do not recognize the axis even after finishing the task. In this instance, someone other than the subjects can give meaning to the axis from a two-dimensional configuration. MDS enables us to find a principle dimension even if the subject does not note the meaning of the dimension, which is one of MDS' benefits. In that sense, the merit of MDS is shared with the proposed method.

⁴http://www.courts.go.jp/app/hanrei_jp/detail7?id=12556.

REFERENCES

- Borg, I., and Groenen, P. J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications*, 2nd Edn. New York, NY: Springer.
- Chen, R. (2009). Design patent map visualization display. *Expert Syst. Appl.* 36, 12362–12374. doi: 10.1016/j.eswa.2009.04.049
- Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika* 33, 469–506. doi: 10.1007/BF02290164
- Holleran, P. A. (1992). "An assessment of font preferences for screen-based text display," in *People and Computers VII*, eds A. F. Monk, D. Diaper, and M. D. Harrison (Cambridge: Cambridge University Press), 447–461.
- Hooper, J. W. (1959). Simultaneous equations and canonical correlation theory. *Econometrica* 27, 245–256. doi: 10.2307/1909445
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 9, 1–27. doi: 10.1007/BF02289565
- Kruskal, J. B. (1964b). Nonmetric multidimensional scaling. *Psychometrika* 29, 115–129. doi: 10.1007/BF02289694
- Levy, S. (1983). A cross-cultural analysis of the structure and levels of attitudes towards acts of political protest. *Soc. Indic. Res.* 12, 281–309. doi: 10.1007/BF.00319806

The proposed method is confined to a two-dimensional space. The two-dimensional space is the easiest dimension to be interpreted and applied to actual issues. Although compressing psychological structure into a two-dimensional space can reduce the amount of information that the original psychological structure preserves, one can estimate the degree of reflection for cognitive space by asking the subjects about the "degree of satisfaction of configuration." The extension of our model to a three-dimensional space may be possible, when the two-dimensional space is not large enough to reflect one's cognitive structure. The development of a software that allows us to place objects in a three-dimensional space and validating of the effectiveness of such a space may be explored in future works.

The proposed method is effective for visual stimuli because subjects can watch all cards and evaluate them simultaneously. Moreover, this method may be applied to conceptual stimuli, such as similarity between nations. We expect that this card arrangement method has versatility and that it will be possible to apply it to various areas. The application to other fields should be explored in future studies.

AUTHOR CONTRIBUTIONS

NN, Hihara, and Hito conceived and designed the study; NN and Hito performed the experiments; NN and Hito analyzed the data; and NN, TS, and Hito wrote the paper.

ACKNOWLEDGMENTS

We thank S. Fujimoto, M. Hamada, N. Okamoto, and Y. Uchida (Kyushu Univ.) for their help with the experiments and T. Aso (Kyushu Univ.) for the fruitful discussions. This research was supported by Grants-in-Aid for Scientific Research (26282006 to Hito).

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Sammon, J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE Trans. Comput.* 18, 401–409. doi: 10.1109/T-C.1969.222678
- Spence, I., and Ogilvie, J. C. (1973). A table of expected stress values for random rankings in nonmetric multidimensional scaling. *Multivar. Behav. Res.* 8, 511–517.
- Tversky, A. (1977). Features of Similarity. *Psychol. Rev.* 84, 327–352. doi: 10.1037/0033-295X.84.4.327
- Wish, M. (1971). "Individual differences in perceptions and preferences among nations," in *Attitude Research Reaches New Heights*, eds C. W. King and D. Tigert (Chicago, IL: American Marketing Association), 312–328.

Conflict of Interest Statement: 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.

Copyright © 2016 Nakatsuji, Ihara, Seno and Ito. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Book Review: “Classifying Reality,” by David S. Oderberg (ed.) (2013)

Paul M. W. Hackett *

School of Communication, Emerson College, Boston, MA, USA

Keywords: classification, categories, mapping sentence, facet theory, mereology, structural ontology, psychological processes, book reviews as topic

A book review on Classifying Reality

Edited by D. S. Oderberg, Chichester: Wiley-Blackwell, 2013.

OPEN ACCESS

Edited and reviewed by:

Tuomas K. Pernu,
King's College London, UK and
University of Helsinki, Finland

***Correspondence:**

Paul M. W. Hackett,
paul.hackett@gmail.com

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 26 February 2015

Accepted: 30 March 2015

Published: 16 April 2015

Citation:

Hackett PMW (2015) Book Review:
“Classifying Reality,” by David S.
Oderberg (ed.) (2013).
Front. Psychol. 6:461.
doi: 10.3389/fpsyg.2015.00461

Classifying reality is an ancient and fundamental ability. Even the most basic living creatures classify their world as do social collectives of living things, most noticeably, human beings through psychological processes. An example of the rudimentary nature of classification exists in the behavior of slime molds. When food is abundant many of the more than 900 types of slime molds exist not as slime but as single-cell organisms. However, if food is in short supply the cells agglomerate and move as a single body. As a collective, slime molds are able to identify food and experiments have demonstrated that through changing shape the mold can reach food in a maze: Branches of the mold that do not terminate at a food source “die off” resulting in an almost singular path to food. Remarkably, these single cell organisms can act as a collective and classify correct or incorrect, advantageous or non-advantageous turns in the maze. At this rudimentary behavioral level slime mold is able to classify reality in order to bring advantage to the collective. This example demonstrates the fundamental nature of classification as a behavioral and biological process. Within psychology classifying has featured large within many areas of the literature but perhaps mainly within developmental and cognitive sub-disciplines.

Notwithstanding the seminal nature of the process of classifying in human life, contemporary books published on the subject of classification are uncommon and I believe that psychologists could benefit from considering the broad perspective assumed in Oderberg's book. His edited volume is a concise collection of writing by contemporary scholars in which each of the six chapters is concerned with an aspect of identifying the structure of reality. Realism proposes that some entities do not need conceptual systems, beliefs, or our linguistic practices, etc., and may be thought of as objectively real. This contrasts with conventionalism, which posits entities are constituted on social agreement. The book contains an eclectic assemblage of writing from authors drawn from a realist perspective as, “... realism about classification stands its ground: all major lines of criticism available to the extreme conventionalist can be addressed.” Tahko (2013, p. 60).

The central question posed in Oderberg's book may be summarized as: Is it possible to classify reality? Here he makes the not universally accepted point that in order to be able to classify reality it is necessary that we establish that reality has clearly extant boundaries to its content. Oderberg's collection does not specifically concentrate upon categories per-se, but rather upon classification, where amongst other questions he asks of the extent to which classification is a fabrication of the mind? However, due to limited space and the title of this special edition I will emphasize categorization. Essays in the first section of the book each chapter reviews

somewhat abstract questions and issues associated with classification within a realist approach. The authors question any general or universal framework for classifying being or the development of a general ontology. Oderberg presents theoretical notions of classification as presented in the late E. J. Lowe's opening chapter on categorical predication. This chapter offers a synopsis and extension of Lowe's thinking on his four-category ontology (Lowe, 2006). Lowe asks how many components would be needed to adequately define a basic categorization of our world and proposes this to be four where any elements within such a system must be mutually exclusive in what they define and that their content cannot be classified through the combination of other elements in the ontology. In a revision of formal logic, Lowe presents a complex categorical ontology of the kinds of things that form the fundamental components of our world. Tahko, in his chapter, questions the conventional notion that there is no best or more realistic form of classification of the fundamental nature of the world. In section two the authors put forth their thoughts on the use of objective classification in science in general and specifically in biology. Authors develop a more applied understanding of classification in chapters such as Stephen Boulter's that posit thoughts upon the classification of biological forms and kinds and in Rosenkrantz's writing that forwards the notion that there are necessary and essential aspects of living things.

In the above review I have demonstrated the breadth and depth of the chapters' contents and I will now evaluate Oderberg's

text as this may be applicable to psychologists. Whilst the book offers a diverse range of thinking on the process of classification there is no concluding chapter on contemporary thinking on classification and its further development and usage. Furthermore, the authors pay little heed to the relationship between classification units (mereology—the study of part-to-part and part-to-whole relationships) either as independent component of classification or as componential sub-units. My work in this area using the mapping sentence as a mereological framework (see, Hackett, 2014) is pertinent in this regard where I suggest that not all classification forms (or categories) are, or should be thought of, as equal. The mapping sentence I put forth as a flexible framework for the incorporation of the combined effects of inter-related or non-independent categories or classifications in a meaningful manner.

It is appropriate to review this book in this special edition on the psychological aspects of categorization as psychological approaches associated with categorization may be identified as a special case of classification. Indeed, the title of this book *Classifying Reality* could have appropriately been the title of this special edition. The psychology study of classification and categories as a separate area of research is neglected and lacks a clear theoretical conception of categorizing processes. Oderberg's volume offers psychologists, and others, an insightful starting point from which to develop research into classification especially when this is read and employed in conjunction with the psychological approach of the mapping sentence.

References

- Hackett, P. M. W. (2014). *Facet Theory and the Mapping Sentence: Evolving Philosophy, Use and Application*. Basingstoke: Palgrave.
- Lowe, E. J. (2006). *The Four Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press.
- Tahko, T. E. (2013). "Boundaries in reality," in *Classifying Reality*, ed D. S. Oderberg (Chichester: Wiley Blackwell), 41–60.

Conflict of Interest Statement: 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.

Copyright © 2015 Hackett. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership