

# BEYOND THE BODY? THE FUTURE OF EMBODIED COGNITION

EDITED BY : Guy Dove

PUBLISHED IN : Frontiers in Psychology



# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2016 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 978-2-88919-797-2

DOI 10.3389/978-2-88919-797-2

## About Frontiers

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

## Frontiers Journal Series

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

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

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

# BEYOND THE BODY? THE FUTURE OF EMBODIED COGNITION

Topic Editor:

**Guy Dove**, University of Louisville, USA

Embodied cognition represents one of most important research programs in contemporary cognitive science. Although there is a diversity of opinion concerning the nature of embodiment, the core idea is that cognitive processes are influenced by body morphology, emotions, and sensorimotor systems. This idea is supported by an ever increasing collection of empirical studies that fall into two broad classes: one consisting of experiments that implicate action, emotion, and perception systems in seemingly abstract cognitive tasks and the other consisting of experiments that demonstrate the contribution of bodily interaction with the external environment to the performance of such tasks.

Now that the research program of embodied cognition is well established, the time seems right for assessing its further promise and potential limitations. This research topic aims to create an interdisciplinary forum for discussing where we go from here. Given that we have good reason to think that the body influences cognition in surprisingly robust ways, the central question is no longer whether or not any cognitive processes are embodied. Instead, other questions have come to the fore: To what extent are cognitive processes in general embodied? Are there disembodied processes? Among those that are embodied, how are they embodied? Is there more than one kind of embodiment? Is embodiment a matter of degree?

There are a number of specific issues that could be addressed by submissions to this research topic. Some supporters of embodied cognition eschew representations. Should anti-representationalism be a core part of an embodied approach? What role should dynamical models play? Research in embodied cognition has tended to focus on the importance of sensorimotor areas for cognition. What are the functions of multimodal or amodal brain areas? Abstract concepts have proved to be a challenge for embodied cognition. How should they be handled? Should researchers allow for some form of weak embodiment? Currently, there is a split between those who offer a simulation-based approach to embodiment and those who offer an enactive approach. Who is right? Should there be a rapprochement between these two groups? Some experimental and robotics researchers have recently shown a great deal of interest in the idea that external resources such as language can serve as form of cognitive scaffolding. What are the implications of this idea for embodied cognition?

This research topic aims to bring together empirical and theoretical work from a diversity of perspectives. Submissions are sought from any of the major disciplines associated with cognitive science, including but not necessarily limited to anthropology, cognitive psychology, computational modeling, linguistics, neuroscience, philosophy, robotics, and social psychology. Researchers are encouraged to submit papers discussing experiments, methods, models, or theories that speak to the issue of the future of embodied cognition.

**Citation:** Dove, G., ed. (2016). *Beyond the body? The Future of Embodied Cognition*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-797-2

# Table of Contents

- 05    *How to go beyond the body: an introduction***  
Guy Dove
- 08    *The co-constitution of the self and the world: action and proprioceptive coupling***  
Olivier Gapenne
- 15    *Mapping the feel of the arm with the sight of the object: on the embodied origins of infant reaching***  
Daniela Corbetta, Sabrina L. Thurman, Rebecca F. Wiener, Yu Guan, and Joshua L. Williams
- 33    *Developing embodied cognition: insights from children's concepts and language processing***  
Michele Wellsby and Penny M. Pexman
- 43    *A perceptual account of symbolic reasoning***  
David Landy, Colin Allen and Carlos Zednik
- 53    *fNIRS in motion—unraveling the neurocognitive underpinnings of embodied numerical cognition***  
Julia Bahnmüller, Thomas Dresler, Ann-Christine Ehlis, Ulrike Cress and Hans-Christoph Nuerk
- 57    *The specificity of action knowledge in sensory and motor systems***  
Christine E. Watson, Eileen R. Cardillo, Bianca Bromberger and Anjan Chatterjee
- 68    *Clustering, hierarchical organization, and the topography of abstract and concrete nouns***  
Joshua Troche, Sebastian Crutch and Jamie Reilly
- 78    *Toward a more embedded/extended perspective on the cognitive function of gestures***  
Wim T. J. L. Pouw, Jacqueline A. de Nooijer, Tamara van Gog, Rolf A. Zwaan and Fred Paas
- 92    *The body and the fading away of abstract concepts and words: a sign language analysis***  
Anna M. Borghi, Olga Capirci, Gabriele Gianfreda and Virginia Volterra
- 105    *Sensory motor mechanisms unify psychology: the embodiment of culture***  
Tamer Soliman, Alison Gibson and Arthur M. Glenberg
- 114    *Action scaling of distance perception is task specific and does not predict "the embodiment of culture": a comment on Soliman, Gibson, and Glenberg (2013)***  
Andrew D. Wilson
- 116    *How intent to interact can affect action scaling of distance: reply to Wilson***  
Tamer M. Soliman and Arthur M. Glenberg



**118 *Linguistic embodiment and verbal constraints: human cognition and the scales of time***

Stephen J. Cowley

**129 *Embodied niche construction in the hominin lineage: semiotic structure and sustained attention in human embodied cognition***

Aaron J. Stutz

# How to go beyond the body: an introduction

Guy Dove \*

*Department of Philosophy, University of Louisville, Louisville, KY, USA*

**Keywords: embodied cognition, grounded cognition, extended cognition, perception, action, concepts**

Embodied cognition represents one of most important theoretical developments in contemporary cognitive science. Many cognitive processes appear to be influenced by body morphology, emotions, and sensorimotor systems. This perspective is supported by an ever increasing collection of empirical studies that fall into two broad classes: one consisting of experiments that implicate action, emotion, and perception systems in seemingly abstract cognitive tasks and the other consisting of experiments that demonstrate the contribution of bodily interaction with the external environment to the performance of such tasks.

Now that embodied cognition is fairly well established, the time seems right for assessing its further promise and potential limitations. This research topic aimed to create an interdisciplinary forum for discussing where we go from here. Given that we have good reason to think that the body influences cognition in surprisingly robust ways, the central question is no longer whether or not some cognitive processes are embodied. Other questions have come to the forefront. To what extent are cognitive processes embodied? Are there disembodied processes? Among those that are embodied, how are they embodied? Is there more than one kind of embodiment? Is embodiment a matter of degree?

## Extending the Research Program

Many of the contributions to this research topic involve experiments that extend the empirical reach of embodied cognition. For instance, Soliman et al. (2013) ambitiously propose that sensorimotor mechanisms can unify explanations at cognitive, social, and cultural levels. They carried out two experiments investigating whether anticipated motor effort can be used to understand cultural differences. Building on earlier work by Proffitt and colleagues implicating an effect of perceived motor effort on visual distance perception (for a review see Proffitt and Linkenauger, 2013), they investigate a cultural motor-effort hypothesis in which relative degree of experience with out-group members can lead to differences in perceived distance. In a commentary, Wilson (2014) suggests that this effect conflicts with the task-relatedness of the effects found by Proffitt and colleagues. Soliman and Glenberg (2014) respond by clarifying how they link their cultural-motor effort hypothesis to the earlier work. Ultimately, further research is needed to settle these issues.

Much of the extant research on concepts within an embodied framework focuses on the binary question of whether or not they are embodied as a general rule. Recently, researchers have come to realize that embodiment might be context-dependent and come in degrees (e.g., Watson and Chatterjee, 2011; Pulvermüller and Garagnani, 2014; Zwaan, 2014). With this potential flexibility in mind, Watson et al. (2014) examined the sensorimotor specificity of action concepts elicited by different exemplars and representational formats. They found that actions appear to be represented at different levels of specificity by visual and motor systems and that the relative recruitment of some sensorimotor brain regions may depend on the format of the stimuli.

Abstract concepts remain a serious challenge for embodied cognition (Dove, 2015). A couple of the contributions address aspects of this challenge. Troche et al. (2014) defend a multidimensional

## OPEN ACCESS

### Edited and reviewed by:

Eddy J. Davelaar,  
Birkbeck, University of London, UK

### \*Correspondence:

Guy Dove,  
guy.dove@louisville.edu

### Specialty section:

This article was submitted to  
Cognitive Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 15 April 2015

**Accepted:** 05 May 2015

**Published:** 21 May 2015

### Citation:

Dove G (2015) How to go beyond the  
body: an introduction.  
Front. Psychol. 6:660.  
doi: 10.3389/fpsyg.2015.00660

approach to abstract concepts. Rather than rely on an intuitive notion of abstractness, they investigated how the meanings of 400 concrete and abstract English nouns are distributed in a multidimensional space using hierarchical cluster analysis. Participants rated the nouns along 12 dimensions. Factor reduction yielded three latent factors that the authors characterize as affective association, perceptual salience, and magnitude. When the original words were plotted for these three factors, abstract and concrete words were associated with unique, but somewhat overlapping, topographies within this space. Borghi et al. (2014) analyze how Italian Sign Language (LIS, *Lingua dei Segni Italiana*) encodes abstract concepts. They argue that the LIS data support the view that abstract concepts are encoded in multiple ways. Some abstract concepts may rely more on metaphors while others may rely more on situations, emotions, or linguistic information.

Despite the clear affinity between constructivist views of cognitive development and embodied cognition, the precise role that embodiment may play in development remains an open question. Corbetta et al. (2014) provide evidence suggesting that the emergence of reaching is a fundamentally embodied process. Infants appear to first learn to make such movements through the haptic and proprioceptive feedback associated with self-produced movements. Vision then maps onto this motor experience and contributes to the emergence of prospective motor control.

Although it is not always acknowledged, the conceptual re-framing of cognition as an embodied activity has important implications with respect to methodology. Bahnmüller et al. (2014) contend that near infrared spectroscopy (NIRS) is better suited to investigating the role that motion plays in embodied cognition than the more commonly used functional magnetic resonance imaging (fMRI).

## New Directions

Several of the contributions are theoretical in nature. These echo many of the themes present in the experimental contributions but also expand the scope of embodied cognition. Some propose stronger versions of the embodiment thesis and others outline new frameworks for integrating embodied cognition with other disciplines.

Pouw et al. (2014) consider embodied theories of the cognitive function of gestures. As they see it, standard embodied accounts are too internalistic because they treat gestures as the epiphenomenal outputs of the sensorimotor processes involved in cognition. Pouw et al. argue that it would be more perspicuous to view gestures in terms of embedded/extended cognition (Kirsh, 1995; Clark, 2013; Wheeler, 2013) and treat them as external tools that can replace or support internal cognitive processes. In a related vein, Landy et al. (2014) defend an embodied account of symbolic reasoning in which external mathematical symbols and formulae serve as targets for action and perception systems. This account, which they refer to as Perceptual Manipulations Theory (PMT), suggests

that mathematical and logical reasoning often involves the sensorimotor systems engaged by physical notations. Perceptual processes exploiting the design features of physical notations underwrite significant aspects of symbolic reasoning. Landy et al. contend PMT is supported by the growing body of evidence demonstrating the manifold ways that sensorimotor processes can influence or interrupt the capacity for symbolic reasoning.

One of the insights behind embodied cognition is that cognitive science has been overly concerned with higher-level cognition. We should instead pay closer attention to lower-level phenomena and consider the cognitive behavior of animals and less complicated agents. When we do, the importance of the body becomes apparent in ways that can be obscured when we focus only on higher-level cognition. Such a bottom-up approach has an underappreciated consequence: it raises significant questions concerning the ontogenetic and phylogenetic emergence of higher-level capacities.

On the ontogenetic front, Wellsby and Pexman (2014) suggest that work needs to be done in order to integrate embodied cognition with the large body of extant research on the development of concepts and language processing in children. They outline several important issues that need to be addressed in order to carry out this research program. Using ideas from radical embodied cognition (Chemero, 2009), Cowley (2014) proposes that there is a symbiotic relationship between linguistic embodiment and external verbal constraints. He offers a distributed-ecological account of how language skills emerge through the dynamic coordination of movement with verbal patterns and social experience.

On the phylogenetic front, Stutz (2014) suggests that an embodied approach can help illuminate the emergence of central human phenotypes such as linguistic communication and symbolic representation. His embodied niche-construction (ENC) hypothesis holds that these are the result of the dynamic co-evolution of embodied forms of cognition and changing environmental interaction. More specifically, it maintains that the capacity to form recursive iconic narratives was an important evolutionary precursor to the emergence of both. Gapenne (2014) defends the hypothesis that proprioception plays a fundamental role in the co-constitution of the self and the world by a cognitive system. He explicitly maintains that the coupling of proprioception and action is an important development in the phylogenesis of even simple organisms.

## Conclusion

The aim of this research topic was to bring together experts from multiple disciplines to discuss the future of embodied cognition. The resulting contributions suggest that embodied cognition is a robust and dynamic research program—one that is focused on addressing recognized challenges, exploring new empirical ground, and expanding its theoretical reach. Taken as a whole, they demonstrate the ongoing fecundity of this approach. Questions certainly remain, but that itself might be a good sign.

## References

- Bahnmueller, J., Dresler, T., Ehlis, A.-C., Cress, U., and Nuerk, H.-C. (2014). NIRS in motion— unraveling the neurocognitive underpinnings of embodied numerical cognition. *Front. Psychol.* 5:743. doi: 10.3389/fpsyg.2014.00743
- Borghi, A. M., Capirci, O., Gianfreda, G., and Volterra, V. (2014). The body and the fading away of abstract concepts and words: a sign language analysis. *Front. Psychol.* 5:811. doi: 10.3389/fpsyg.2014.00811
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Clark, A. (2013). “Gesture as thought,” in *The Hand, an Organ of the Mind: What the Manual Tells the Mental*, ed Z. Radman (Cambridge, MA: MIT Press), 255–268.
- Corbetta, D., Thurman, S. L., Wiener, R. F., Guan, Y., and Williams, J. L. (2014). Mapping the feel of the arm with the sight of the object: on the embodied origins of infant reaching. *Front. Psychol.* 5:576. doi: 10.3389/fpsyg.2014.00576
- Cowley, S. J. (2014). Linguistic embodiment and verbal constraints: human cognition and the scales of time. *Front. Psychol.* 5:1085. doi: 10.3389/fpsyg.2014.01085
- Dove, G. O. (2015). Three symbol ungrounding problems: abstract concepts and the future of embodied cognition. *Psychon. Bull. Rev.* doi: 10.3758/s13423-015-0825-4. [Epub ahead of print].
- Gapenne, O. (2014). The co-constitution of the self and the world: action and proprioceptive coupling. *Front. Psychol.* 5:594. doi: 10.3389/fpsyg.2014.00594
- Kirsh, D. (1995). The intelligent use of space. *Artif. Intell.* 73, 31–68. doi: 10.1016/0004-3702(94)00017-U
- Landy, D., Allen, C., and Zednik, C. (2014). A perceptual account of symbolic reasoning. *Front. Psychol.* 5:275. doi: 10.3389/fpsyg.2014.00275
- Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., and Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Front. Psychol.* 5:359. doi: 10.3389/fpsyg.2014.00359
- Proffitt, D. R., and Linkenauger, S. A. (2013). “Perception viewed as a phenotypic expression,” in *Tutorials in Action Science*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 171–197. doi: 10.7551/mitpress/9780262018555.003.0007
- Pulvermüller, F., and Garagnani, M. (2014). From sensorimotor learning to memory cells in prefrontal and temporal association cortex: a neurocomputational study of disembodiment. *Cortex* 57, 1–21. doi: 10.1016/j.cortex.2014.02.015
- Soliman, T., Gibson, A., and Glenberg, A. M. (2013). Sensory motor mechanisms unify psychology: the embodiment of culture. *Front. Psychol.* 5:885. doi: 10.3389/fpsyg.2013.00885
- Soliman, T. M., and Glenberg, A. M. (2014). How intent to interact can affect action scaling of distance: reply to Wilson. *Front. Psychol.* 5:513. doi: 10.3389/fpsyg.2014.00513
- Stutz, A. J. (2014). Embodied niche construction in the hominin lineage: semiotic structure and sustained attention in human embodied cognition. *Front. Psychol.* 5:834. doi: 10.3389/fpsyg.2014.00834
- Troche, J., Crutch, S., and Reilly, J. (2014). Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Front. Psychol.* 5:360. doi: 10.3389/fpsyg.2014.00360
- Watson, C. E., Cardillo, E. R., Bromberger, B., and Chatterjee, A. (2014). The specificity of action knowledge in sensory and motor systems. *Front. Psychol.* 5:494. doi: 10.3389/fpsyg.2014.00494
- Watson, C. E., and Chatterjee, A. (2011). The functional neuroanatomy of actions. *Neurology* 76, 1428–1434. doi: 10.1212/WNL.0b013e3182166e2c
- Wellsby, M., and Pexman, P. M. (2014). Developing embodied cognition: insights from children’s concepts and language processing. *Front. Psychol.* 5:506. doi: 10.3389/fpsyg.2014.00506
- Wheeler, M. (2013). “Is cognition embedded or extended?” in *The Hand, an Organ of the Mind: What the Manual Tells the Mental*, ed Z. Radman (Cambridge, MA: MIT Press), 269–301.
- Wilson, A. D. (2014). Action scaling of distance perception is task specific and does not predict “the embodiment of culture”: a comment on Soliman, Gibson, and Glenberg (2013). *Front. Psychol.* 5:302. doi: 10.3389/fpsyg.2014.00302
- Zwaan, R. A. (2014). Embodiment and language comprehension: reframing the discussion. *Trends Cogn. Sci.* 18, 229–234. doi: 10.1016/j.tics.2014.02.008

**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 Dove. 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.



# The co-constitution of the self and the world: action and proprioceptive coupling

Olivier Gapenne\*

CNRS, BioMécanique et BioIngénierie, UMR 7338, Université de Technologie de Compiègne, Compiègne, France

## Edited by:

Guy Dove, University of Louisville,  
USA

## Reviewed by:

David Vaughn Becker, Arizona State  
University, USA

Guy Dove, University of Louisville,  
USA

## \*Correspondence:

Olivier Gapenne, CNRS,  
BioMécanique et BioIngénierie, UMR  
7338, Université de Technologie de  
Compiègne, CS 60319, 60203  
Compiègne Cedex, France  
e-mail: olivier.gapenne@utc.fr

This article proposes a theoretical reflection on the conditions for the constitution of a distinction between the self and the world by a cognitive system. The main hypothesis is the following: proprioception, as a sensory system that is habitually dedicated essentially to experience of the body, is conceived here as a coupling which is necessary for the dual and concomitant constitution of a bodily self and of a distal perceptual field. After recalling the singular characteristics of proprioceptive coupling, three lines of thought are developed. The first, which is notably inspired by research on sensory substitution, aims at emphasizing the indispensable role of action in the context of such perceptual learning. In a second part, this hypothesis is tested against opposing arguments. In particular, we shall discuss, in the context of what Braitenberg called a synthetic psychology, the emergence of oriented behaviors in simple robots that can be regulated by sensory regulations which are strictly external, since these robots do not have any form of “proprioception.” In the same vein, this part also provides the opportunity to discuss the argument concerning a bijective relation between action and proprioception; it has been argued by others that because of this strict bijection it is not possible for proprioception to be the basis for the constitution of an exteriority. The third part, which is more prospective, suggests that it is important to take the measure of the phylogenetic history of this exteriority, starting from unicellular organisms. Taking into account the literature which attests the existence of proprioception even amongst the most elementary living organisms, this leads us to propose that the coupling of proprioception to action is very primitive, and that the role we propose for it in the co-constitution of an exteriority and self is probably already at work in the simplest living organisms.

**Keywords: proprioception, sensory substitution, enaction, perception, coupling, self-world duality, cybernetics**

## INTRODUCTION

Inspired by the conjunction between the traditions of constructivism and phenomenology, which has been formulated and elaborated recently in the framework of the paradigm of enaction (Varela et al., 1991), this article proposes a reflection on the conditions for the constitution of a double perceptual polarity: that of the self (mainly a bodily self here), and that of a structured exteriority. In other words, how it is that a cognitive agent manages to constitute a “referential impression” of the lived world at the same time that it specifies itself. This constitution, or the genesis of a structured experience, comprises two aspects: the first concerns the fundamental properties of the objects that are co-constructed (self and/or world), such as substantiality, distality, figurability, tangibility, or yet again a sense of sameness; the second concerns the properties of the perceptual field itself as well as its englobing character (the fact that the agent experiences the feeling of being inside). We will not here exhaustively address all these properties. Rather, we propose to focus on the *initial* and *generic* conditions for this constitution of an organized process of appearing: firstly at the level of perceptual consciousness; and then at the level of a generalizing, imaginative, and anticipatory consciousness. First of all, we will recall the importance of “bodily action” as action produced by

an agent, and inducing sensory effects at the level of the same agent. This activity, conceived as sensory-motor or kinesthetic coupling, characterizes the concrete and continuous mode of relation that the agent entertains with its body and its environment (the dimension of what is present). The role of this coupling is to introduce a necessary variation which will form the basis for an activity of synthesis which will allow not only for feeling but also for the appearance of objects. A reminder of the situation of sensory substitution will serve as an example for this aspect.

Then – and this will be at the heart of this article – when one wishes to account for the constitution of the distinction between the self and the world, there is a necessity for the acting agent to make a distinction between two sources of variation in the sensory signals that affect it: those that are related to its own activity, and those that arise from the environment (considering that the perceived organization of this environment is not pre-defined). We may note that an absence of distinction, or a confusion, between these two sorts of signal directly threatens the agent since it favors the constitution of erroneous perceptions which may be deleterious, and are at the very least unsettling as in the case of illusions ofvection and self-motion. Thus, going further toward a definition of the mechanisms of the constitution of this phenomenological



dissociation between self and world, we propose a mechanism of “filtering and calibration” which allows an agent, when its sensory organs are submitted to variations in their states, to be able to attribute these variations either to its own activity (and thus as effects of its actions), or to events over which it has no control. In this way we develop the hypothesis, following on from the reflections of Poincaré (1902, p. 84) on the construction of perceptual and conceptual space, that the singularity of proprioception lies in the fact that it is a firm reference-point which enables this process able to play this role of “filtering and calibrating” (Declerck and Gapenne, 2009; Gapenne, 2010a,b; Blanchard et al., 2013). This will lead us to reflect upon the organized behaviors of certain artificial agents which do not possess proprioception, and to critically discuss theses which claim that it is possible to constitute spatiality solely on the basis of external sensory inputs.

Finally, in conclusion, we will redefine action as not being limited to the motor dimension of effective action, but as deriving from an organization where performance and sensation are necessarily coupled, and to postulate that motor-proprioceptive coupling plays a foundational role in the construction and the genesis of the enactive process of partitioning the self and the world, the inside and the outside, and so on.

## ACTION-SENSATION COUPLING IN SENSORY SUBSTITUTION

Amongst the various fields of research which have confirmed the importance of embodiment and action, as starting-points for the constitution of a process of appearing, the work of Bach y Rita in the 1960s on sensory substitution by means of a special device (TVSS: tactile vision sensory substitution) holds a prime place (Bach Y Rita, 1972). By actively using this device (see below for a detailed description), blind or blindfolded participants are able to perceive distal events (the position and the form of a 3D object) as in vision and to improve significantly their performances (discriminate objects in a scene and manage the interposition) by learning. Right from the start, many authors such as Paillard (1971) did not fail to emphasize the interest of this work, which opens up the possibility for the precise experimental study of the genesis of a form of perception which derives from action. This study appeared all the more original in that it mobilized proximal sense-organs (in this case, tactile sensor) in the constitution of an experience of an object at a distance without any direct contact. Although many summaries of these studies have already been published (e.g., Kaczmarek et al., 1991), we consider that it is useful to reformulate here the principle of sensory substitution. Technically, sensory substitution requires the insertion of an activator or stimulator (or a whole set of activators or stimulators) as an intermediary between two sensory systems, one artificial and the other natural. In other words, the “substitution” involves a doubling of the stimulation and thus a doubling of the transduction<sup>1</sup>: an artificial transduction (via

a sensory device = transduction 1), and a natural transduction (via a functional sensory system = transduction 2). This double transduction, via the insertion of an artificial captor and activator, makes it possible to provide access to a sensory flow which would not be available without this technical mediation. In the pioneering work of Bach-y-Rita et al. (1969) on the TVSS, it was a question of providing blind persons with access to an optical flow, via a camera (transduction 1), with electro-mechanical stimulators which relayed the signal from the camera and stimulated natural sensory organs which were available, i.e., the tactile receptors of the skin (transduction 2). As can be appreciated immediately, and as many subsequent developments have shown in practice (for relatively recent reviews see Wall and Brewster, 2006; Visell, 2009), this principle of substitution can theoretically substitute any sort of flow by any other (visual–auditory, auditory–tactile, tactile–tactile, etc).

However, the use of these instruments rapidly revealed that the substitution is not limited to this double transduction in the sense of a two-stage transfer of input signals to the nervous system. Firstly, it is imperative that the signals that are transmitted should be subject to variation. Secondly, and this is the really essential point, the “substitution” only becomes effective if this variation is amenable to interpretation; and the key condition for this is that the variation in question should be determined by the user. It must therefore be well understood that the constitution of the properties, and in particular the spatial properties of the flow that is substituted (for example, vision being substituted by the tactile modality) does not derive from simply capturing the spatiality inherent in the organization of the network of activators which deliver the signals (for example, a square 20 x 20 matrix of 400 activators in the case of the TVSS). In this sense, the logic of the constitution of perceptual experience, and more generally of cognitive experience, via this type of device cannot be limited solely to the double transduction of signals whose variation arises from external events. This variation must be an *active* variation, i.e., the variation must be produced and controlled by the agent. Thus, the “substitution,” as a process which is equipped, must also include the tool of an *inverse* double transduction corresponding to the action produced by the body with respect to the instrument, an action producing a movement of the instrument with respect to the environment.

It is thus essential to understand that the “substitution” cannot be solely sensory; this has led us to propose that the substitution is rather perceptual, in the sense that it involves a moto-sensory<sup>2</sup> coupling whose closure is ensured by the technical system on one hand and the user on the other. In addition – and this is in a way a consequence of the preceding point – the process involved is not properly speaking a “substitution,” but rather what we have called a “supplementation” (Lenay et al., 2003), this latter term being conceptually more adequate. And as a matter of fact, the system that is called “tactile–visual substitution” does not give rise to a truly *visual* experience as such.

<sup>1</sup>This generic term designates any mechanism which performs the conversion of a signal of one sort into an equivalent signal of another sort. Thus, any sort of sensory organ (photoreceptor, semi-circular canal, or whatever) performs a transduction, which is different for each of them.

<sup>2</sup>Although the term “sensory-motor” is more frequent in the literature, we prefer here the inverse formula, besides having the merit of emphasizing the primacy of the action, it also affirms both its role in producing variation in the sensory input, and the importance of the agency of the movement.



The instrument, in particular when it is actively taken in hand, opens up an unprecedented space of experience, which makes it possible to interpret certain properties of the “novel” moto-sensory flow. And in the case of the TVSS, it is remarkable that the instrumented activity makes it possible to interpret distal spatial qualities on the basis of proximal tactile signals. Guarniero (1974) evidences that after several hours of use, a blind user is able to recognize simple objects at a distance, including moving objects, and to interpret certain events as interpositions.

A final point that is worth mentioning is that the stimuli delivered by the tactile stimulators are not forces of a sort which would constrain the movements of the subject; this is in contrast to devices such as the robotic arm PHANTOM Desktop. With the TVSS, the stimulation consists of a pressure on the skin, but it does not deliver a return of effort of a kind which could guide the movement. This is an essential point because, although it involves a tactile activator, the TVSS is an interface which is “gestural,” and in this sense much closer to visual gestures. Indeed, the movements of the ocular globe are produced without any constraint from the optical flow, since this flow does not deliver any forces such that the movement of the ocular globe would be mechanically affected and guided. In other words, the tactile stimulations of the TVSS do not directly constrain the movements of the agent. Thus, in the two cases, the control of the movement must be actively produced by the agent – and this is a quite general situation. In this context, a gesture (an organized exploratory movement) can be minimally described as an attractor where each state must be defined by at least two parameters: a definite position of the point of action in (x, y, z) co-ordinates; and a value of the sensation (0 or 1) indicating the absence or presence of an event in the environment. The temporal succession of these states (x,y,z,e) describes a trajectory that we may define as a “gesture,” or alternatively as a “strategy” (Stewart and Gapenne, 2004). In this situation, what the subject receives at each point in time is just a sensation (or a set of sensations), and the mere projection of this sensation onto the sensory organ is not sufficient to initiate perceptual activity. If the subjects do succeed in perceiving “objects,” it can only be through their active exploration, and by integrating over time their movements, the tactile sensations, and their kinesthetic sensations. Thus, the situation of perceptual supplementation is exemplary because, quite besides the technical innovation, it makes it possible to re-create at a micro-developmental scale a situation of perceptual learning. Even though this learning does not have exactly the same meaning for an adult and for a newborn child, we can nevertheless follow through the necessary steps for the mastery of a new mode of coupling.

In another technical context, inspired by the work of Meijer (1992), Auvray et al. (2005, 2007) has proposed a description of the steps involved in the appropriation of a device by sighted adult subjects. Without going into the fine details of the succession of all these stages, let us consider the first two which are of particular interest here. The first stage is called “contact”; it involves learning the sensory-motor regularities necessary to stabilize and to actively maintain perceptual contact with the stimulus. As for the second stage, labeled “distal attribution,” it corresponds to understanding the origin of the sensations as deriving from the fact of

making contact with an object situated in the perceptual space opened up by the tool. This second stage is perhaps unfortunately labeled, since it risks confusing the fact that the variation in the sensations has an origin which is *distinct* (i.e., not related to the determination of my actions) as compared to an origin which is *spatially distant*, which is of course not the same thing. In this situation, as in the original experiment of Epstein et al. (1986), the participants using a sensory substitution device but not being informed about its functioning are asked for the nature of what they perceived and had to make a choice among several scenarios (e.g., “sensors, located on my head and hand, record the locations of my head and hand and produce different stimulation intensity levels whenever those locations change.” or “a camera, located in front of me, detects both hand and head movements and sends a signal to the device whenever movement is initiated.”) that proposed a rationale for what was happening. The point of interest is that the subjects produce sensory variations as a result of their own movements; but, taking into account the fact that the subjects are ignorant as to the experimental setup, the situation remains somewhat ambiguous so that the interpretation of the variation in the stimuli is not necessarily that of a determination through agency. And even when it is, the subjects have great difficulty in considering that the source of these variations may be external and distant. It is clearly apparent that whereas at the stage of contact the subjects often succeed, in the experiments of Epstein et al. (1986) and Auvray et al. (2005), in expressing their consciousness of the relation between their actions and the reafferent sensations, this is because the source is fixed and cannot produce a stimulation unknown to the subject if the latter is immobile and not stimulated. Nevertheless, the sensitivity to the spatio-temporal coincidence between the movement and the tactile reafference does not seem to be so obvious to all the subjects. This point is important, since it indicates that even in such favorable conditions the interpretation in terms of agency is not guaranteed with an external source, and it is necessary to introduce certain conditions of manipulating the coupling (for example by giving the possibility of interposing a screen between the sensory cap-tor and the source) in order to lift the ambiguity (Auvray et al., 2005).

To sum up this section, and referring to the work on sensory substitution, we will note three main points. Firstly, *modulo* the necessary movement by a suitably equipped agent, it is possible to constitute a distinct, distal appearance. Secondly, this appearance is not reducible to an analysis of the tactile sensations or of the movements produced in order to determine them; in both cases, the tactile and kinesthetic sensations are “forgotten” and replaced by a consciousness focused on the events in the environment. Thirdly, if the subjects are not informed about the properties of the coupling system (for example the TVSS), and are not informed about what there is to be perceived by specifying explicitly that the source is clearly positioned “out there” at a distance, it seems that the experience of agency is not guaranteed. This being so, with respect to our question concerning the constitution of the self/world distinction, the analyses which have been carried out so far by means of the experiments of sensory substitution/perceptual supplementation only provide us with partial answers as to the conditions of this constitution.

One of the reasons for this is that the studies in this domain have been concerned above all with the constitution of the “object” pole; the other pole, that of the “subject,” is referred to an implicit, pre-reflexive register which at best plays a role of motivation, and not really of exposition (Husserl, 1989). It is to be noted immediately here that the problem arises from the impossibility of having simultaneous experience of the two poles; in fact, certain studies have clearly shown that it is possible to have recourse to the principle of sensory substitution in order to constitute/recover bodily experiences (Tyler et al., 2003). In line with this, the “subject” pole can also be recovered in this way, to the extent that it refers to a kinesthetic, bodily experience. In this light, we come to realize that there is a point which has remained obscure in all these analyses: and this is, to understand *how* the sensory flow generated by the active movement by the agent, which determines the variation in the flow, can actually be partitioned by the agent. One way to lift the veil of mystery would be to consider that the deployment of each movement is always associated with a *double* reafferent flow (here, a tactile flow and a proprioceptive flow). One of these reafferent flows (the tactile flow) would be contingent, and the other one (the proprioceptive flow) would be absolute – at least to a first approximation. The hypothesis would then be that the proprioceptive system contributes to a filtering, since it provides the agent with a non-ambiguous indication as to whether he/she is active or not. We shall develop this hypothesis and make it more precise in the next section. In order to close the present section, we will remark again that if the tool can be “forgotten” when it contributes to the accession to an experience of the self and/or an exteriority, this “forgetting” also concerns the tactile sensations as such. It would seem interesting to delve more deeply into this “disappearance from experience” which occurs at the level of receptors such as the retina or the cochlea. In the case of a prolonged and intensive use of the TVSS, would one arrive at a stage when the tactile sensations would have become just as inaccessible as retinal sensations?

## THE SINGULARITY OF PROPRIOCEPTION

In the matter of proprioception, it is important to be very precise (Stillman, 2002). Only too often, and wrongly, “proprioception” is misleadingly over-represented as the perception of self as an embodied, acting agent. But it is obvious, as indicated by the unfortunate expression “proprioceptive function” as coined by Gibson, that this sort of perception of bodily activity and the self involves many (and indeed, *a priori*, all) perceptual systems. For this reason, rather than the term “proprioceptive function,” we prefer the term “kinesthetic function” which does properly refer to the multimodal experience of the body at rest or in movement, static or dynamic. We will reserve the term “proprioception” as one specific perceptual system among others, which is indeed involved in the experience (and the regulation) of movement, posture and balance; but as we shall see, proprioception also does more than this. Anatomically, the proprioceptive system mobilizes sensory organs, afferent innervations, and specific cortical structures which are known in part today (McCloskey, 1978; Hogervorst and Brand, 1998; Romaiguère et al., 2003). A notable feature of this system is that all the sensory organs are localized in the core of effectors

(muscles, tendons, articulations) involved in the maintenance and the animation of the skeleton. It is thus a case not just of relative proximity, but of genuine contiguity between the sensory organ and the effector. It is thus important to understand that variation in the activity of the proprioceptive organs (neuromuscular spindles, neurotendinous organs, or articular receptors), variation which is necessary for them to function, is intimately related to variation in the activity of the effector itself. This has led proprioception to be called “muscular sense” ever since its first description by Bell (1826), including mainly a sensitivity to movement and to position. The specificity of proprioception derives from the fact that all the other sensory organs respond to variations essentially linked to mechanical, chemical, optical, or other flows which come from the environment. More precisely, since a living organism is never completely static (physiological tremor, ocular micro-nystagmus), sensory organs can receive variations in input whose amplitude cannot be directly related to the amplitude of movements of the agent (Lockhead, 1992). The crucial point here is thus that variation in the stimulation of all the sensory organs, with the sole exception of proprioception, is linked to bodily engagement but is always liable to be compounded with the effects of events external to the agent. In other words, variation in the activity of the sensory organs, which is itself linked to variation of their source, is always potentially composite and ambiguous quite simply because the source of this variation is potentially dual (and, in the event, almost always is dual since there is a mix of variation due to the agent itself and variation due to external events). The proprioceptive system thus has this prime singularity, that it is always activated by deformations of the body, and (in natural conditions) by nothing else. As experimental studies in humans and animals have already suggested empirically, the consequence of this is that if proprioception plays an unquestionable role in the perception of bodily events (Farrer et al., 2003), it can also play a role in the perception of external events and, more fundamentally, in the genesis of the perception of such events (Buisseret et al., 1988; Roll et al., 1991). We therefore suggest that by having at its disposal a moto-sensory system strictly associated with its own activity, the agent possesses a powerful tool for filtering and calibrating signals for which it does not control the determinism.

Philipona et al. (2003), far from any immediately phenomenological considerations, have taken up this line of argument on a strictly formal basis, and have proposed an algorithm, based on inputs and outputs, which is apparently able to deduce the geometry and the dimensions of an external space without any *a priori* knowledge. The calculating artifact (a virtual poly-articulated robot) has at its disposal input signals coming from two types of sensors: sensors which are sensitive to changes in the positions of the articulated segments of the robot, and sensors which are sensitive to the presence of light in the unknown virtual environment. In addition, effector organs controlled by the algorithm produce the movements of the robot. In the first instance, the algorithm distinguishes two types of signal according to whether they are related to its own movements or not. In other words, the algorithm can discriminate between exteroceptive signals which are produced when the robot is static, and proprioceptive signals which possess a bijective relation with its

own movement (“certain inputs react always in the same way to motor command,” Philipona et al., 2003, p. 3). Then, in a second instance, the robot (or rather its “brain”) is able to distinguish between two sorts of exteroceptive signals: *exafferent* signals which are independent of its own movement (when the robot is static and the sensors are subject to variations), which enables the induction of a vector called “*representation of the state of the environment*”; and signals which are associated with movements of the robot (whose sensors are subject to variations related to *reafferent* signals), which enables the induction of a vector called “*representation of the exteroceptive body*.”

Now although this study does have the interest of proposing a possible mathematical formulation of the distinctions which an organism can make in order to perceive itself as different from its environment, it has serious limitations. In particular, it does not treat the case of a double source of stimulation when the robot is in movement (a *combination* of exafferent and reafferent signals), which is in the end the crucial situation for a living organism, and which is at the core of the dilemma we have to deal with. Moreover, the relation between proprioception and exteroceptive reafferents is envisaged merely as a possible intersection. From our point of view, the constitution of a genuine process of appearing (i.e., the microgenesis of perception) requires a genuine articulation, and not just a contingent intersection between entities that are presupposed to be distinct. Finally, besides the hypothesis of a bijective relation between action and sensation in the case of proprioception (see below), and its limitation in this model to a capture of position (even if movement and position are to some extent correlated), the hypothesis that the motor commands – which will prime the moto-sensory coupling and thus prime the subsequent inferences realized by the “brain” – are produced “at random” remains mysterious. Where do these commands come from? Why do they take the form that they do? Are they generated by a “program”? As I will say below, this conception of commands as pure effectuation does not seem adequate in the case of living organisms.

A second singularity of proprioception is that these sensors do not seem to be submitted to the activity of an efferent sensory system, as are all other sensory systems (e.g., Warr, 1975). In addition, the activity of proprioceptive receptors does not seem to be modulated by anything other than the activity of the effectors to which they are linked. The receptors or the primary and secondary sensory nerve-endings situated in the equatorial zone of the muscular fibers present a variation in their potential as a function of the modulation of the tension of the muscular tension. And even in the case of the gamma loop, the neurons emanating from the anterior horn of the spinal cord are moto-neurons which modulate the stretching of the fibers, but they are in no way sensory efferent fibers which modulate the activity of the sensory nerve-endings themselves. This anatomical particularity has functional consequences. The activation of afferent proprioceptive fibers can modulate the behavior of the receptors of other sensory systems via their action at the level of central nuclei from which efferent fibers leave toward the other sensory receptors. Conversely, the other sensory systems are not able to carry out such a modulation, other than

indirectly via the modulation of the tension of the muscular fibers.

## THE ABSENCE OF PROPRIOCEPTION AND THE BIJECTION ACTION/PROPRIOCEPTION

In order to discuss the theoretical proposition formulated above concerning the role of proprioception, and maybe to contest it, we shall now consider two arguments which go against it: one of these argument is empirical and factual, and the other is theoretical. The first argument refers to the possibility of producing spatially organized behavior without any recourse to proprioception; the second posits that the constitution of a space is impossible if it is admitted that the relation action-proprioception is bijective.

In his essay in synthetic psychology, Braitenberg (1986) presents some very simple robotic architectures based on direct connections between sensors and effectors, which are nevertheless sufficient for the mobile robots to exhibit distinctive behaviors, such as attraction and repulsion, with respect to a source. At no point in his short and fascinating text does Braitenberg even so much as mention the very idea of proprioception – which leads him, in fact, to put forward some very internalist and representationalist propositions. We may recall here that the famous “tortoises” of Gray Walter (*Machina Speculatrix*) were likewise bereft of any proprioception (they possessed only a shock-sensor), and were already able to exhibit behavior such as “return to the nest,” an “attractive” site where the tortoise could recharge its energy; this site possessed a light which served as an external source for guiding the tortoise. Let us consider then this case of a displacement toward a source of light. The robots were equipped with a photo-electric cell (a photo-sensitive sensor); detection of the light was supposed to produce exploratory movements which here were of two types, “translation” or “rotation.” The composition of these two sorts of movement produces a sinusoidal (or ellipsoidal) trajectory, whose amplitude theoretically tends to decrease as the robot approaches the source. What can we learn from the emergent behaviors produced by these automata? It is clearly a case of emergence, in the sense that the trajectory produced by the agent, and described by the observer, is in no way programmed as such (even though it results from the operation of an electronic circuit), and it is not learned. These behaviors demonstrate that an agent, even an artificial agent, can produce spatially organized behavior without any recourse to “proprioceptive” signals concerning its own material architecture and its own movement. This self-organization does however, have some limits, in particular concerning the choice of the material architecture and the possibilities of action which are associated with it. We may note that, unlike the virtual robot of Philipona et al. (2003) described above, these robots do not have any proprioception and so the problem of “partitioning” simply does not arise. Moreover, the problem of portioning “external” signals as arising from the movement of the agent versus that of the environment cannot be resolved by intersecting external and proprioceptive flows of sensation. So what, after all, does this tropism toward a light-source tell us? It indicates that the action of the agent (the activation of a motor producing the rotation of the wheels) can be controlled by the capture of a contingent “external” signal on

which feedback is applied. But then, with respect to our hypothesis concerning the deleterious consequences of confusion concerning the source of variation, why in the case of these robots does this not cause totally aberrant behavior? When the photo-electric cell is activated, the robot cannot “interpret” this activation as being necessarily related to its own rotation (the light-source is fixed), because it does not have any signals concerning its own movement. So what could possibly constitute a “pathological” behavior in this case? This strictly external guidance of the actions which are successively produced rests on the tolerance of a fusion of the sources of contingency: the light-source can be displaced by the experimenter, or the movement of the robot can produce a displacement of the sensor, such that it is no longer in phase with the source. And in fact, an examination of the concrete situations reveals that the regulation occurs in the succession of these two modes of variation, and does not tolerate well their concurrence. However, and this is a key point, the great majority of natural situations do expose the agents to the simultaneity of the variations.

Of course, this tropism toward a light-source is reminiscent of the way bacteria climb a glucose gradient; we will come back to this point, to suggest that the management of this simultaneity by a living organism is not of the same order as the Braitenberg robots, and as in the case of micro-organisms, does not need a central nervous system to be achieved.

The argument concerning the bijection action-sensation is in a way the counterpoint to the preceding question. If one admits the existence of an agent which would possess only proprioception, such an agent would not be able to have access to any variations other than those produced by its own actions, and it would therefore be in a situation where the variations are totally determined (Piaget, 1937; Lenay, 2006). In this case, no opening toward the exterior would be possible, and neither would an access to the bodily self on the basis of the actual variations. This argument is often invoked, on the one hand to affirm that proprioception alone, in and of itself, cannot open the way to spatiality; and on the other hand, it constitutes a risk of a return to a representationalist conception of bodily experience. Both of these risks are real. However, this hypothetical situation and the associated risks should be put in due perspective. Firstly, there is no known living organism whose organization is founded strictly and solely on proprioception. All known living organisms do have two sorts of sensors, those that are proprioceptive, the others which are sensitive to events which are totally or partly independent of the actions of the organism. The question is thus not so much that of a total determinism of the moto-proprioceptive loop, but rather that of the articulation between this loop and the others. Secondly, one can question the status of a possible bijection; and also ask questions about the bijection itself. If the hypothetical bijection supposes that the motor command, specifying a precise value for a parameter of position, speed or other, has the effect of producing a corresponding unique value at the level of the sensor, this supposition postulates anew that the command/action is a matter of pure effectuation, and tends to deny the importance of the differential of the activity of the sensor. As for the bijection itself, it may be doubted whether it could ever actually be realized, not only because the bandwidth for proprioceptive

sensors is limited and their response not so reliable (Wann and Ibrahim, 1991), but also and above all because of the principle of functional ambiguity which refers to the radical impossibility for a command to totally anticipate the concrete realization of the action. In particular, gravitation and friction always leave a certain degree of uncertainty concerning the movement which will actually occur. These variations, which cannot be determined by the command, are actually a condition for the possibility of constituting an experience of the body/self – even if, as we have already said, this kinesthetic experience involves the set of sensory organs as a whole.

## LIFE AND THE SELF-WORLD DUALITY

In this article we have proposed that the constitution of an experience of the distinction between the self and the external world supposes that the agent has at its disposal a way of coupling its means of action and its means of sensation; the latter being sensitive to variations in the signal that are related, or not, to the effects of actions produced by the agent itself. We have also postulated that moto-proprioceptive coupling plays a decisive role in this constitution, to the extent that it allows for the advent of a referent with respect to which other sensory signals can be sorted and calibrated. We have insisted on this function of sorting, because it seems to us to be indispensable, *via* action, in the constitution of two distinct poles of experience, that of the subject and that of the object. On this point, we wish to draw attention to the fact that even the simplest forms of life (even before the advent of a nervous system) possess both a system of action and a double form of sensors (proprioceptive and others). Thus, we may venture to suggest that the hypothesis we develop here, which is valid for complex perceptual systems, actually corresponds to a mechanism which is much more general and which is common to all forms of life as they exist from the unicellular scale onward (Iscla and Blount, 2012; Lebois et al., 2012). Thus life, in its primary organization, never exists in a pure feed-forward mode; pure effectuation does not seem to exist; this is in the end compatible with the circular forms of organization characteristic of the later cybernetic approaches. It remains to launch an enquiry into the genesis of the sensor/effector partition in the course of the advent of life itself.

On this basis, and in coherent fashion at the theoretical level, we are led to formulate the three following points:

- (1) With reference to the theory of autopoiesis, the paradigm of enaction poses that cognition implies an organization that is proper to living organisms.
- (2) This organization, which involves a characteristic circularity at the level of metabolism (a network of elements which produce the elements necessary for its own functioning), replays the scheme of circularity at the nascent sensory-motor level in terms of the relation between effectors and sensors.
- (3) The notion of action as pure effectuation no longer having a place (outside the context of the mechanistic conception of automata), it could give way to the concept of enaction which is more favorable to an effort at the characterization of the organization and the lived experience of living and thinking agents.



By way of conclusion, it appears that the next theoretical step will aim at developing the conception of a form of enactive memory which escapes from the bounds of current coupling, without reducing it to a simple representation that can be activated on an occasional basis. Such a memory could be the basis for justifying the appearance of the self and the world.

## ACKNOWLEDGMENTS

The author thanks warmly John Stewart for his translation of the text, and the latter and Gunnar Declerck for their comments.

## REFERENCES

- Auvray, M., Hanneton, S., and O'Regan, J. K. (2007). Learning to perceive with a visuo-auditory substitution system: localization and object recognition with 'the voice'. *Perception* 36, 416–430. doi: 10.1068/p5631
- Auvray, M., Hanneton, S., Lenay, C., and O'Regan, J. K. (2005). There is something out there: distal attribution in sensory substitution, twenty years later. *J. Integr. Neurosci.* 4, 505–521. doi: 10.1142/S0219635205010102
- Bach Y Rita, P. (1972). *Brain Mechanism in Sensory Substitution*. New York: Academic Press.
- Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., and Scadden, L. (1969). Vision substitution by tactile image projection. *Nature* 221, 963–964. doi: 10.1038/221963a0
- Bell, C. (1826). On the nervous circle which connects the voluntary muscles with the brain. *Philos. Trans. Roy. Soc. (Lond.)* 116, 163–173 doi: 10.1098/rstl.1826.0016
- Blanchard, C., Roll, R., Roll, J. P., and Kavounoudias, A. (2013). Differential contributions of vision, touch and muscle proprioception to the coding of hand movements. *PLoS ONE* 8:e62475. doi: 10.1371/journal.pone.0062475
- Braitenberg, V. (1986). *Vehicles: Experiments in Synthetic Psychology*. Bradford: A Bradford Book.
- Buisseret, P., Gary-Bobo, E., and Milleret, C. (1988). Development of the kitten visual cortex depends on the relationship between the plane of eye movements and visual inputs. *Exp. Brain Res.* 72, 883–948. doi: 10.1007/BF00248503
- Declerck, G., and Gapenne, O. (2009). Actuality and possibility: on the complementarity of two registers in the bodily constitution of experience. *Phenomenol. Cogn. Sci.* 8, 285–305. doi: 10.1007/s11097-009-9128-4
- Epstein, W., Hugues, B., Schneider, S., and Bach y Rita, P. (1986). Is there anything out there? A study of distal attribution in response to vibrotactile stimulation. *Perception* 15, 275–284. doi: 10.1068/p150275
- Farrer, C., Franck, N., Paillard, J., and Jeannerod, M. (2003). The role of proprioception in action recognition. *Conscious. Cogn.* 12, 609–619. doi: 10.1016/S1053-8100(03)00047-3
- Gapenne, O. (2010a). Altération de la matrice proprioceptive et psychose naissante: une hypothèse. *Evol. Psychiatr.* 75, 383–393. doi: 10.1016/j.evopsy.2010.06.005
- Gapenne, O. (2010b). "Kinaesthetics and the construction of perceptual objects," in *Enaction: A New Paradigm for Cognitive Science*, eds J. Stewart, O. Gapenne, and E. Di Paolo (Cambridge: MIT Press), 183–218.
- Guarniero, G. (1974). Experience in tactile vision. *Perception* 3, 101–105. doi: 10.1068/p030101
- Hogervorst, T., and Brand, R. (1998). Current concepts review: mechanoreceptors in joint function. *J. Bone Joint Surg.* 80, 1365–1378.
- Husserl, E. (1989). *Chose et Espace. Leçons de 1907*, trans. J.-F. Lavigne (Paris: PUF).
- Iscla, I., and Blount, P. (2012). Sensing and responding to membrane tension: the bacterial MscL channel as a model system. *Biophys. J.* 103, 169–174. doi: 10.1016/j.bpj.2012.06.021
- Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., and Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Trans. Biomed. Eng.* 38, 1–16. doi: 10.1109/10.68204
- Lebois, F., Sauvage, P., Py, C., Cardoso, O., Ladoux, B., Hersen, P., et al. (2012). Locomotion control of *Caenorhabditis elegans* through confinement. *Biophys. J.* 102, 2791–2798. doi: 10.1016/j.bpj.2012.04.051
- Lenay, C. (2006). Enaction, externalisme et suppléance perceptive. *Intellectica* 43, 27–52.
- Lenay, C., Gapenne, O., Hanneton, S., Marque, C., and Genouëlle, C. (2003). "Sensory substitution, limits and perspectives," in *Touch for Knowing*, eds Y. Hatwell, A. Streri, and E. Gentaz (Amsterdam: John Benjamins), 275–292.
- Lockhead, G. R. (1992). Psychophysical scaling: judgments of attributes or objects. *Behav. Brain Sci.* 15, 543–601. doi: 10.1017/S0140525X00069934
- McCloskey, D. I. (1978). Kinaesthetic sensibility. *Physiol. Rev.* 58, 763–820.
- Meijer, P. B. L. (1992). An experimental system for auditory image representations. *IEEE Trans. Biomed. Eng.* 39, 112–121. doi: 10.1109/10.121642
- Paillard, J. (1971). Les déterminants moteurs de l'organisation de l'espace. *Cah. Psychol.* 14, 261–316.
- Philipona, D., O'Regan, K., and Nadal, J. P. (2003). Is there something out there? Inferring space from sensorimotor dependencies. *Neural Comput.* 15, 2029–2049. doi: 10.1162/089976603322297278
- Piaget, J. (1937). *La Construction du Réel chez l'Enfant*. Neuchâtel: Delachaux et Niestlé.
- Poincaré, H. (1902). *La Science et l'Hypothèse*. Paris: Flammarion.
- Roll, R., Velay, J.-L., and Roll, J.-P. (1991). Eye and neck proprioceptive messages contribute to the spatial coding of retinal input in visually oriented activities. *Exp. Brain Res.* 85, 423–431. doi: 10.1007/BF00229419
- Romaiguère, P., Anton, J. L., Roth, M., Casini, L., and Roll, J.-P. (2003). Kinaesthetic activates both motor and parietal cortical areas in humans: a parametric fMRI study. *Cogn. Brain Res.* 16, 74–83. doi: 10.1016/S0926-6410(02)00221-5
- Stewart, J., and Gapenne, O. (2004). Reciprocal modelling of active perception of 2-D forms in a simple tactile-vision substitution system. *Minds Mach.* 14, 309–330. doi: 10.1023/B:MIND.0000035423.93112.b2
- Stewart, J., Gapenne, O., and Di Paolo, E. (2010). *Enaction: A New Paradigm for Cognitive Science*. Cambridge: MIT Press. doi: 10.7551/mitpress/9780262014601.001.0001
- Stillman, B. C. (2002). Making sense of proprioception: the meaning of proprioception, kinaesthesia and related terms. *Physiotherapy* 88, 667–646. doi: 10.1016/S0031-9406(05)60109-5
- Tyler, M., Danilov, Y., and Bach y Rita, P. (2003). Closing an open-loop control system: vestibular substitution through the tongue. *J. Integr. Neurosci.* 2, 159–164. doi: 10.1142/S0219635203000263
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Visell, Y. (2009). Tactile sensory substitution: models for enaction in HCI. *Interact. Comput.* 21, 38–53. doi: 10.1016/j.intcom.2008.08.004
- Wall, S. A., and Brewster, S. (2006). Sensory substitution using tactile pin arrays: human factors, technology and applications. *Signal Process.* 86, 3674–3695. doi: 10.1016/j.sigpro.2006.02.048
- Wann, J. P., and Ibrahim, S. F. (1991). Does limb proprioception drift? *Exp. Brain Res.* 91, 162–166. doi: 10.1007/BF00230024
- Warr, W. B. (1975). Olivocochlear and vestibular efferent neurons of feline brainstem – their location, morphology and number determined by retrograde axonal transport and acetylcholinesterase histochemistry. *J. Comp. Neurol.* 161, 159–181. doi: 10.1002/cne.901610203

**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.

Received: 31 January 2014; accepted: 27 May 2014; published online: 12 June 2014.  
Citation: Gapenne O (2014) The co-constitution of the self and the world: action and proprioceptive coupling. *Front. Psychol.* 5:594. doi: 10.3389/fpsyg.2014.00594  
This article was submitted to Cognitive Science, a section of the journal *Frontiers in Psychology*.  
Copyright © 2014 Gapenne. 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.



# Mapping the feel of the arm with the sight of the object: on the embodied origins of infant reaching

Daniela Corbetta<sup>1\*</sup>, Sabrina L. Thurman<sup>2</sup>, Rebecca F. Wiener<sup>2</sup>, Yu Guan<sup>2</sup> and Joshua L. Williams<sup>3</sup>

<sup>1</sup> Director Infant Perception-Action Laboratory, Department of Psychology, The University of Tennessee, Knoxville, TN, USA

<sup>2</sup> Department of Psychology, The University of Tennessee, Knoxville, TN, USA

<sup>3</sup> Department of Psychology, Armstrong State University, Savannah, GA, USA

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Claudia Scorolli, University of Bologna, Italy

Penny M. Pexman, University of Calgary, Canada

Nicholas A. Holt, University of Louisville, USA

## \*Correspondence:

Daniela Corbetta, Department of Psychology, The University of Tennessee, 303D Austin Peay Building, Knoxville, TN 37996, USA  
e-mail: dcorbett@utk.edu

For decades, the emergence and progression of infant reaching was assumed to be largely under the control of vision. More recently, however, the guiding role of vision in the emergence of reaching has been downplayed. Studies found that young infants can reach in the dark without seeing their hand and that corrections in infants' initial hand trajectories are not the result of visual guidance of the hand, but rather the product of poor movement speed calibration to the goal. As a result, it has been proposed that learning to reach is an embodied process requiring infants to explore proprioceptively different movement solutions, before they can accurately map their actions onto the intended goal. Such an account, however, could still assume a preponderant (or prospective) role of vision, where the movement is being monitored with the scope of approximating a future goal-location defined visually. At reach onset, it is unknown if infants map their action onto their vision, vision onto their action, or both. To examine how infants learn to map the feel of their hand with the sight of the object, we tracked the object-directed looking behavior (via eye-tracking) of three infants followed weekly over an 11-week period throughout the transition to reaching. We also examined where they contacted the object. We find that with some objects, infants do not learn to align their reach to where they look, but rather learn to align their look to where they reach. We propose that the emergence of reaching is the product of a deeply embodied process, in which infants first learn how to direct their movement in space using proprioceptive and haptic feedback from self-produced movement contingencies with the environment. As they do so, they learn to map visual attention onto these bodily centered experiences, not the reverse. We suggest that this early visuo-motor mapping is critical for the formation of visually-elicited, prospective movement control.

**Keywords:** reaching, eye-tracking, human infants, visuo-motor mapping, embodiment, longitudinal study, skill emergence, object-directed visual attention

## INTRODUCTION

Reaching for objects is a fundamental skill that emerges in infancy around 3–5 months of age. Understanding how this skill forms and develops has been a core area of study in developmental psychology since the 1930s (Halverson, 1931; Piaget, 1936/1952; Gesell and Amatruda, 1946). Indeed, the onset of object-directed reaching marks an important transition in the development of infants' voluntary activity and provides essential foundations for the development and refinement of future motor, perceptual, and cognitive behaviors (Bushnell and Boudreau, 1993; von Hofsten, 1993, 2009). Despite extensive research in this area, the process by which infants learn to bring their arm in contact with a wanted object is still open to much investigation. For the longest time, the emergence of reaching was thought to be under the control of visual guidance. It was assumed that infants needed to see their hand in order to steer it toward the target (e.g., White et al., 1964; von Hofsten, 1979; Bushnell, 1985). But in recent decades, researchers have begun to question the guiding role of vision for the emergence of infant reaching. Some demonstrated that from

their earliest attempts, infants can reach in the dark toward a glowing target without seeing their hand (Clifton et al., 1993). This suggested that infants rely primarily on proprioceptive information, not vision, to begin controlling and directing their arm toward a specific location in space (Thelen et al., 1993; Robin et al., 1996). As a result, recent accounts have begun to emphasize a more embodied process of learning to reach in contrast to the visually-guided approach that has dominated the field for several decades.

This shift toward an embodied account of learning to reach, however, leaves some questions unanswered regarding the actual role of vision in the emergence of infant reaching and particularly how vision and action map onto each other. We know that in daily, lighted surroundings, when vision is available for reaching, infants do fixate the objects (McCarty and Ashmead, 1999). They do so even weeks before reaching onset (von Hofsten, 1984). Specifically, when an object is within arm's reach and being fixated, pre-reaching infants already begin to display object-oriented changes in their arm and hand movements compared to



when they are not fixating the object (von Hofsten, 1984) or to when objects are not present (Bhat et al., 2005). Finally, despite clear evidence that infants can purposefully reach for a glowing object in the dark without seeing their hand (Clifton et al., 1991, 1993), studies have revealed an effect of vision on the formation of goal-directed movements when target or arm are occluded (Clifton et al., 1994; McCarty and Ashmead, 1999; Pogetti et al., 2013). Thus, this body of work suggests that eye and hand interact with one another well before the emergence of purposeful reaching, and continue to do so afterwards. What remains unclear is how looking at the object and bringing the hand to that location occurs at first when infants perform their initial intentional attempts to hit the target. What visuo-motor mapping process allows this to happen?

The goal of this paper is to examine anew the role of vision in relation to the emergence of goal-directed reaching in infancy, particularly in light of the more recent embodied accounts on learning to reach. We ask how do infants figure out how to map the feel of their arm to a specific location identified visually if infants' first reaching attempts are mainly controlled proprioceptively? Does vision provide any specific information in this process *prior* to reaching onset that could help tune infants' arm movements to the target location? Does proprioceptive control of the arm, from reach onset, improve such that infants become increasingly more successful and more accurate at bringing their hand toward the object area attended visually? Such scenario would be in line with our current understanding of the early process of learning to reach. It presupposes that vision is prospective and that progression in the development of reaching is a matter of learning how to improve movement control to align the movement endpoint to the visually attended target area. But could it be the other way around, that vision is mapping onto the proprioceptive movement experience of the infant? This other scenario would offer a more consistent embodied account of learning to reach by assuming that the use of vision for the control of future-oriented actions could possibly originate from infants' initial and self-produced proprioceptive movement experiences. A third scenario could also be that vision and action map onto one another in a more reciprocal fashion. This paper aims to examine these hypotheses on the developmental origins of object-directed visuo-motor mapping in infancy. We first review how previous research on infants learning to reach has addressed the question of perceptual-motor mapping. Then, we present preliminary, first-time longitudinal data on object-directed looking (captured via eye-tracking) and reaching in three infants that we followed weekly throughout the transition to reaching, to examine how the above scenarios play out. We attempt to gain insights on the process underlying the formation of visuo-motor mapping at reach onset (1) by identifying whether looking patterns at the target objects prior to the onset of reaching can help predict the formation of early goal-directed movement, (2) by tracking whether these looking patterns at the object change in the weeks following reach onset, and (3) by examining if there is some spatial correspondence between the history of looking patterns at the object and the history of point of hand-object contact after reach onset that could support one of the suggested scenarios. As we will show, these preliminary data further extend

previous embodied accounts of infants learning to reach. They suggest the possibility that mapping the feel of the hand with the sight of the object occurs by learning to align visual attention to the point of first hand-object contact, and not the reverse, as previously thought. We discuss the implication of these findings for the development of prospective control from an embodied perspective.

## LEARNING TO REACH FROM A VISUALLY-GUIDED ACCOUNT

Traditional accounts on the development of infant reaching greatly emphasized the role of vision in the process of guiding the hand toward the target. Piaget (1936/1952) was the first to describe this visually-guided process from observing his children. He reported that the emergence of reaching was elicited by the simultaneous perception of the hand and object in the same visual field. From that point, infants actively learned to match the sight of their hand to the sight of the object by coordinating two initially isolated schemes—the one for looking and the one for grasping. This combined scheme reflected a new level of functioning between vision and action, and marked the naissance of goal-directed actions. This view, that vision of the hand and target were critical for the emergence of infant reaching, was later heralded by a number of studies.

White et al. (1964) described the developmental steps leading to the emergence of visually-guided reaching by following infants longitudinally in a state hospital over their first 6 months of life. They reported several occurrences of infants alternating glances between hand and object in the months preceding reach onset. At reach onset, they noticed that these glances were used to guide the hand to the object. However, in the following weeks, they indicated that these glancing patterns dropped fairly rapidly and infants were able to lift their arm quickly from out of view to reach for the target. Assumingly, a more direct visuo-motor match had formed after a few months of visually-guided practice.

Subsequent studies recorded the kinematics of infants' reaching trajectories. They found that infants' early reaching trajectories were poorly controlled and contained many corrections and changes in direction before the hand attained the target (von Hofsten, 1979, 1991). Such indirect trajectories were interpreted as in line with the visually-guided reaching hypothesis, that vision was needed initially to actively steer the hand step-by-step closer to the target. Some studies even manipulated vision by using mirrors and displacement prisms to perturb infants' eye-hand coordination during reaching (McDonnell, 1975, 1979; Lasky, 1977). Results indicated that only older infants were affected by the mirrors/prisms. Researchers concluded that young infants did not experience a disruption in perceptual-motor coordination because they were visually monitoring their displaced hand in relation to the displaced target through the prisms, which was considered in support of the visually-guided hypothesis.

In sum, these earlier studies agreed that infants learned to reach via a top-down, visually-guided process, as if the mind was "teaching" the hand where to move in space to contact the target. Visually-guided reaching declined after months of intensive practice and gave way to visually-elicited reaching assuming a more direct spatial match between felt arm and seen object (Bushnell, 1985).

## LEARNING TO REACH FROM AN EMBODIED ACCOUNT

Today, researchers agree that learning to reach toward a wanted target is a protracted process that involves much practice over many months before infants can perform smooth and fully adapted movement patterns (Thelen et al., 1996; Konczak and Dichgans, 1997; Corbetta and Snapp-Childs, 2009). However, findings from these recent decades disagree with the premises that vision and action are separated and need to be coordinated through visual guidance in order to develop goal-directed reaching. Two lines of work contributed to this change in view.

Clifton and colleagues (Perris and Clifton, 1988; Clifton et al., 1991, 1994) found that infants can reach in the dark toward glowing or sounding objects without seeing their hand. Further, they investigated whether not seeing the hand would delay the emergence of reaching (Clifton et al., 1993). They followed infants weekly for a month prior to the onset of reaching. They found that infants who were presented with glowing objects in the dark over the weeks began to reach at approximately the same time as infants who were presented with objects in the light. This confirmed that vision of the hand was not needed to direct it to a specific spatial location even at reach onset.

The second line of studies related to trajectory formation and the circuitous hand paths typical of infants' early reaches. Thelen et al. (1993, 1996) tested four infants weekly in standard lighted conditions through the transition to reaching and subsequently throughout the end of their first year of life. They found that the initial distortions in hand trajectory were not the result of visual guidance of the hand, but rather the product of infants' inability to adequately calibrate the speed of their arm movements to the desired goal (see also Konczak et al., 1995). For example, when infants produced reaching movements with excessive speed, important motion dependent forces were generated throughout the joints and segments of the arm, which in turn acted as internal perturbations to movement coordination and contributed to drag the hand away from its intended goal. In order to counteract these disruptive forces and attain the object, infants needed to break these forces in movement and steer their hand toward the target, thus causing the observed changes in trajectory. Breaking of the movement speed and steering of the hand was not done by visual control, because infants continued to fixate on the target during this process. It was accomplished by modulating muscle forces. In subsequent weeks, as infants continued to practice reaching, they began to alter the speed of their reaching movements, suggesting that they were attempting to figure out how to calibrate their movement speed to the intended goal. This revealed an embodied learning process that involved many trials and errors, through which infants proprioceptively experienced a wide range of movements, some fast, some slow, thereby testing the dynamic boundaries of their movement in relation to the goal. Infants learned to map their intrinsic movement dynamics to the intended target goal by remembering the ones that led to good outcomes, and increasingly selecting these good solutions in the production of future attempts (Sporns and Edelman, 1993; Thelen, 1995).

These newer lines of work indicated that infants do not learn to reach via a top-down process where the mind commands the

body, but rather do so by controlling the proprioceptive feel and intrinsic dynamics of their arm movement in relation to a goal located in space. This is a deeply embodied dynamic process in which mind and body work in concert, and in which a more exact mapping between intentions and arm movement forms through repeated sensory-motor experience, producing a behavior that becomes increasingly direct and tuned to its intended goal (Chiel and Beer, 1997; Corbetta, 2009).

## THE MISSING LINK: MAPPING THE FEEL OF THE HAND WITH THE SIGHT OF THE OBJECT

What remains unclear from this prior body of work is how infants discover how to meet their intentions by mapping the proprioceptive sensations of their moving arm to a visually detected location in space. When beginning to reach, and reproducing this behavior, infants display a new intentional skill never performed before. How does looking at the object (even if performed in the dark toward a glowing object) and bringing the hand in that specific location come together in the first place?

The embodied accounts reviewed above have somewhat downplayed the critical role of vision for learning to reach despite abundant evidence indicating that visual input matters for reaching. As mentioned earlier, when infants are approaching reach onset, they fixate the target object intensely (von Hofsten, 1984, 1986). They continue to do so at reach onset and thereafter while improving arm control (Williams, 2009, 2011). Blind infants, who cannot build visual experience from birth, develop reaching at a later age (Bigelow, 1986; Troester and Brambling, 1993). Additionally, a large literature supports the prospective role of vision in the planning and execution of future-oriented actions (Jeannerod, 1988). Adult studies that used eye-tracking in the context of goal-directed movement activities have shown that the eyes usually precede the action; they aid selecting ahead of time the location of the action, but also (among other things) where and how the action should occur (Land et al., 1999; Johansson et al., 2001; Horstmann and Hoffmann, 2005; Rosander and Von Hofsten, 2011). Such prospective control of vision has been documented in infants reaching as well, for example, for identifying objects' spatial locations, (Morronegiello and Rocca, 1989), for picking up object-related information (Lockman et al., 1984; von Hofsten and Fazel-Zandy, 1984; Witherington, 2005; Berthier and Carrico, 2010), intercepting moving objects (von Hofsten, 1983; Rosengren et al., 1988), and adjusting movement in precision tasks (Carrico and Berthier, 2008; Berthier and Carrico, 2010). Vision was even found important to stimulate infants' motivation to develop active search strategies (Bojczyk and Corbetta, 2004). Such work, however, contrasts with other findings suggesting that the use of vision for movement planning and execution in infancy does not occur before 6 months of age (Berthier and Carrico, 2010) and may even continue to develop until the second year of life, especially in precision tasks (Carrico and Berthier, 2008). This raises the question of how infants learn to map the feel of their arm with the sight of the target. Indeed, it is not known if at reach onset infants control the proprioceptive feel of their arm to approximate a spatial location that is visually defined, or, if it could be the other way around, that infants map their visual attention to their

proprioceptive movement experience; or maybe even a combination of both, that is, vision and proprioception are mapping onto each other. Given the reviewed evidence, it seems critical to reevaluate the role of vision in the formation of early goal-directed movements, particularly around the emergence of reaching.

In this paper, we focus on the period around reaching onset to address three goals. Prior to reach onset, we investigate whether infants simply visually attend the location of the object without specific pattern of visual exploration of the object *per se*, or whether they already examine the shape or physical properties of object in certain ways, casting the possibility of a pre-nascent visual selective process in preparation for learning to reach. We examine how infants' object-directed visual behaviors develop following reach onset, when rapid changes in arm control are taking place. Additionally, by analyzing object-directed visual attention throughout the transition to reaching, we aim to gain new insights into the visuo-motor mapping process that underlies the emergence of infant reaching. Based on the existing literature, we see three possible scenarios that could account for how vision and action may come together when infants begin to reach for an object purposefully, for the first time. We present these scenarios first and then evaluate them against preliminary longitudinal data on the looking and reaching behaviors of three infants over an 11-week period.

## POSSIBLE SCENARIOS AND PREDICTIONS

- **Scenario 1 (or the prospective control hypothesis).** This scenario assumes that vision dictates where the hand needs to go (a sort of visually-elicited hypothesis). This would imply that infants learn to map the feel of their arm onto the sight of the object by learning to control their arm increasingly accurately to bring it where visual attention on the object is directed. Predictions from this scenario would imply that a visual selection process or particular pattern of fixations at the object could possibly form and become more observable in the weeks preceding reach onset. At reach onset and in subsequent weeks, as infants increase proprioceptive control of their arm, they would gradually figure out how to better align the transport of their hand to the vicinity of the object area visually attended. The developmental origins for such a visuo-motor mapping process can be found in the initial eye-hand coordination and pre-reaching skills of newborns (von Hofsten, 1982; van der Meer et al., 1995b), but also in their developing predictive visual abilities (von Hofsten, 1980, 2005; van der Meer et al., 1994, 1995a). Such early prospective control abilities and initial eye-hand coordination could strengthen in the weeks preceding reach onset and facilitate the transition to reaching. As object-directed vision would become more selective prior to reach onset, infants' intents toward the object could also increasingly materialize. Initially, such intention could be instantiated by the use of poorly controlled arm movements and rapid, inaccurate swipes at the target. But with rapid improvement in proprioceptive arm control infants would become increasingly capable of nearing and even matching with their hand the visually pre-selected spatial location. In this scenario, movement would align to vision.
- **Scenario 2 (or the embodied account hypothesis).** This scenario would assume that infants can take advantage of their proprioceptive experience prior to reach onset to discover how to direct their arm to a specific location. Vision would become calibrated to the action as the result of accidental events. Early in life, infants move their arms around in seemingly unintentional ways. Doing so, they can hit objects inadvertently, without seeing them, but receive immediate haptic and proprioceptive feedback about the posture and location of their arms and hands in space. Such accidental events could eventually cause the infants to direct their regard toward the location where hand-object contact occurred. Hence, vision could begin to be associated to the arm movement experience of the child as a consequence of the haptic and proprioceptive feedback received. After many repetitions of such events, a basic form of movement intentionality could emerge when the feel of an arm movement in a given direction would happen to match the sight of an object in that same direction, but the experience would still be proprioceptively-guided, not visually-elicited. As vision would continue to map onto these emergent, successful, proprioceptively controlled goal-directed movements, a more accurate selective or prospective role of vision could develop, mainly as a consequence of connecting what is seen to what is felt, not the reverse. Predictions from this scenario would entail that infants can demonstrate a relatively accurate sense of their arm projection in space at reach onset. We would expect little or no specific pre-selective looking patterns at the object in the weeks preceding the emergence of reaching. Over time, we should see increased visual attention directed toward where the hand is reaching as a result of infants discovering how to align vision to their actions and beginning to anticipate where to bring their arm in space. This sensory-motor scenario can be rooted in infants' early sense of their own body and movements (Rochat, 1998; Rochat and Morgan, 1998). Neonates respond to touch and orient themselves in a sophisticated manner from the first days of life (Rochat and Hespos, 1997). In subsequent months, infants continue to discover the action possibilities of their body through movement explorations in time, space, and self-perception that are not necessarily goal-directed. By doing so, they learn to detect contingencies between their felt actions and the interesting visual outcomes they may cause and observe (Angulo-Kinzler, 2001), including relating arm movements to object touch, even if they occurred accidentally at first and without looking directly. These early movement experiences and contingencies could contribute to the formation of new action patterns increasingly associated with a stronger and more accurate sense of limb movements in space without needing explicit visual guidance. This would be consistent with studies on infant reaching in the dark, even for unseen, auditory perceived objects (Perris and Clifton, 1988; Clifton et al., 1991). Furthermore, such scenario, where vision is being mapped onto action, would be compatible with some current embodied mind views. Through early bodily experiences, infants could form extended space representations of their actions (Borghi et al., 2013), "mesh" them with visual experiences and varied contexts, and remember them for future actions (Glenberg, 1997), and thus, develop a

cognition for action that would initially be deeply body based (Wilson, 2002).

- *Scenario 3 (co-mapping of sight and feel).* This third scenario would correspond to a mix of the two described above and would not assume any dominance of vision over proprioception (scenario 1), or proprioception over vision (scenario 2), but would rather cast the emergence of reaching as the product of a continuous process where both prospective vision and proprioceptive feel of the arm experienced in the month before reach onset become progressively integrated. Predictions should show that both sight of the object and feel of the arm are increasingly mapped onto each other, but are not related to particular visual looking trends prior to reaching onset, nor to any movement tendency after reach onset.

To examine the plausibility of these scenarios, we documented the looking (captured via eye-tracking) and reaching patterns of three infants that we began to see from the age of 2–2.5 months old (that is prior to the emergence of reaching). We followed them until they were 12 months of age, but for the purpose of this report we focus only on an 11-week period around the transition to reaching. Each week, infants were presented with 3D objects that they could visually scrutinize for up to 5 s before they would be allowed to reach for them. Infants were presented with five kinds of objects. Here we describe in detail the results related to a drumstick-shaped object (a sphere attached to the end of a rod) and contrast them with those of a plain rod with no distinct features. For each week, we report how looking patterns were distributed on the objects. When infants began to reach, we documented where they brought their hand to make the first contact with the object and related it to the looking patterns. We also compared their performance to a group of 9-month-old infants tested in the same conditions. Because 9-month-olds have more reaching experience and demonstrate decent prospective control in reaching (Lockman et al., 1984; von Hofsten and Fazel-Zandy, 1984; Piérait-Le Bonniec, 1985; Bloch, 1988; von Hofsten and Rönqvist, 1988), they constitute a good developmental norm.

METHODS

PARTICIPANTS

Eighteen infants participated in this study. Fifteen of them (6 females) were 9 months old ( $\pm 1$  week) at the time of testing. They were seen only once and their data were used in this report to provide a developmental reference norm. The other three infants (2 females) were followed longitudinally from about 2 months of age, and up to the end of their first year of life. This report presents the 11-week period around the transition

to reaching (that is, 5 weeks prior to reaching onset, the week of reaching onset, and 5 weeks following reaching onset). **Table 1** summarizes the ages (in weeks) at which we obtained useable eye-tracking data and when reach onset occurred. For infant MC, week 10 was used in replacement for missing data at week 11. Infant ME only provided useable eye-tracking data prior to reaching at week 20. Infant AC had missing data at weeks 11 and 12 prior to reach onset. All infants were recruited from the Greater Knoxville, Tennessee area (USA), via formal mailings, follow-up phone calls, or various forms of personal contact. Parents voluntarily enrolled their infants in the study and informed consent was collected for all infants. Infants were born full term, and were free of visual or motor impairments. All participating infants were White, except for one longitudinal infant who was African American. Parents were given \$10 and a photograph of their child at each visit, and received a certificate of participation.

MATERIALS

Testing sessions were completed in a well-lit room. A custom-designed infant seat reclined 10 degrees from vertical was used for infant seating. It provided full trunk support via a 15-cm-wide padded foam strap wrapped around the infants’ torso and allowed free-range arm and leg movements. A small pillow was used for the head. Before infants could support the weight of their own heads, infants were seated in their caregivers’ lap. When transitioned to the infant seat, caregivers sat nearby in another chair. Both MC and ME had already transitioned to the infant seat for collection of the data reported. AC transitioned to the seat at week 19, thus was the only infant who provided data while on her mother’s lap.

To minimize ambient distractions, a custom-designed, black, tri-fold, wooden theater was positioned directly in front of the infants (see **Figure 1A**). The theater had an opening in the center panel, precisely sized to display a black 15-inch flat-screen monitor mounted on an adjustable arm. The monitor was used for eye calibration. When the flat-screen monitor was removed from the center opening, dual layers of black curtains were positioned to conceal it. A rear curtain, always closed, provided a consistent black backdrop throughout the testing session and concealed the experimenter behind who was presenting the objects to the infants through the opening. The front curtain was opened and closed by this experimenter by using hidden strings located behind the theater in order to reveal the objects.

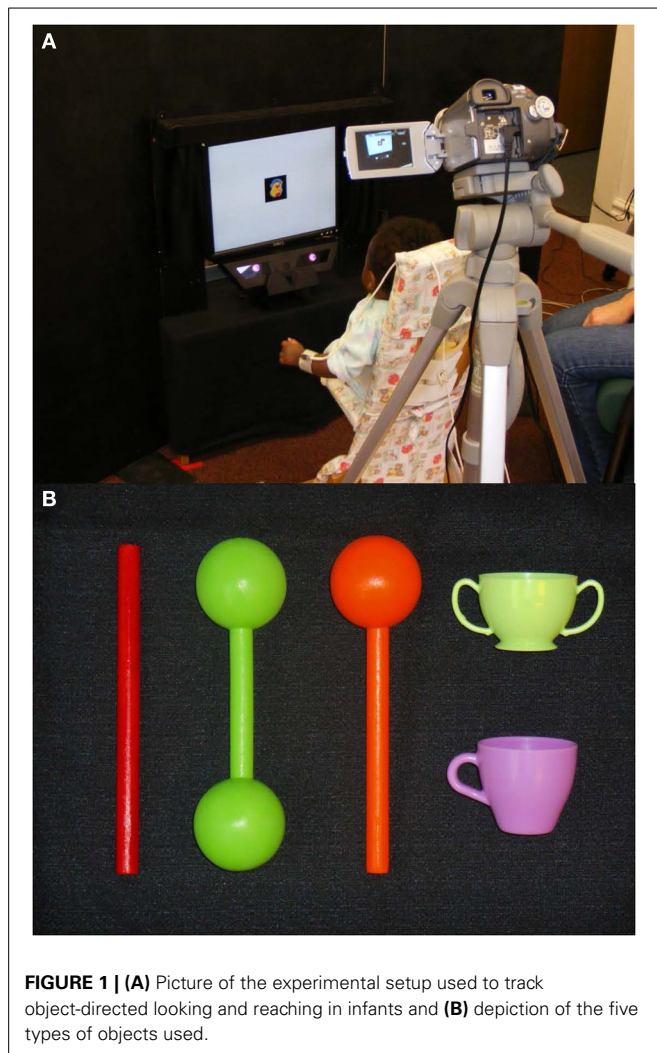
A Tobii x50 remote eye-tracker (Tobii Technology, Inc., Danderyd, Sweden) was located at the bottom of the presentation window, directly under the flat-screen monitor to capture infants’ eye movements during calibration and object presentations. The

Table 1 | Ages (in weeks) for the three longitudinal infants when tested over the 11-week period.

Infant ID	Weeks prior reach onset					Reach onset	Weeks after reach onset				
MC	10	12	13	14	15	<b>16</b>	17	18	19	20	21
ME	-	-	-	-	20	<b>21</b>	22	23	24	25	26
AC	10	-	-	13	14	<b>15</b>	16	17	18	19	20

Weeks of reach onset are marked in bold. A hyphen indicates that no useable data were collected on that particular week.





eye-tracker was positioned at a 60 cm distance from the infants' eyes and its angle was adjusted to accommodate the height of the infants' eyes (usually between 60 and 70 degrees). The eye-tracker, operated through Tobii software (Studio v. 2.0.8), used an infrared light source on the cornea relative to the center of the pupil. Estimated directions of visual fixation and saccade gaze were recorded at a rate of 50 Hz and then were superimposed onto a live video recording of the infants' visual scene, which was captured by a digital camera located directly behind the infant.

Reaching behavior was recorded with three cameras. A small, black webcam facing toward the infant and secured on top of the presentation opening recorded the infants' faces, arms, and hands. This webcam view was merged and saved with the live scene recording containing the infants' looking behaviors. Two additional video cameras were situated on the right and left sides of the infants. They were connected to a Digital Video Switcher (Datavideo Corp., Whittier, CA, USA), which merged the left and right side camera views into one split-screen arrangement and then recorded with an added image frame counter (Horita, Mission Viejo, CA, USA) on a VCR. All camera views, (side reaching cameras, scene camera, and webcam) were synchronized to

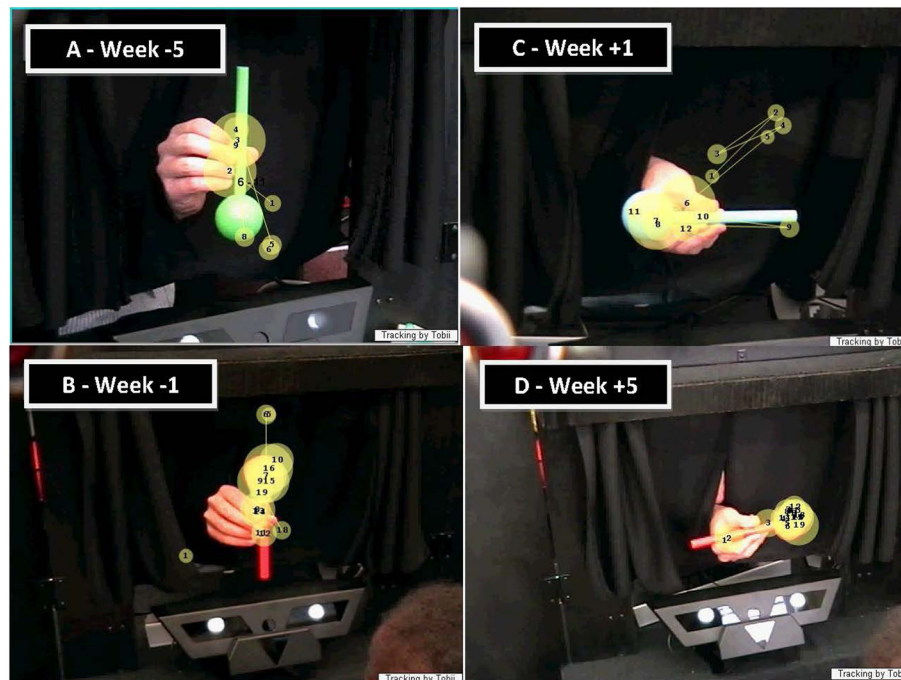
each other using a small custom-made diodes system (Corbetta et al., 2012).

Infants were offered five different types of objects (see **Figure 1B**): plain rods (18.5 cm long  $\times$  1 cm wide), drumsticks (similar plain rods, 13.5 cm long, with one 5 cm diameter sphere added to one of its ends), dumbbell-shaped objects (made of two 5 cm diameter spheres attached to each ends of a 8.5 cm long rod), small cups (5  $\times$  5 cm with one or two 3  $\times$  1.5 cm handle(s) on the side), and plain spheres (5 cm diameter). The relatively large sizes of these objects were chosen in order to elicit scanning patterns on the objects and enable us to identify if visual selection processes are at work before infants reach for the objects. Most objects were wooden and painted with solid, bright, colorful, non-toxic paint. The cups were made of solid non-toxic plastic. The solid colors ensured that infants would direct attention to the shape of objects. Due to print space constraints, preliminary data from the plain rod and drumstick objects are fully displayed in this report, results for the other objects are discussed in conclusions.

### PROCEDURE

While seated, infants were shown a Sesame Street video ([www.sesamestreet.org](http://www.sesamestreet.org)) playing on the flat-screen monitor positioned in the theater window. When the infant's attention focused on the monitor, the angle of the eye-tracker and the distance between the infant eyes and the eye-tracker were adjusted. Once the capture of the infant's pupils displayed a clear and stable signal, eye calibration using five points began. Calibration points were located at the four corners and center of the monitor. Colorful pictures of objects moving and sounding in concert were displayed consecutively in each of the five areas until the infant had looked at each location for 3–5 s. If any calibration points were missing or inaccurate for either eye, those points were repeated until eye calibration was accurate on at least four out of five points for both eyes. Occasionally, three points were used. When sounds and pictures on the monitor were not sufficient at holding infants' attention to the calibration areas, the experimenter shook small rattles in front of the target areas. Calibration typically lasted between 3 and 10 min.

After calibration, the monitor was moved out of the infants' view behind the theater, the rear curtain was placed in the back of the open window, and the front curtains were closed to hide the object presentation area. The presenting experimenter sat behind the theater and began each trial by holding an object in place at the center of the calibrated area, right in front of the rear curtain. Once the object was in place, the experimenter gave a verbal signal to a second experimenter located in an adjacent room who was running the eye-tracker. This other experimenter provided an auditory signal when gaze data collection was triggered and the presenter opened the front curtain to reveal the object (see **Figures 2A–D**). The presenting experimenter, while holding the object steadily in the calibrated window, observed the infant's live gaze on the object from the monitor behind the theater. The object was held out of the infants' reach to approximate as much as possible 5 s of active looking at the scene. Then, the presenter moved the object into the infants' reaching space and the trial ended either when the infant made contact with the object (if capable of reaching), or after a few seconds of holding the object



**FIGURE 2 | Illustration of the developmental progression of infant MC's looking patterns directed at the drumstick object.** The object is displayed in all four orientations. The numbered dots on the image indicate the points of fixations, their sizes indicate the duration of the fixations, and the lines correspond to saccades. **(A)** example of object-directed looking pattern

performed on week –5 prior to reach onset, **(B)** example of object-directed looking pattern performed on week –1 prior to reach onset, **(C)** example of object-directed looking pattern performed on week +1 following reach onset and **(D)** example of object-directed looking pattern performed on week +5 following reach onset.

in close arm range to the child (in weeks prior reach onset). If infants reached, they were given 10–15 s to continue touching the object while held by the experimenter (infants cannot grasp objects this young), after which the caregiver took the object away and placed it in a bucket behind the theater out of the infant's view. The next trial proceeded in the same manner.

All objects were presented in both horizontal and vertical orientations. The drumstick had four possible orientations with the sphere located at each one of the four cardinal points while the other objects had two. Each object and orientation were presented twice following a random order, thus the drumstick was presented up to eight times while the other objects up to four times. The same object and orientation were never presented twice consecutively.

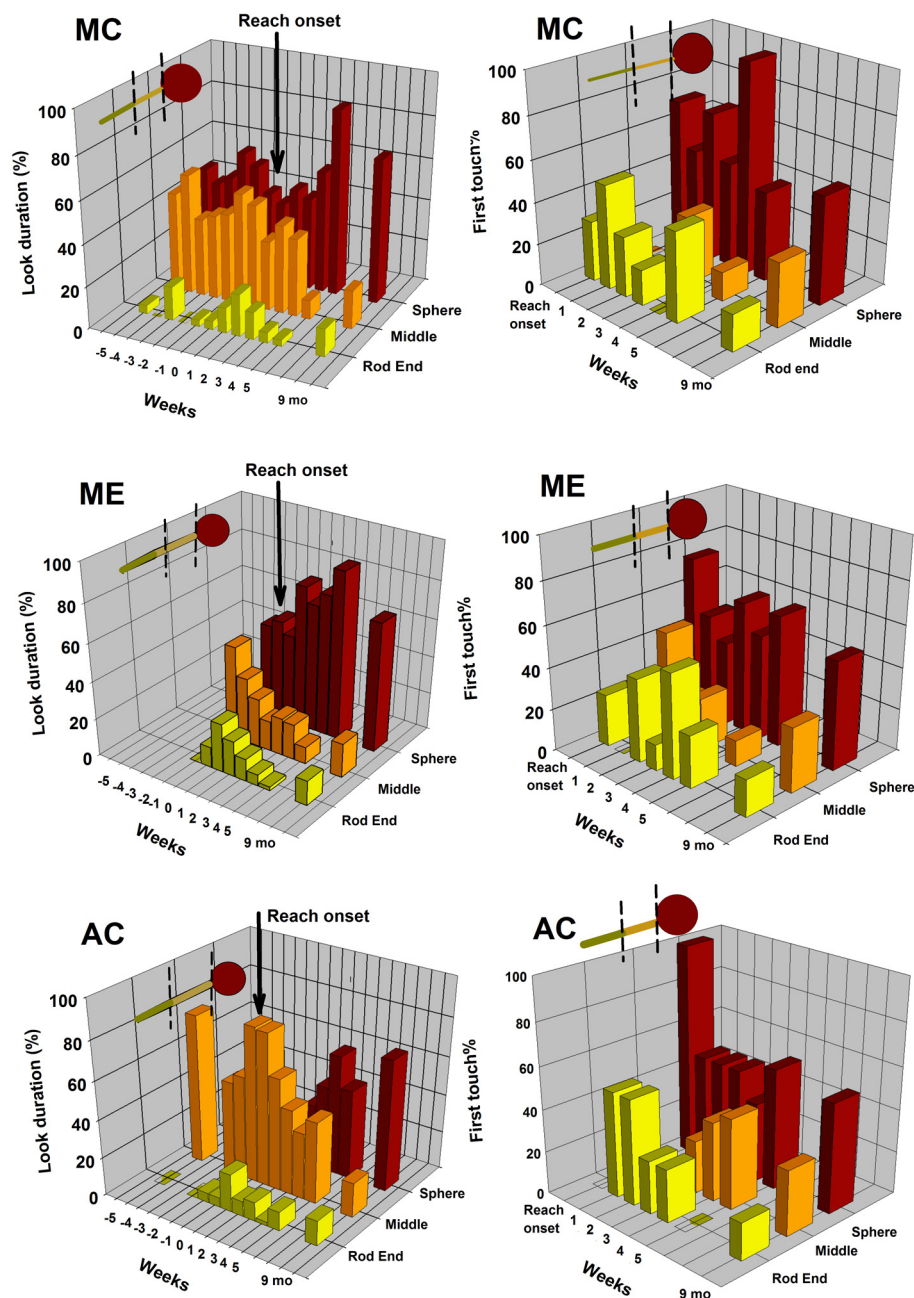
### BEHAVIORAL CODING AND ANALYSES

All reaching and looking video recordings were imported into and coded in The Observer XT, v 9.0 (Noldus Information Technology Inc., VA, USA). Coding was performed by trained independent observers who identified the onset/offset of fixation points according to predefined regions or areas of interest on the objects and also coded the point of first hand/object contact according to these same predefined object areas (see **Figures 3, 4**). The coding of looking and reaching were performed independently, in separate passes, to control for possible influences from coding one behavior as a function of the other. Coding of the looking patterns was limited to the time of object exposure in the calibrated

window from the moment the curtain opened (revealing the whole object at the center of the theater window) to the moment the presenter began moving the object into the infant's reaching space. Plain rods were divided into three equivalent areas of interest such that when presented horizontally there was a left, middle, and right region and when presented vertically there was a top, middle, and bottom section. The drumsticks' three regions corresponded to a left or right sphere, left or right rod end, and middle rod (when horizontal), or to a top or bottom sphere, top or bottom rod end, and middle rod (when vertical). Looking behavior was coded conservatively by attributing looking to a related object area only when the centers of the fixations were located on the object. Fixation centers located right on the edge of the object were still coded as object-directed fixations, but fixation centers right outside of the object border or located on the hand of the experimenter holding the object were not. We adopted this off-line coding because identifying where the center point of fixation was on the video, specifically in the area where the hand was holding the toy, was easier to determine. Moreover, if the hand holding the toy happened to move slightly, the coders could always and promptly track where the object boundaries were. Finally, the point of hand-object contact, that was coded separately, could later be exported with the looking data in the same spreadsheet. Two dependent measures were extracted from this coding:

- *Looking duration at different object regions.* Looking duration was the accumulated time infants visually attended each predefined region of the objects during each object presentation.





**FIGURE 3 | Drumstick shaped object.** The 3D bar graphs display the distributions of looking patterns on the object (graphs on the left) and distribution of first touched object area (graphs on the right) for each of the

three longitudinal infants. Data are plotted by object area (sphere, middle rod, and end rod) and by week of observation. The corresponding results for a group of 9-month-old infants are provided for comparison purposes.

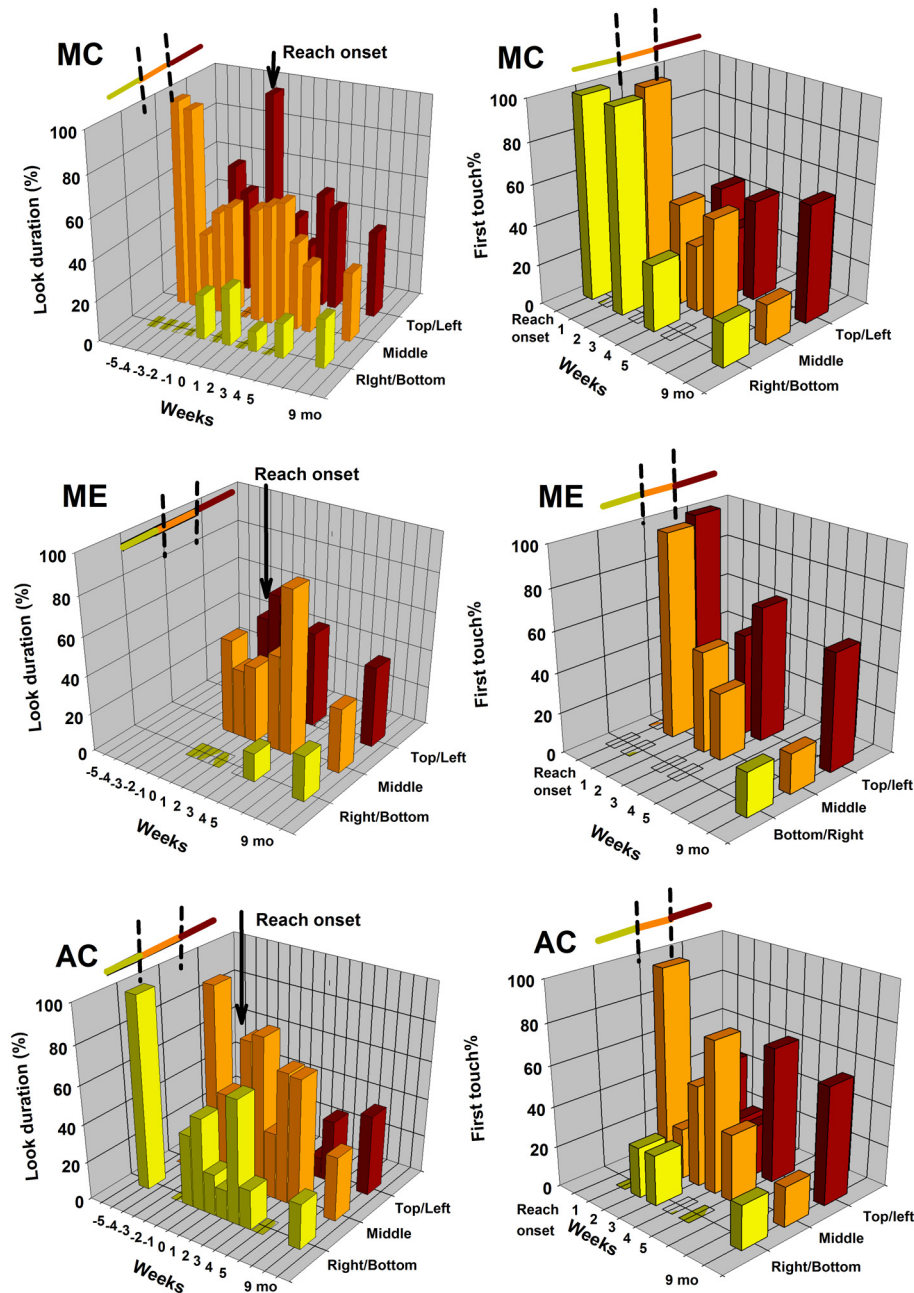
This coding excluded times when the infant looked at the hand of the experimenter holding the object and when they looked elsewhere on the scene. This duration was normalized as a function of the total looking time on the object during the trial. Inter-observer reliability performed on 20% of the data sample was 93.11% for the longitudinal infants and 91.43% for the 9-month-olds.

- *Location of first hand-object contact.* The location of the first hand-object contact corresponded to the object pre-defined

region where it occurred. Inter-observer reliability performed on 20% of the data sample was 80% and 96.7% for the longitudinal and 9-month-old infants, respectively.

#### Description of data corpus

We succeeded at collecting active looking behavior at the scene in all three longitudinal babies within the neighborhood of the 5 s targeted [average overall active looking time per trial and baby in seconds:  $MC = 5.32$  ( $SD = 2.51$ ),  $ME = 5.076$  ( $SD =$



**FIGURE 4 | Plain rod object.** The 3D bar graphs are displaying the distributions of looking patterns on the object (graphs on the left) and distribution of first touched object area (graphs on the right) for each of the three longitudinal

infants. Data are plotted by object area (bottom/right end rod, middle rod, and top/left end rod) and by week of observation. The corresponding results for a group of 9-month-old infants are provided for comparison purposes.

1.80), and  $AC = 8.21$  ( $SD = 3.09$ )). However, looking behavior was not solely directed at the object, it could be directed at the experimenter's hand holding the object or at the surrounding scene, and, in some trials, infants never looked at the object. We eliminated trials with no or minimal looking at the object, which constituted 13.25% of our data sample, and did not consider looking times that were not directed at the object (i.e., hand and surrounding). Our final data samples and average looking durations at the objects for the longitudinal

infants over the 11 weeks used in this report corresponded to:  $MC = 209$  trials, object-directed average looking time = 2.51 s ( $SD = 1.45$ ),  $ME = 105$  trials, object-directed average looking time = 2.42 s ( $SD = 1.53$ ), and  $AC = 145$  trials, object-directed average looking time = 2.76 s ( $SD = 1.57$ ). The drumsticks and rods used for this report constituted 47% of this overall sample. ME and AC produced less object-directed useable data for some weeks preceding reaching onset, which resulted in missing data for those weeks (see **Table 1**).

### Statistical analyses strategy

Statistical analyses were focused on capturing trends and developmental changes between periods before and after the onset of reaching within each infant. The strategy adopted was considered the best possible approach given the absence of statistical procedures allowing for the analysis of single subject data. This strategy accounted for the fact that our data are non-parametric normalized proportions, and that all measures are dependent. We first examined if there were predominant looking or reaching behavior at specific object areas as a function of pre- or post-reach onset using a Friedman test. If significant, we followed with pairwise Wilcoxon between object areas to determine where on the objects differences in looking and reaching resided. Development trends within pre- and post-reaching periods were assessed using linear curve estimations on the looking and reaching distributions. To approximate as much as possible an equal number of observations for the weeks prior and the weeks following reaching onset, the pre-reaching period included the 5 weeks before reach onset and the week of reach onset. The post-reaching period included the 5 weeks following reach onset and the week of reach onset. Also,

because of low power (analyses performed on 6-week periods at best), we report significance at the 0.05 level, but also *p*-values up to 0.07 level to denote trends toward significance. The 9-month-old data were not included in these longitudinal data analyses. However, we ran Mann–Whitney tests to assess whether the looking and reaching behaviors of the longitudinal infants differed from those of the 9-month-old infants.

## RESULTS

### LOOKING AND REACHING AT THE DRUMSTICK

Figure 3 displays the looking and reaching results for the drumstick-shaped object. The 3D bar graphs on the left correspond to the distributions of accumulated looking duration at this object as a function of the three pre-defined object areas (rod end, rod middle, or sphere), the week of testing (−5 to −1 = weeks prior reach onset, 0 = reach onset, 1 to 5 = weeks after reach onset), and infant (MC top graph, ME middle, and AC bottom). The corresponding 3D bar graphs on the right side of this figure display these infants' reaching distributions in relation to where they made first hand contact with the object (rod end, rod middle,

**Table 2 | *P*-values obtained from the statistical tests applied to (1) the individual distributions of accumulated looking directed to each of the three areas of the drumstick (sphere, middle rod, end rod) for the pre- and post-reaching periods, and, (2) *P*-value of the statistics applied to the individual distributions of the first hand/object contacts.**

DRUMSTICK				
Looking		Infant MC	Infant ME	Infant AC
Pre-reaching period		(weeks 10–16)	(weeks 16–21)	(weeks 10–15)
<i>N</i> (weeks)		6	2	4
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Sphere vs. middle rod vs. end rod	0.009	— <sup>†</sup>	0.039
Wilcoxon	Sphere = middle rod	ns	—	ns
	Sphere > end rod	0.028	—	0.068
	Middle rod > end rod	0.027	—	0.066
Looking		Infant MC	Infant ME	Infant AC
Post-reaching period		(weeks 16–21)	(weeks 21–26)	(weeks 15–20)
<i>N</i> (weeks)		6	6	6
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Sphere vs. middle rod vs. end rod	0.009	0.006	0.032
Wilcoxon	Sphere ≥ middle rod	ns	0.028	ns
	Sphere ≥ end rod	0.027	0.028	ns
	Middle rod ≥ end rod	0.028	ns	0.027
Reaching		Infant MC	Infant ME	Infant AC
		(weeks 16–21)	(weeks 21–26)	(weeks 15–20)
<i>N</i> (weeks)		6	6	6
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Sphere vs. middle rod vs. end rod	0.013	0.041	0.040
Wilcoxon	Sphere > middle rod	0.027	0.042	0.042
	Sphere > end rod	0.066	0.059	0.068
	Middle rod = end rod	ns	ns	ns

ns = *p*-value > 0.07; <sup>†</sup>no statistics applied for lack of data.

or sphere) from the week of reach onset (weeks 0–5). In addition, all six bar graphs display the corresponding data for the group of 9-month-olds for the purpose of comparison. On all graphs, object orientations were collapsed together.

The *p*-values of the statistical analyses performed on these longitudinal data following the strategy outlined above are presented in **Tables 2, 4**. **Table 2** shows that all the Friedman tests that were applied to each of these longitudinal looking and reaching distributional data were significant, meaning that all three infants looked and reached at this object respective pre-defined areas differentially pre- and post-reaching. Wilcoxon tests revealed the following trends. For the *pre-reaching looking period*, both MC and AC divided their object-directed visual attention mainly between the sphere and the middle of the rod. Their amounts of looking at those two areas were significantly greater than at the end of the rod. No test was ran on ME's pre-reaching looking period due to only 2 weeks of useable data up to reach onset. AC's *p*-values for that period were nearing significance. For the *post-reaching looking period*, visual attention to the drumstick was

still mainly directed toward the sphere area, middle rod, or both depending on the child. MC's and AC's looking patterns were still mainly distributed between sphere and middle rod, while ME's visual attention was mainly directed to the sphere. Wilcoxon tests performed on the *reaching patterns* indicated a significant bias toward more frequent first touches at the sphere area. All three babies directed their hand and made first contact more frequently with the sphere than the middle rod (significant trend), and end rod (nearing trend). There were no differences in frequency of first touches between middle and end rod areas.

Developmental trends in *looking behavior* assessed with linear curve estimations (**Table 4**) only revealed significant changes over time for the post-reaching periods. All three babies did not change looking behavior before reach onset, however, following reach onset, all three babies similarly significantly increased amount of looking at the sphere, while significantly decreasing amount of looking at the middle of the rod. No developmental trends were detected for looking at the end of the rod; this object area continued to be poorly visually attended even after reaching onset.

**Table 3 | *P*-values obtained from the statistical tests applied to (1) the individual distributions of accumulated looking directed to each of the three areas of the plain rod (top/left, middle rod, right/bottom rod) for the pre- and post-reaching periods, and, (2) *P*-value of the statistics applied to the individual distributions of the first hand/object contacts.**

PLAIN ROD				
Looking		Infant MC	Infant ME	Infant AC
Pre-reaching period		(weeks 10–16)	(weeks 16–21)	(weeks 10–15)
<i>N</i> (weeks)		6	2	4
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Top/left vs. middle vs. right/bottom	0.027	— <sup>†</sup>	ns
Wilcoxon	Top/left vs. middle	ns	—	—
	Top/left vs. right/bottom	0.042	—	—
	Middle vs. right/bottom	0.043	—	—
Looking		Infant MC	Infant ME	Infant AC
Post-reaching period		(weeks 16–21)	(weeks 21–26)	(weeks 15–20)
<i>N</i> (weeks)		6	5	6
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Top/left vs. middle vs. right/bottom	0.070	0.069	0.030
Wilcoxon	Top/left vs. middle	—	—	0.028
	Top/left vs. right/bottom	—	—	ns
	Middle vs. right/bottom	—	—	ns
Reaching		Infant MC	Infant ME	Infant AC
		(weeks 16–21)	(weeks 21–26)	(weeks 15–20)
<i>N</i> (weeks)		6	4	5
Statistical test	Toy areas compared	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Friedman	Top/left vs. middle vs. right/bottom	ns	ns	ns
Wilcoxon	Top/left vs. middle	—	—	—
	Top/left vs. right/bottom	—	—	—
	Middle vs. right/bottom	—	—	—

ns = *p*-value > 0.07; <sup>†</sup>no statistics applied for lack of data.

Interestingly, by week 5 after reach onset, the looking patterns distributions at the drumstick in those three infants closely approximated the looking patterns distribution of the 9-month-old group. This older group displayed significantly longer looks at the sphere (Friedman,  $p < 0.0001$ ). The linear curve estimations performed on the *reaching* data did not reveal consistent significant developmental trends across babies, except for AC who increased her object contacts at the middle. For all three babies, the predominant tendency to touch the sphere more frequently remained about the same over the 6 weeks post-reaching period. The 9-month-old infants also displayed significantly more first touches at the sphere (Friedman,  $p < 0.027$ ).

To compare the looking and reaching trends of the longitudinal infants with those of the 9-month-old group, we collapsed the 9-week period (3 weeks before and 6 weeks after reach onset) into three 3-week periods' averages corresponding to: prior to reach onset for looking only (we used week 20 for ME), right after reach onset, and the last 3 weeks post-reaching for looking and reaching. For *looking behavior*, the developmental trends described above were confirmed. The Mann–Whitney tests revealed significant differences between longitudinal and 9-month-old infants for looking at the middle and the sphere areas of the drumstick in the weeks preceding and just following reach onset (sphere prior

reach onset,  $p < 0.021$ , sphere at reach onset,  $p < 0.038$ , middle rod prior reach onset,  $p < 0.011$ , middle rod at reach onset,  $p < 0.028$ , all two-tailed). However, those group differences were no longer significant for the last 3 weeks post-reaching (sphere post-reach,  $p = 0.628$ , middle rod post-reach,  $p = 0.173$ , two-tailed). There were no significant differences between groups for looking at the rod end of the drumstick (all two-tailed  $p$ 's  $> 0.374$ ). Thus, the longitudinal infants looking patterns at the drumstick, initially different from those of the 9-month-old infants prior to and right around reach onset, became increasingly more similar to those of the 9-month-olds by week 4–6 after reach onset. For *reaching behavior*, there were no significant differences between groups (all two-tailed  $p$ 's  $> 0.107$ ), verifying that distribution of the point of first hand/object contact of the longitudinal infants did not differ from those of the 9-month-old group.

In sum, all three longitudinal infants looked and reached at the drumstick differentially over the 11-week period, however, developmental change over time was only observed in relation to the looking pattern performed after reach onset. Infants increased their visual attention toward the sphere, and in a 6-week span, approximated the distributional looking pattern displayed by 9-month-old infants. Interestingly, for reaching, more frequent first contacts at the sphere were present from the week of reach

**Table 4 | Developmental trends in looking distribution and first hand/object contact over the 6 weeks up to reach onset and 6 weeks from reach onset.**

Looking	Drumstick			Plain rod		
	Infant MC (weeks 10–16)	Infant ME (weeks 16–21)	Infant AC (weeks 10–15)	Infant MC (weeks 10–16)	Infant ME (weeks 16–21)	Infant AC (weeks 10–15)
<b>Pre-reaching period</b>						
<i>N</i> (weeks)	6	2	4	6	2	4
Area of look duration	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Sphere/top-left	ns	– <sup>†</sup>	ns	ns	– <sup>†</sup>	ns
Middle rod	ns	–	ns	0.026	–	ns
End rod/bottom-right	ns	–	ns	ns	–	ns
<b>Looking</b>	<b>Infant MC</b>	<b>Infant ME</b>	<b>Infant AC</b>	<b>Infant MC</b>	<b>Infant ME</b>	<b>Infant AC</b>
<b>Post-reaching period</b>	<b>(weeks 16–21)</b>	<b>(weeks 21–26)</b>	<b>(weeks 15–20)</b>	<b>(weeks 16–21)</b>	<b>(weeks 21–26)</b>	<b>(weeks 15–20)</b>
<i>N</i> (weeks)	6	6	6	6	5	6
Area of look duration	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Sphere/top-Left	0.047	0.023	0.033	ns	ns	ns
Middle rod	0.035	0.019	0.009	ns	ns	ns
End rod/bottom-right	ns	ns	ns	ns	ns	ns
<b>Reaching</b>	<b>Infant MC</b>	<b>Infant ME</b>	<b>Infant AC</b>	<b>Infant MC</b>	<b>Infant ME</b>	<b>Infant AC</b>
	<b>(weeks 16–21)</b>	<b>(weeks 21–26)</b>	<b>(weeks 15–20)</b>	<b>(weeks 16–21)</b>	<b>(weeks 21–26)</b>	<b>(weeks 15–20)</b>
<i>N</i> (weeks)	6	6	4	6	4	5
Area of look duration	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Sphere/top-left	ns	ns	ns	0.032	ns	ns
Middle rod	ns	ns	0.006	ns	ns	ns
End rod/bottom-right	ns	ns	ns	ns	– <sup>‡</sup>	ns

*P*-values from linear trend testing are displayed by object area and by infants. <sup>†</sup>no statistics applied for lack of data; <sup>‡</sup>no look at the bottom/right end rod.



onset. This reaching trend was maintained over the 6 weeks of reaching and was similar to the distribution of points of contacts displayed by the 9-month-old group.

### LOOKING AND REACHING AT THE PLAIN ROD

**Figure 4** displays the looking and reaching results for the plain rod. As for the drumstick data, the 3D bar graphs on the left correspond to the distributions of accumulated looking duration at this object as a function of the three pre-defined object areas (each end and middle areas). Since there was no specific shape asymmetry to this object, we arbitrarily collapsed the vertical and horizontal presentation trials by merging the amount of looking at the top with the amount of looking at the left, and amount of looking at the bottom with the amount of looking at the right. As for **Figure 3**, distribution of looking (left graphs) and reaching (right graphs) per object areas are displayed similarly as a function of the week of testing, and by infant following the same order. Again, all six bar graphs display the corresponding looking and reaching data for the group of 9-month-old infants for this same object.

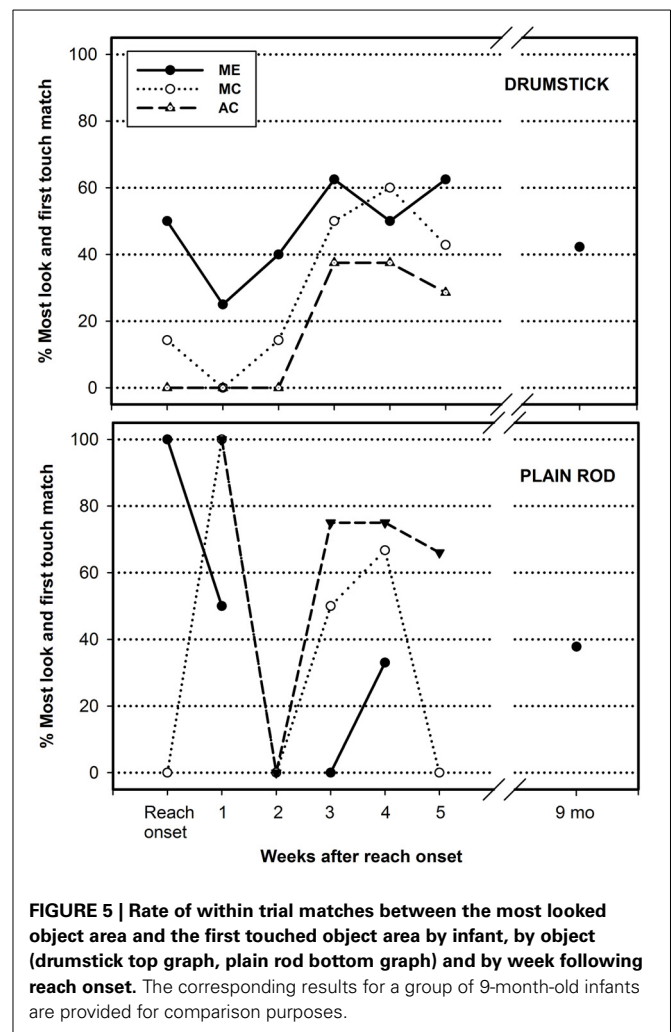
The corresponding *p*-values of the statistical analyses performed on these longitudinal data are presented in **Tables 3, 4**. **Table 3** reveals that very few Friedman tests applied to these individual looking and reaching distributional data reached significance. Thus, as a whole, infants did not display consistent preferred looking biases for the plain rod in the periods preceding and following reach onset, neither did they display consistent biases in reaching. With the exception of MC for the pre-reaching period, and AC for the post-reaching period, infants seemed to have more week-to-week fluctuating looking and fluctuating reaching distributions on the rod with no specific object areas attracting consistently greater looking or reaching behaviors.

Likewise, **Table 4** shows that the linear curve estimations applied to these data revealed almost no developmental trends over the 6-week periods preceding and following the emergence of reaching. The only significant linear change over time observed in looking was for infant MC, who reduced visual attention to the middle of the rod during the pre-reaching period. Also, MC was the only one to display a significant linear change in reaching; she increased her amount of first hand contact with the top/left of the rod over the 5 weeks following reach onset.

In sum, compared to the drumstick that yielded looking and reaching patterns that seemed to gravitate predominantly toward the sphere in all three infants, the plain rod seemed to entice more random trends. Note that for the 9-month-old infants, looking patterns on the rod were also more distributed across all three object pre-defined areas (Friedman,  $p = 0.891$ ). Reaching, in that older age group was biased toward the top/left rod area (Friedman,  $p < 0.027$ ). A similar trend can be seen in the longitudinal infants, although it is present only for week 5 post-reaching. None of the group comparisons between longitudinal and 9-month-old infants revealed significant difference in looking and reaching behavior for any of the object areas and any of the collapsed 3-week periods.

### VISUAL-MOTOR MAPPING FOLLOWING THE ONSET OF REACHING

The data presented above reflected changes in looking and reaching behaviors independently. To address the question of the



**FIGURE 5 | Rate of within trial matches between the most looked object area and the first touched object area by infant, by object (drumstick top graph, plain rod bottom graph) and by week following reach onset.** The corresponding results for a group of 9-month-old infants are provided for comparison purposes.

mapping between the feel of the hand and the sight of the object, looking and reaching behaviors needed to be linked to each other. To address this, we performed a trial-by-trial analysis to examine whether there was a direct spatial correspondence between the areas of the object visually attended the most (the most looked area) and the location where the hand made the first contact with the object (area of first touch). The number of trials corresponding to a direct spatial match between the most looked area and the area of first hand contact were normalized as a function of the total number of trials collected for a given object. These data are reported in **Figure 5** by week from reach onset, and for each longitudinal infant separately, drumstick on top and plain rod at the bottom. These same data for the 9-month-old group are displayed on these graphs for the purpose of comparison.

We performed a linear curve estimation on these data to assess changes in the rate of spatial look-reach match over time. For the drumstick, the rate of matching between looking and reaching revealed a 2 to 3-fold increase over the observed 6-week period, but reached significance only for AC ( $p < 0.048$ ). It neared significance for MC ( $p = 0.055$ ) and was not significant for ME ( $p = 0.209$ ). For the plain rod, there was no significant developmental trend observed. Mann-Whitney tests comparing the 3-week



averages of the longitudinal infants with the data of 9 months old group revealed a nearing group difference ( $p = 0.065$ ) for the first 3 weeks following reach onset in the drumstick. All other comparisons (last 3 weeks for drumstick and plain rod group comparisons) were not significant (all  $p$ 's  $> 0.244$ ).

The last analysis assessed where on the object the look-reach match occurred to determine if visuo-motor matches occurred randomly on any areas of the object or if they were focalized to one or two specific areas of the objects. To do so, we considered only the trials that yielded a look-reach spatial match. For the drumstick, 88% of the look-reach matches documented over the 6-week period were aimed at the sphere area of the object, the remaining 12% occurred at the middle rod area. There were no look-reach matches corresponding to the end of the rod for this object. For the plain rod, 77% of the matches occurred at the middle of the rod, the remaining 23% percent was spread at either end areas of the rod. These numbers are combining all three infants and all weeks. These trends indicate that when spatial matches occurred between looking and reaching, they did not occur randomly on either area of the object, but were mainly focused on the sphere area for the drumstick and the middle area for the plain rod. These trends were present from the first weeks of reaching in all three infants and lasted up to the last week of reaching reported.

## DISCUSSION

The goal of this paper was to begin evaluating how infants map the feel of their arm with the sight of the object when they perform their first goal-directed reaching movements. To examine this mapping process, we tracked the object-directed looking behavior (via eye-tracking) and point of first hand-object contact in three infants that we followed weekly over an 11-week period throughout the transition to reaching. With these data, we evaluated three possible scenarios that offered different levels of balance between the respective roles that vision and proprioception could have at the emergence of infant's reaching. The prospective control hypothesis assumed a more predominant role of vision over proprioception, the embodied hypothesis assumed a more predominant role of proprioception or sensory-motor experience over vision, and finally, a more mutually balanced contribution of both vision and proprioception acting in concert was considered as a third possible hypothesis. For each hypothesis, we made specific predictions. Here we discuss our findings against these predictions.

- *Scenario 1 (or the prospective control hypothesis)*. This scenario assumed that vision would develop its prospective control role prior to reach onset. Reaching would form as a result of the infants increasingly figuring out how to proprioceptively guide their hand toward the area where visual attention is directed. Predictions consistent with this scenario were that (1) vision would reveal specific selective looking trends at the objects in the weeks preceding reach onset, (2) that after reach onset the looking trends would persist, and (3) that change over time would occur in the reaching behavior as a result of the more successful spatial alignment (or mapping) of the action end-point onto the visually selected object area. For the drumstick,

we observed a looking bias prior to reach onset directed toward two object areas—the sphere and the middle rod areas. This looking bias grew more specific in the direction of the sphere after reach onset, while all three infants maintained a relatively steady reaching bias at the sphere from reach onset and thereafter. Thus, for this object, a developmental change was observed in the looking behavior following reach onset but not in the reaching behavior *per se*, which is inconsistent with this scenario's predictions. Similar trends were not observed for the plain rod. In fact, very few significant results were reported for this plain rod object, suggesting that object shape may have interacted with this perceptual-motor mapping process, a point discussed further below, also in relation to the other objects that we did not present.

- *Scenario 2 (or the embodied account hypothesis)*. This scenario assumed that infants build an extended, proprioceptive based sense of where to move their arm in space through initially undirected, nonetheless active movement experiences prior to reach onset. As infants happen to touch objects in their surroundings, at first accidentally (and without looking), they discover how to connect their visual attention to the consequential proprioceptive and haptic feel of their arm and hand in the environment. This would allow the formation of some kind of proprioceptive knowledge of how to direct arm movements in space by associating spatial vision to movement perception, leading to reach onset. The predictions we made in relation to this scenario were that infants (1) would not necessarily show specific object-related visual trends prior to reach onset, but (2) would demonstrate a more accurate/consistent spatial aiming ability in their reaching movement from reach onset, as a result of their acquired proprioceptive spatial movement experience in the weeks preceding reach onset. We also expected a change in visual attention toward the touched area (rather than a change in movement aiming toward the looked area) after reach onset as a product of successful reaching. The drumstick results were in line with these predictions. All three infants displayed a greater first touch trend at the sphere from reach onset. They maintained this aiming trend over the post-reaching weeks. Developmental change occurred in the looking pattern after reach onset. Over the post-reaching weeks, visual attention increasingly shifted toward the sphere, the preferred first touch area. Again, similar trends were not observed for the plain rod.
- *Scenario 3 (co-mapping of sight and feel)*. This scenario assumed that both vision and action would map onto each other following reach onset. With this scenario, we did not predict a dominant trend in looking prior to reaching nor did we predict a dominant trend in reaching from reach onset. Rather, we expected initial random looking and reaching patterns that would reciprocally map onto one another over time. Both vision and reaching would adjust and morph into a more precise perceptual-motor mapping over time. Infants would increasingly try to look to where they aim, while trying to aim to where they look at the same time. None of our data revealed concomitant changes in both looking and reaching over the observed developmental period to support this third scenario.

## POSSIBLE IMPLICATIONS FOR THE PROCESS OF LEARNING TO REACH AND THE DEVELOPMENT OF PROSPECTIVE CONTROL IN INFANCY

The looking and reaching behaviors documented through the transition of reaching in those three infants seem to point toward the embodied hypothesis, but as noted, this was only for the drumstick-shaped object; these findings did not extend to the plain rod object. Here, we discuss these results in relation to the other objects that we have not presented and evaluate the significance and limitations of these preliminary findings for our understanding of visuo-motor mapping in infants learning to reach.

We found the trend reported for the drumstick at first provocative. The long held belief in the infant reaching literature (even from prior embodied accounts) has been that infants have poor control of their arm at reach onset, hence the documented indirect trajectories (Thelen et al., 1993, 1996; von Hofsten, 1991). As a result, assumptions were that, from reach onset, infants first learned how to bring their arm more successfully in contact with the object and only after extensive practice, did they learn to refine their arm control to direct their hand more accurately and more smoothly toward the target taking into account its physical characteristics (Lockman et al., 1984; von Hofsten and Fazel-Zandy, 1984). According to such assumptions, we would have expected more random points of first hand-object contacts for any object following reach onset. The fact that, for the drumstick (and for the other objects also, as we will see), all three infants succeeded touching the sphere more predominantly from the beginning, suggests that infants are somewhat capable of aiming their movement in space more accurately than thought before. This result is particularly striking given that the sphere orientation was randomly presented in one of four possible cardinal locations. Thus, to touch the sphere first more often and from the first week of reaching, the infants had to have developed some basic control ability to direct their arm to those different locations in space. Related to this finding was the fact that the observed increase in look-reach match also occurred predominantly at the sphere area, not at other object areas, and that this match increase seemed to result more from an augmented visual attention toward the sphere across weeks, not a change in touch rate at the sphere.

If we think a little more about this result on reaching accuracy, we realize that it may not be so unexpected after all. Prior studies on infant reaching have typically used small objects for reaching and shown that at reach onset infants can hit such smaller objects. Thus, in a way, prior studies have already demonstrated that infants are capable of some spatial movement accuracy. However, this ability never came into clear focus, possibly because no studies had observed how infants could begin reaching for larger objects offering choices in points of contact.

We also think that the object shapes and spatial arrangement of their distinct features mattered in driving the responses observed. When spheres or larger parts were present (as in the drumsticks, dumbbells, or cups), the looking and reaching responses were more skewed toward the sphere(s) or cup bowl. Skewed responses appeared stronger when the bigger part of the object was one, as in the drumsticks or cups. For the cups, for instance, looking and reaching were heavily directed at the bowl of the cup, not the handle(s). This trend for the cups was also present

in the 9-month-old group (there were no significant differences between groups). Shape features also seemed to engender more developmental changes (as we saw for the drumstick). For example, for the dumbbell-shaped object, two out of the three infants (MC and ME) displayed growing visual attention toward the two spheres located at each end of the rod in the post-reaching weeks compared to the pre-reaching weeks. This developmental change was not as strong as the one reported for the drumstick shaped object due to the fact that, for the dumbbell, visual attention was being increasingly split between two sphere locations (instead of one as in the drumstick). For MC, the sphere/middle rod/sphere looking distributions for the dumbbell went from 24/76/0 on week -5 to 48/13/39% on week 5. For ME, the pattern distribution was 66/34/00 at week -1 and was 45/0/55% at week 5. The looking distributions at week 5 for the dumbbell were not significantly different from those of the 9-month-old group (41/18/41%). Reaching, on the other hand, was already directed toward one of the two spheres more frequently from reach onset (MC sphere/middle/sphere percent reaching = 71/2/27% at reach onset and 25/0/75% at week 5 post-onset, ME sphere/middle/sphere percent reaching = 6/35/59% at reach onset and 75/0/25% at week 5 post-onset; 9-month-olds = 46/18/36%). Finally, as a result of two visually attended areas, but only one touched area, the rate of look-reach match for the dumbbell was not showing as a consistent progression over time as reported for the drumstick with one sphere. But this low rate of look-reach match was not void of trends. We reported that for the drumstick, when matches between looking and reaching occurred, they occurred in great majority at the sphere location. For the plain rods, even though there was no strong, consistent progression between looking and reaching matches, when matches occurred, they happened at the middle of the rod area. The same area of match consistency was found for the dumbbell and cup objects. For the dumbbell, when look and reach spatially matched, they occurred 80% of the time on one of the two spheres (only 20% of the matches were performed in the middle rod area), and for the cups, 94% of the look-reach matches occurred at the bowl area (again, these trends are consistent with what we observed with the 9-month-olds). Thus, from all of the above, it appears that object shape drove infants' reaching responses and visual attention differentially, otherwise we would not have obtained such response trends and regularities within and across objects. Furthermore, a steady reaching trend from reach onset was observed for nearly all objects when distinct shape features were present, while it was not always the case for looking.

We also did not expect infants showing points of object contact so similar to those of the 9-month-olds right from reach onset. And we did not expect infants' looking patterns to change so quickly to resemble those of the 9-month-olds in just a few weeks. These findings were surprising but also suggest clues to our understanding of the process of visuo-motor mapping at reach onset.

First, we think that these data show that from reach onset infants can project their hand toward a future location in space successfully and can display a certain level of endpoint accuracy, similar to those of more experienced reachers. This supports the

interpretation that, infants must have developed some kind of proprioceptive spatial knowledge of their arm movement prior to learning to reach. As discussed earlier, such motor knowledge could have formed as a consequence of accidental events involving the arm and the eyes, which possibly created useful visuo-motor contingencies that helped the development of an extended sense of the arm movement in space (see also Borghi et al., 2013 on the concept of extended embodied mind). We know that blind infants who cannot use vision to spatially calibrate their actions in space are delayed in their development of goal-directed skills (Bigelow, 1986), while infants who have had visual experience of the world prior to reach onset can begin reaching in the dark, even at first with minimal visual information (Clifton et al., 1993). We think that while the formers may be deprived from building such extended proprioceptive mapping of their actions in space as a result of their lack of vision, the latter benefit from being able to associate visual experience to their movement experience, thus, allowing them to reach successfully in the dark in response to a seen target, but also in response to auditory cues (Clifton et al., 1991). Clearly, given our procedure, and the fact that we gave time to the infants to look at the objects prior to allowing them to reach for them, we cannot rule out that by doing so, we may have enhanced their visual attention to the objects, which could possibly explain the selective object-related responses we observed. It is possible, that with such object attention enhancement, we allowed infants to consider the spatial properties of the objects more fully than if they were not given time to look at the object prior to aiming for them. Obviously, the present results need to be substantiated on a larger sample and extended to other task contexts to fully understand the underlying processes of early perceptual-motor matching. But, the fact that all three infants in our setup displayed similar trends on many of these measures is remarkable in our opinion.

Second, the fact that for the drumstick the greatest point of hand-object contact seemed more consistent over the weeks, while the point of greatest visual attention grew to align with the point of hand-object contact over the weeks, suggests that vision may not have been the main driving factor in setting initial motor goal accuracy. However, as spatial mapping between vision and action strengthened over the weeks, vision may have become more predictive in defining the point of where to bring the hand in contact with the object. Again, this was suggested by the progressive alignment of vision onto the preferred contacted area for the drumstick. Such alignment could reflect an increasing ability of vision to become more selective and more predictive of where the hand is being directed as motor and visual spatial outcomes are being paired repeatedly during early reaching responses. The prospective role of vision could originate from these infants' initial embodied reaching experiences. Indeed, it could be possible that infants' movement experience and associated resulting action's outcomes drove the needs of vision to begin detecting ahead of time where the hand should go. In the case of reaching for the drumstick, infants could have learned to direct their visual attention increasingly toward the sphere area, perhaps because it met some valued outcome. For example, infants may have preferred to touch the sphere because it provided greater haptic experience than the thinner rod to which the sphere was attached.

As infants gained experience at reaching and touching, vision became increasingly attuned to these features and began performing more searches for these special features, thereby becoming more selective and predictive for reaching. Such interpretation is consistent with a number of studies on infants' self-produced actions and their understanding of actions in the physical and social world that suggest, in similar ways, that infants' active experiences can drive changes in their attention and perception of the world (Cicchino and Rakison, 2008; Rakison and Krogh, 2012). Another study also found that infants' observational experience of others' actions does not lead to the same understanding as when acting themselves (Gerson and Woodward, 2014). Thus, findings from these studies are consistent with our stand that vision alone, may initially not provide the best source of information in the context of goal-directed actions, but experience acquired through early sensory-motor activity may foster a discovery and understanding of the world that could eventually translate into a more cognitive or visual knowledge of the world (see also, Campos et al., 2000 on motor activity and mind).

Future studies are necessary to extend our observations to more infants and wider contexts to examine the validity of the embodied scenario we propose. Most useful, we think, will be studies examining vision during the movement of reaching itself, something we did not do in our longitudinal observations. Such observations will be essential to disentangle the respective role of vision and arm control in infants' first reaching attempts. Prior evidence, in 9-month-old infants, where the recording of infants' eye-movements directed to a target were paired with the arm movement kinematics corresponding to reaching for that same object, pointed to the production of object-specific looking patterns closely matching movement corrections toward that object (Corbetta et al., 2012). It is unknown whether infants at reach onset can perform such eye-hand corrections during movement. Detecting whether such on-line attentional patterns and movement corrections also occur in young infants at reach onset will be important to continue to understand how infants discover how to map the feel of their arm with the sight of the object.

## REFERENCES

- Angulo-Kinzler, R. M. (2001). Exploration and selection of intralimb coordination patterns in 3-month-old infants. *J. Mot. Behav.* 33, 363–376. doi: 10.1080/00222890109601920
- Berthier, N. E., and Carrico, R. L. (2010). Visual information and object size in infant reaching. *Infant Behav. Dev.* 33, 555–566. doi: 10.1016/j.infbeh.2010.07.007
- Bhat, A., Heathcock, J., and Galloway, J. C. (2005). Object-oriented changes in hand and joint kinematics during the emergence of purposeful reaching. *Infant Behav. Dev.* 28, 445–465. doi: 10.1016/j.infbeh.2005.03.001
- Bigelow, A. E. (1986). The development of reaching in blind children. *Br. J. Dev. Psychol.* 4, 355–366. doi: 10.1111/j.2044-835X.1986.tb01031.x
- Bloch, H. (1988). Early coordinations and visual anticipation. *Infant Behav. Dev.* 11, 365–368.
- Bojczyk, K. E., and Corbetta, D. (2004). Object retrieval in the 1st year of life: learning effects of task exposure and box transparency. *Dev. Psychol.* 40, 54–66. doi: 10.1037/0012-1649.40.1.54
- Borghi, A. M., Scorolli, C., Caligiore, D., Baldassare, G., and Tummolini, L. (2013). The embodied mind extended: using words as social tools. *Front. Psychol.* 4:214. doi: 10.3389/fpsyg.2013.00214
- Bushnell, E. W. (1985). The decline of visually guided reaching during infancy. *Infant Behav. Dev.* 8, 139–155. doi: 10.1016/S0163-6383(85)80002-3

- Bushnell, E. W., and Boudreau, J. P. (1993). Motor development and the mind: the potential role of motor abilities as a determinant of aspects of perceptual development. *Child Dev.* 64, 1005–1021. doi: 10.2307/1131323
- Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., and Witherington, D. (2000). Travel broadens the mind. *Infancy* 1, 149–219. doi: 10.1207/S15327078IN0102\_1
- Carrico, R. L., and Berthier, N. E. (2008). Vision and precision reaching in 15-month-old infants. *Infant Behav. Dev.* 31, 62–70. doi: 10.1016/j.infbeh.2007.07.005
- Chiel, H. J., and Beer, R. D. (1997). The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends Neurosci.* 20, 553–557. doi: 10.1016/S0166-2236(97)01149-1
- Cicchino, J. B., and Rakison, D. H. (2008). Producing and processing self-propelled motion in infancy. *Dev. Psychol.* 44, 1232–1241. doi: 10.1037/a0012619
- Clifton, R. K., Muir, D. W., Ashmead, D. H., and Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Dev.* 64, 1099–1110. doi: 10.2307/1131328
- Clifton, R. K., Rochat, P., Litovsky, R. Y., and Perris, E. E. (1991). Object representation guides infants' reaching in the dark. *J. Exp. Psychol. Hum. Percept. Perform.* 17, 323–329. doi: 10.1037/0096-1523.17.2.323
- Clifton, R. K., Rochat, P., Robin, D. J., and Berthier, N. E. (1994). Multimodal perception in the control of infant reaching. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 876–886. doi: 10.1037/0096-1523.20.4.876
- Corbetta, D. (2009). "Brain, body, and mind: lessons from infant motor development," in *Toward a Unified Theory of Development: Connectionism and Dynamic Systems Theory Re-Considered*, eds J. P. Spencer, M. Thomas, and J. McClelland (New York, NY: Oxford University Press), 51–66.
- Corbetta, D., Guan, Y., and Williams, J. L. (2012). Infant eye-tracking in the context of goal-directed actions. *Infancy* 17, 102–125. doi: 10.1111/j.1532-7078.2011.00093.x
- Corbetta, D., and Snapp-Childs, W. (2009). Seeing and touching: the role of sensory-motor experience on the development of infant reaching. *Infant Behav. Dev.* 32, 44–58. doi: 10.1016/j.infbeh.2008.10.004
- Gerson, S. A., and Woodward, A. L. (2014). Learning from their own actions: the unique effect of producing actions on infants' action understanding. *Child Dev.* 85, 264–277. doi: 10.1111/cdev.12115
- Gesell, A., and Amatruda, C. (1946). *Developmental Diagnosis*. London: Hamilton.
- Glenberg, A. M. (1997). What memory is for. *Behav. Brain Sci.* 20, 1–55.
- Halverson, H. M. (1931). An experimental study of prehension in infants by means of systematic cinema records. *Genet. Psychol. Monogr.* 10, 107–283.
- Horstmann, A., and Hoffmann, K. P. (2005). Target selection in eye-hand coordination: do we reach to where we look or do we look to where we reach? *Exp. Brain Res.* 167, 187–195. doi: 10.1007/s00221-005-0038-6
- Jeannerod, M. (1988). *The Neural and Behavioural Organization of Goal-Directed Movements*. Oxford: Clarendon Press.
- Johansson, R. S., Westling, G., Backstrom, A., and Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *J. Neurosci.* 21, 6917–6932.
- Konczak, J., Borutta, M., Topka, H., and Dichgans, J. (1995). The development of goal-directed reaching in infants: hand trajectory formation and joint torque control. *Exp. Brain Res.* 106, 156–168. doi: 10.1007/BF00241365
- Konczak, J., and Dichgans, J. (1997). The development toward stereotypic arm kinematics during reaching in the first 3 years of life. *Exp. Brain Res.* 117, 346–354. doi: 10.1007/s002210050228
- Land, M. F., Mennie, N., and Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception* 28, 1311–1328. doi: 10.1068/p2935
- Lasky, R. E. (1977). The effect of visual feedback of the hand on the reaching and retrieval behavior of young infants. *Child Dev.* 48, 112–117. doi: 10.2307/1128888
- Lockman, J. J., Ashmead, D. H., and Bushnell, E. W. (1984). The development of anticipatory hand orientation during infancy. *J. Exp. Child Psychol.* 37, 176–186. doi: 10.1016/0022-0965(84)90065-1
- McCarty, M. E., and Ashmead, D. H. (1999). Visual control of reaching and grasping in infants. *Dev. Psychol.* 35, 620–630. doi: 10.1037/0012-1649.35.3.620
- McDonnell, P. M. (1975). The development of visually guided reaching. *Percept. Psychophys.* 18, 181–185. doi: 10.3758/BF03205963
- McDonnell, P. M. (1979). Patterns of eye-hand coordination in the first year of life. *Can. J. Psychol.* 33, 253–267. doi: 10.1037/h0081724
- Morronegiello, B. A., and Rocca, P. T. (1989). Visual feedback and anticipatory hand orientation during infants' reaching. *Percept. Mot. Skills* 69, 787–802. doi: 10.2466/pms.1989.69.3.787
- Perris, E., and Clifton, R. (1988). Reaching in the dark toward sound as a measure of auditory localization in infants. *Infant Behav. Dev.* 11, 473–491. doi: 10.1016/0163-6383(88)90007-0
- Piaget, J. (1936/1952). *The Origins of Intelligence in Children [La naissance de l'intelligence chez l'enfant]*. New York, NY: Basic Books.
- Piéraut-Le Bonniec, G. (1985). From visual-motor anticipation to conceptualization: reaction to solid and hollow objects and knowledge of the function of containment. *Infant Behav. Dev.* 8, 413–424. doi: 10.1016/0163-6383(85)90005-0
- Pogetti, L. S., Souza, R. M., de Tudella, E., and Teixeira, L. A. (2013). Early infant's use of visual feedback in voluntary reaching for a spatial target. *Front. Psychol.* 4:520. doi: 10.3389/fpsyg.2013.00520
- Rakison, D. H., and Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age? *Dev. Sci.* 15, 43–53. doi: 10.1111/j.1467-7687.2011.01096.x
- Robin, D. J., Berthier, N. E., and Clifton, R. K. (1996). Infants' predictive reaching for moving objects in the dark. *Dev. Psychol.* 32, 824–835. doi: 10.1037/0012-1649.32.5.824
- Rochat, P. (1998). Self-perception and action in infancy. *Exp. Brain Res.* 123, 102–109. doi: 10.1007/s002210050550
- Rochat, P., and Hespos, S. J. (1997). Differential rooting response by neonates: evidence for an early sense of self. *Early Dev. Parenting* 6, 105–112.
- Rochat, P., and Morgan, R. (1998). Two functional orientations of self-exploration in infancy. *Br. J. Dev. Psychol.* 16, 139–154. doi: 10.1111/j.2044-835X.1998.tb00914.x
- Rosander, K., and Von Hofsten, C. (2011). Predictive gaze shifts elicited during observed and performed actions in 10-month-old infants and adults. *Neuropsychologia* 49, 2911–2917. doi: 10.1016/j.neuropsychologia.2011.06.018
- Rosengren, K. S., Pick, H. L., and Von Hofsten, C. (1988). Role of visual information in ball catching. *J. Mot. Behav.* 20, 150–164. doi: 10.1080/00222895.1988.10735439
- Sporns, O., and Edelman, G. M. (1993). Solving Bernstein's problem: a proposal for the development of coordinated movement by selection. *Child Dev.* 64, 960–981. doi: 10.2307/1131321
- Thelen, E. (1995). Motor development. A new synthesis. *Am. Psychol.* 50, 79–95. doi: 10.1037/0003-066X.50.2.79
- Thelen, E., Corbetta, D., Kamm, K., Spencer, J. P., Schneider, K., and Zernicke, R. F. (1993). The transition to reaching: mapping intention and intrinsic dynamics. *Child Dev.* 64, 1058–1098. doi: 10.2307/1131327
- Thelen, E., Corbetta, D., and Spencer, J.-P. (1996). Development of reaching during the first year: role of movement speed. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 1059–1076. doi: 10.1037/0096-1523.22.5.1059
- Troester, H., and Brambring, M. (1993). Early motor development in blind infants. *J. Appl. Dev. Psychol.* 14, 83–106. doi: 10.1016/0193-3973(93)90025-Q
- van der Meer, A. L., van der Weel, F. R., and Lee, D. N. (1994). Prospective control in catching by infants. *Perception* 23, 287–302. doi: 10.1068/p230287
- van der Meer, A. L., van der Weel, F. R., Lee, D. N., Laing, I. A., and Lin, J. P. (1995a). Development of prospective control of catching moving objects in preterm at-risk infants. *Dev. Med. Child Neurol.* 37, 145–158. doi: 10.1111/j.1469-8749.1995.tb11984.x
- van der Meer, A. L. H., van der Weel, F. R., and Lee, D. N. (1995b). The functional significance of arm movements in neonates. *Science* 267, 693–695. doi: 10.1126/science.7839147
- von Hofsten, C. (1979). Development of visually guided reaching: the approach phase. *J. Hum. Mov. Stud.* 5, 160–178.
- von Hofsten, C. (1980). Predictive reaching for moving objects by human infants. *J. Exp. Child Psychol.* 30, 369–382. doi: 10.1016/0022-0965(80)90043-0
- von Hofsten, C. (1982). Eye-hand coordination in the newborn. *Dev. Psychol.* 18, 450–461. doi: 10.1037/0012-1649.18.3.450
- von Hofsten, C. (1983). Catching skills in infancy. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 75–85. doi: 10.1037/0096-1523.9.1.75
- von Hofsten, C. (1984). Developmental changes in the organization of pre-reaching movements. *Dev. Psychol.* 20, 378–388. doi: 10.1037/0012-1649.20.3.378
- von Hofsten, C. (1986). Early spatial perception taken in reference to manual action. *Acta Psychol.* 63, 323–335. doi: 10.1016/0001-6918(86)90077-6



- von Hofsten, C. (1991). Structuring of early reaching movements: a longitudinal study. *J. Mot. Behav.* 23, 280–292. doi: 10.1080/00222895.1991.9942039
- von Hofsten, C. (1993). Prospective control: a basic aspect of action development. *Hum. Dev.* 36, 253–270. doi: 10.1159/000278212
- von Hofsten, C. (2005). “The development of prospective control in tracking a moving object,” in *Minnesota Symposium on Child Psychology*, eds J. J. Rieser, J. J. Lockman, and C. A. Nelson (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 51–89.
- von Hofsten, C. (2009). Action, the foundation for cognitive development. *Scand. J. Psychol.* 50, 617–623. doi: 10.1111/j.1467-9450.2009.00780.x
- von Hofsten, C., and Fazel-Zandy, S. (1984). Development of visually guided hand orientation in reaching. *J. Exp. Child Psychol.* 38, 208–219. doi: 10.1016/0022-0965(84)90122-X
- von Hofsten, C., and Rönqvist, L. (1988). Preparation for grasping an object: a developmental study. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 610–621. doi: 10.1037/0096-1523.14.4.610
- White, B. L., Castle, P., and Held, R. (1964). Observations on the development of visually directed reaching. *Child Dev.* 35, 349–364. doi: 10.1111/j.1467-8624.1964.tb05944.x
- Williams, J. L. (2009). *The Effects of Task Exposure and Simulated Grasping Experience Upon the Emergence of Reaching*. Master Project, University of Tennessee.
- Williams, J. L. (2011). *The Impact of Repeated Sensory-Motor Experience with Multimodal Objects Upon the Emergence of Infant Reaching*. Dissertation project, University of Tennessee.
- Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322
- Witherington, D. C. (2005). The development of prospective grasping control between 5 and 7 months: a longitudinal study. *Infancy* 7, 143–161. doi: 10.1207/s15327078in0702\_2

**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.

Received: 01 April 2014; accepted: 23 May 2014; published online: 11 June 2014.

Citation: Corbetta D, Thurman SL, Wiener RF, Guan Y and Williams JL (2014) Mapping the feel of the arm with the sight of the object: on the embodied origins of infant reaching. *Front. Psychol.* 5:576. doi: 10.3389/fpsyg.2014.00576

This article was submitted to Cognitive Science, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Corbetta, Thurman, Wiener, Guan and Williams. 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.



# Developing embodied cognition: insights from children's concepts and language processing

Michele Wellsby and Penny M. Pexman\*

Language Processing Laboratory, Department of Psychology, University of Calgary, Calgary, AB, Canada

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Serge Thill, University of Skövde, Sweden

Daniela Corbetta, University of Tennessee, USA

## \*Correspondence:

Penny M. Pexman, Language Processing Laboratory, Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB T2N1N4, Canada  
e-mail: pexman@ucalgary.ca

Over the past decade, theories of embodied cognition have become increasingly influential with research demonstrating that sensorimotor experiences are involved in cognitive processing; however, this embodied research has primarily focused on adult cognition. The notion that sensorimotor experience is important for acquiring conceptual knowledge is not a novel concept for developmental researchers, and yet theories of embodied cognition often do not fully integrate developmental findings. We propose that in order for an embodied cognition perspective to be refined and advanced as a lifelong theory of cognition, it is important to consider what can be learned from research with children. In this paper, we focus on development of concepts and language processing, and examine the importance of children's embodied experiences for these aspects of cognition in particular. Following this review, we outline what we see as important developmental issues that need to be addressed in order to determine the extent to which language and conceptual knowledge are embodied and to refine theories of embodied cognition.

**Keywords: developmental science, embodied cognition, language development, sensorimotor processing, action, concepts**

Embodied cognition (EC) is a broad term used to describe a class of theories within cognitive science, many of which emphasize the importance of sensorimotor experience gained through our bodily interactions with the environment for acquiring and representing conceptual knowledge (Borghi and Cimatti, 2010). That is, contrary to classical cognitive theories, which deemphasized the importance of the body for cognitive processing and posited that cognition strictly involved the processing of abstract and amodal symbols, EC theories tend to assume that our actions and bodily experiences are crucial to our cognitive processing. According to EC theories, direct sensorimotor interactions are essential for gaining knowledge and developing cognitive capabilities (Engel et al., 2013), and higher order and offline cognitive processing (i.e., removed from the environment) involve re-enactment of the bodily states from previous experience (Foglia and Wilson, 2013).

Theories of EC have become a prominent way of conceptualizing cognitive processing and have been particularly influential in reconceptualising and explaining adult language processing. A large number of studies have now provided evidence that when comprehending language, adults simulate the meaning implied in words and sentences [e.g., implied motion (Glenberg and Kaschak, 2002), object orientation (Stanfield and Zwaan, 2001), object affordances (Myung et al., 2006)]. Thus, adults use sensorimotor information gained through their experiences with the world to represent concepts and comprehend language. There continues to be debate, however, about whether sensorimotor experiences comprise conceptual knowledge and language or whether accessing this information merely activates sensorimotor areas epiphenomenally. In the adult literature, there are now a number of variants of EC theories that posit different degrees of embodiment and disembodiment (e.g., Mahon

and Caramazza, 2008; Meteyard et al., 2012). These theories can be viewed along a continuum ranging from strongly embodied to disembodied, differing in their assumptions about the nature of the relationship between sensorimotor and cognitive processing.

The disembodied end of the spectrum is represented by what is essentially the classical cognitive perspective described above, which posits that sensorimotor experiences are not involved in cognitive processing (Meteyard et al., 2012). From a developmental perspective, this end of the spectrum would be represented by the view that, while sensorimotor experiences might be important for infants' earliest learning, cognition becomes progressively more abstract and less embodied with development. At the other end of the spectrum, a strong embodied account suggests that cognition is constituted in action and sensorimotor processing (Glenberg and Gallese, 2012), and that our conceptual representations are dependent on sensorimotor experiences. From a strong embodied perspective, cognitive processing involves a recreation of direct sensory experience (Meteyard et al., 2012), in childhood and beyond.

An alternative view is taken by secondary embodiment theories, which propose that sensorimotor areas of the brain are activated as a by-product of cognitive processing, through spreading activation (Mahon and Caramazza, 2008). From this perspective, sensorimotor activation during cognitive processing is a passive consequence of, as opposed to a necessity for, representing a concept. Finally, a weak embodied account suggests that conceptual representations are partially comprised of sensorimotor knowledge, as sensorimotor interactions help to ground concepts during initial knowledge acquisition. However, activation of this same sensorimotor information is not required for conceptual

processing; rather, representations are abstracted from the initial experience, and then are organized to form conceptual knowledge (Gennari, 2012; Meteyard et al., 2012).

There is also growing support for hybrid or pluralist theories that add to or combine different components of the embodiment spectrum (e.g., Paivio, 1990; Barsalou et al., 2008; Louwerse and Jeuniaux, 2010). For example, Dove (2011) proposed that an embodied approach, in which conceptual representations consist mainly of simulation of previous sensorimotor experience (perceptual symbols), is more useful for certain concepts than others; specifically, for concrete concepts as compared to abstract concepts. Dove emphasized that in order for an embodied theory to adequately provide an explanation of abstract concepts, language and linguistic symbols would be important. Thus, concepts comprise both sensorimotor representations, gained through previous embodied experience, and also what Dove called dis-embodied representations, gained from our experience with language. From this perspective, our knowledge of concepts is not only comprised of our sensorimotor experience but also how we use language. By this view, concrete concepts are comprised of embodied sensorimotor information from previous interactive experience with objects and the environment (perceptual symbols), as well as dis-embodied sensorimotor information from our experience using language (linguistic symbols), whereas our understanding of abstract concepts is mainly comprised of information from our experience with language.

Taking a slightly different perspective, Pulvermuller and Garagnani (2014) proposed that different types of cognitive processing could involve different degrees of embodiment, such that while long-term memory is embodied and is grounded in sensorimotor systems, working memory relies less on those systems. In a similar vein, Zwaan (2014) proposed that rather than arguing for or against a particular version of embodiment, we instead need to investigate the relative importance of sensorimotor information and symbolic representations in different contexts for language processing. In particular, language comprehension that is relatively more embedded in the environment will likely involve more embodied processing.

Thus, it is evident that there are multiple theories of embodiment, which differ in how much emphasis is placed on sensorimotor experiences for conceptual and language processing. It seems possible that a lifespan perspective could afford new insights on these issues by examining the developmental trajectory of how sensorimotor experiences shape language and conceptual processing. In addition, rather than simply taking an “embodied versus disembodied” stance, it is essential to determine specific details surrounding when and how sensorimotor representations are involved in language and conceptual processing (Willems and Francken, 2012). As we will discuss below, children initially use sensorimotor information to gain conceptual knowledge. By examining how and when sensorimotor information is important for children's linguistic and conceptual understanding, and determining if and when they shift away from a reliance on this sensorimotor knowledge as their cognition becomes more sophisticated and more abstract, it seems likely that developmental research could help advance EC theories more generally.

As described above, EC theories can be viewed along a continuum with regards to the emphasis placed on the role of embodied experience, and numerous studies have demonstrated that embodied knowledge plays some sort of role in adult concepts and language processing. However, there has been less research conducted to examine embodied effects in children's conceptual and linguistic processing, and less discussion of the implications for EC theories in research examining cognitive development. Although developmental research does not often use the term “embodied cognition” when describing children's cognitive processing, the notion that sensorimotor experience is essential to conceptual and linguistic knowledge is not a novel idea in the developmental field. Kontra et al. (2012) proposed that “theories of embodied cognition have the potential to deepen our understanding of the mechanisms underlying early developmental changes driven by action experience” (p. 738); in addition, we propose that to refine theories of EC, it is essential to consider the insights that can be gleaned from developmental research, examining children's sensorimotor experiences and how those experiences shape their knowledge.

In this paper, we will first review developmental theories and recent evidence from the developmental literature that highlight the importance of sensorimotor experience early on in childhood for the development of later cognitive skills and abilities. By sensorimotor experience, we refer to a range of experiences that typically involve an action being performed on an object, either by a child directly, or through observation of another's action. The experience is multisensory, primarily derived from visual, tactile, and proprioceptive senses. We have chosen to focus on this characterization of sensorimotor experience (which is quite broad) because this is what has typically been examined in child development research. Certainly, grounding of conceptual information could involve other systems, such as emotions (e.g., Pulvermuller, 2013), but there is as yet little research on how children ground the meaning of language and concepts through emotion (we return to this in our final “issues to be addressed” section). Additionally, although there are numerous aspects of development we could examine, we limited our review to emphasize research on children's language and conceptual processing. These areas will be our focus because language and concepts have been at the center of much of the debate between strong, weak, and secondary theories of EC (Zwaan, 2014). We first review these findings, and then describe what we see as pertinent issues that need to be addressed in order that EC theories can be further refined and advanced as theories of lifelong cognitive development.

## THE IMPORTANCE OF SENSORIMOTOR EXPERIENCE IN DEVELOPMENT

Although EC theories have not been prominent in the developmental literature, the notion that sensorimotor experience is essential to child development is certainly not a new concept. The proposal that sensorimotor information initially drives cognitive development was an important aspect of Piaget's work, and he emphasized the influential role of children's interactions with their environment (Piaget, 1952; Laakso, 2011). Piaget argued that in early infancy, sensorimotor experiences are an essential aspect of

learning, and later cognitive processes develop from these sensorimotor abilities. The general idea emphasized in the developmental literature is that infants are embodied learners, and use sensorimotor information to gain knowledge about their world (Laakso, 2011). It has been proposed that infants develop a representational system as a result of early perceptual and motor interactions with their environment (Meltzoff, 1990). These early representations are considered the building blocks that allow embodied learning to continue throughout childhood. Whereas Piaget proposed that children go on to develop concepts that are independent of their sensorimotor experience, others have argued that as children develop increased cognitive and physical capabilities, their sensorimotor interactions with the environment continue to be important for language processing and increased conceptual understanding (Gibbs, 2006).

While few would challenge the claim that infants and young children initially use sensorimotor knowledge and interactions with their environment to acquire information, the extent to which embodied experience is relevant for higher-order cognitive functioning (e.g., language processing) in childhood has been less widely considered. Given the results of adult studies it seems likely, however, that EC theories can ultimately explain how sensorimotor knowledge is beneficial for early sensory learning in infancy, for motor and action development through childhood, and for language and higher-order cognitive functioning in school-aged children (Kontra et al., 2012).

Further, a theory emphasizing the importance of embodiment across the life-span would propose that the role of embodiment in conceptual processing is always present; the influence of sensorimotor experiences does not stop or change fundamentally throughout development, it just may become more refined and flexible over time (Antonucci and Alt, 2011). Indeed, the role of embodiment in conceptual processing is considered by some developmental theorists to be continuous, as conceptual representations across the lifespan are composed of perceptual and action experiences (Thelen, 2008), and the successive development of sensation, action, and language across childhood into adulthood is influenced by the experiences that a child has in their environment (Borghi and Cimatti, 2010). Embodied experiences contribute to a dynamic grounding of cognition over the lifespan that allows children and adults to learn language and represent concepts based on previous sensorimotor interactions (Thelen, 2008). Children interact with their environment and learn concepts, and language can then be mapped onto these representations (e.g., Glenberg and Gallese, 2012). There is evidence that children's sensorimotor experience and actions towards objects directly influence their word and concept learning (e.g., O'Neill et al., 2002; Smith, 2005). Although it appears likely that conceptual knowledge is grounded in the environment from infancy onward (Zwaan and Kaschak, 2009), with sensorimotor interactions continuously shaping cognitive processing, there has been little integration of developmental findings with theories of EC to explain the relationship between cognition and bodily experience across development (Gabbard, 2013).

There is longitudinal evidence for the relationship between children's early sensorimotor (in particular, action) experiences

and later higher-order cognitive functioning. For example, Bornstein et al. (2013) recently reported the results of a longitudinal study conducted to examine motor exploration behavior in infancy and how this behavior predicted academic abilities in adolescence. Bornstein et al. (2013) measured the motor-exploratory competence (movement, balance, and locomotion) and exploratory activity of five-month-old infants. Longitudinal data showed that infants with higher scores on the motor exploration variables at 5-months of age had higher scores on intellectual and academic measures at 4-, 10-, and 14-years of age. While there are likely multiple mediating factors, it is probable that the infants with relatively high motor competency and exploratory behavior had more opportunities for sensorimotor interactions with objects and with their environment. For instance, infants who are able to sit and maintain balance while manipulating objects are able to acquire multimodal sensorimotor information about objects (Smith, 2013). This increased embodied experience could facilitate sustained attention, richer interactive experiences, and more instances of adults labeling objects, which all contribute to greater knowledge of objects in the environment. In turn, vocabulary, attention, and knowledge could all be enhanced, resulting in positive long-term cognitive outcomes like those observed by Bornstein et al. (2013). Thus, evidence suggests that there are benefits of exploration and increased motor activity in infancy (i.e., embodied interactions) for later cognitive development. Of course, this type of research does not allow us to make inferences about the types of embodied experiences that are most important, but that is better achieved by the experimental studies on this topic, reviewed next.

## THE INFLUENCE OF EARLY SENSORIMOTOR EXPERIENCE ON CHILDREN'S CONCEPTS AND WORD LEARNING

### NOUNS

It is widely agreed that before children acquire language, they build conceptual representations based on their sensorimotor experiences with the world (Antonucci and Alt, 2011). Once infants are able to sit and manipulate objects, they are able to acquire information about objects based on motor, tactile, visual, and auditory input (Smith, 2013). Through active exploration with the environment, children develop an increased understanding of the functions of objects and how they can be manipulated. This knowledge of semantic features and object affordances helps children to differentiate objects more easily, and eventually to learn words by mapping labels onto representations based on previous experiences (Scofield et al., 2009).

Findings from research examining infants' and children's interactions demonstrate effects of specific types of sensorimotor experience on categorization and word learning. In particular, the manner in which children act on objects, with regards to actions performed and sensorimotor experience obtained, influences how these objects are conceptualized (Smith, 2005). In Smith's (2005) study, 2-year-old children were introduced to an exemplar object called a "wug," which the experimenter labeled while moving the object either horizontally or vertically. Some of the children were also given the opportunity to move the object themselves, in the same direction. Following this, children were



asked to select the wug from two novel objects: an object that was the same height as the exemplar, but extended horizontally, and an object that was the same diameter as the exemplar, but extended vertically. For the children who had manipulated the object themselves, there was an interaction between the direction they had moved the object and the object they selected as the wug: children who had watched and then moved the wug horizontally chose as the wug the novel object that was extended horizontally, and vice versa. Interestingly, there was no such effect for the children who only watched the experimenter move the objects.

Smith (2005) offered an embodied explanation for these findings, by proposing that the way the children manipulated and acted on the object comprised part of their conceptual representation for that object. For the children who interacted with the object, the sensorimotor experience created a mental representation of the object based on the action performed, which influenced their judgment of the objects' shape; this motor information was later simulated when the children viewed the novel objects and had to make a categorization decision.

Other studies have also demonstrated that the way in which objects are held and manipulated influences the aspects of that object that are relevant for children's categorization (Smith et al., 2007). In one study, 2-year-old children were taught a novel label for an object with a hinge and were given the opportunity to interact with the object. When children were then presented with similarly shaped objects without a hinge and objects that differed in shape but had a hinge, the children were more likely to extend the novel object label to the other objects with hinges. Thus, the functional knowledge gained through interaction with objects can determine how objects and categories are formed.

Additional research has examined how spatial location and body positioning influence word learning (Smith and Samuelson, 2010). Children between 18 and 24 months of age were presented with two unlabeled objects one at a time, one to their right, and one to their left. Following this, the objects were removed and a label was provided to one of the empty locations where an object had previously been presented (e.g., "modi"). When the children were later shown both of the objects in new locations and asked to select the named object ("where is the modi?"), the majority of the children selected the object that had been presented in the location where the label was provided. Thus, children associated the object's location with its label, suggesting that visuospatial experience with the object's location (and not just with the object itself) influences word learning. Interestingly, changing the children's posture from sitting to standing decreased their ability to map the label to the object. This finding suggested, further, that children's body posture also played a role in linking label to object.

To further examine the influence of sensory experience and body posture on object learning, Morse et al. (2010) extended the Smith and Samuelson (2010) paradigm to the field of developmental robotics. Morse et al. (2010) replicated the Smith and Samuelson (2010) experiments using a robot, and reported that the robot's categorization performance was comparable to the children's performance in the Smith and Samuelson study. Taken together, these results indicate that sensory representations, as

well as proprioceptive information about body posture, are both important factors when learning to categorize and map labels to objects. The robotics simulations provide additional insight about the sensory representations that are involved in category learning.

## VERBS

One general theme in the developmental literature is that interactions with the environment play an important role in verb learning. For instance, according to Glenberg and Gallese (2012), children's understanding of verbs is grounded in bodily actions and sensorimotor experiences. That is, a verb like "give" would be understood in infancy from concrete experiences of giving objects to parents/caregivers; the meaning would be grounded in these actions. Children's bodily actions towards other people are also related to their understanding of abstract verbs, such as "love" or "hate," that do not appear to being grounded in one specific action. These verbs can be associated with observable bodily behaviors (such as showing affection) that can help ground understanding of the emotional content associated with the verb meaning (Smith et al., 2007).

Moving beyond infancy, the role of sensorimotor experience in verb learning has been directly examined in young children, with findings indicating that there are differences in brain activation as a result of whether verbs were learned through self-performed or observed actions (James and Bose, 2011; James and Swain, 2011). Children aged 5- to 7-years were taught novel verbs either by actively performing the action while repeating the verb label out loud, or by watching an experimenter perform the action while the experimenter repeated the verb label. Then, children were presented with auditory and visual information from the objects (e.g., verb label, video of the action being performed) during fMRI scanning. When the action label was auditorily presented, motor areas in the brain (including regions associated with grasping objects) were activated only for the verbs the children had learned through self-action, not for verbs learned through passive observation (James and Swain, 2011). The same pattern of findings was observed when viewing videos of the actions, with greater activation for actively learned verbs in areas associated with tool use, integrating motor information, and visual processing (James and Bose, 2011). These findings suggest that sensorimotor movements evoked when learning language are reactivated during recognition. Further, it appears that in order for perception and action to become linked and for motor representations to be re-activated when action verbs are heard, children may need to have actively interacted with objects.

## ADJECTIVES

Studies have also examined the influence of children's sensorimotor experience with objects when learning other parts of speech, such as adjectives. For example, two-year-old children were taught novel adjectives (e.g., spiny, spongy) by an adult using either referential gestures toward an object (i.e., pointing to an object) or descriptive gestures (e.g., using tactile gestures, such as squeezing the spongy object; O'Neill et al., 2002). On each trial, an animal toy was given to the child and an adjective was provided. When providing the adjective, the experimenter either gestured with the toy to

illustrate the property or pointed to the toy. Thus, the descriptive gestures provided sensorimotor information about the adjective, through observation as well as any actions the child made toward the toy. In contrast, the referential point did not provide this sensorimotor information and only acted as an attentional cue.

On test trials, children were presented with two toys and asked for one displaying a specific property (e.g., "Give me the lumpy toy"). The children who were taught adjectives using descriptive gestures performed better at test, and additionally, descriptive gestures were especially helpful at teaching adjectives that did not correspond to visual properties. That is, observation of descriptive gestures was more beneficial for teaching adjectives such as lumpy and spongy, where tactile experiences are essential to meaning, as compared to adjectives such as spiny, for which the meaning can be inferred through visual inspection. Interestingly, more accurate performance in the test trials was not related to the amount of sensorimotor interaction the children had during the teaching trials. There was, however, a positive relationship between performance and interaction at test. That is, the children who were taught adjectives by viewing descriptive gestures used this sensory information in the test trials to perform the gesture themselves and, presumably, to determine which object fit with the adjective they were asked to identify (O'Neill et al., 2002). Thus, although the children in both conditions interacted with the objects during the training trials, the children who were in the descriptive gesture condition seemed to use these gestures as a cue to focus on that specific property of the object. It seems likely that the children who observed descriptive gestures gained tactile information about the objects that allowed them to ground the meaning of the adjective.

#### QUALIFYING THE BENEFITS OF SENSORIMOTOR EXPERIENCE

Although there is evidence that sensorimotor experience supports children's word learning, there is also evidence that this is not always the case. Tare et al. (2010) examined how manipulative features influenced children's learning of novel animal names from picture books. That is, 20-month-old children were taught labels for novel animals using one of three picture books: a book with drawings of animals, a book with photos of animals, or a book with drawings of animals and manipulative features with which the children could interact (e.g., a flap to pull up to reveal an animal). The children who were taught the animal name using a picture book with realistic photos demonstrated the most accurate learning, while children who were read the picture book with manipulative features had the least accurate learning. A similar pattern was observed in a second study with 30- and 36-month-olds who were read the same books but were also taught facts about the animals (e.g., birds like to eat worms). These findings indicate that having children interact with attention-capturing features like pop up flaps may not always be beneficial for word learning, particularly when the sensorimotor experience obtained does not correspond with the information to be learned.

It is also possible that certain kinds of sensory information may be more relevant for learning certain word classes. For instance, it has been suggested that functional information may be relatively more important than sensory information for distinguishing between inanimate objects (Warrington and Shallice,

1984). Results from recent robotics work suggest additional differences between the information that is important for learning words of different classes. In a study by Yuruten et al. (2013), a robot interacted with objects using different manipulations to learn nouns and adjectives. The authors examined the relevance of different object features, and determined that object affordances were more important for learning adjectives, while object appearance was more important for learning nouns. This suggests, again, that specific kinds of sensorimotor experiences are useful for learning different kinds of concepts.

While an EC account would propose that previous interactions with an object comprise the representation for that object (e.g., our representation of the concept "car" consists of our previous experiences interacting with cars; Barsalou, 1999), it seems likely that some kinds of interactions are more influential than others. It is not yet known whether sensorimotor experience is always linked to the representation of a word and is beneficial for language learning regardless of whether this sensorimotor knowledge is directly involved in the specific word meaning. For example, holding a pencil provides sensorimotor information about its hardness, but does this improve children's ability to later label this object, compared to simply observing a pencil or being told the function of a pencil? It may be that any sensorimotor information leads to acquisition of a richer semantic representation, and therefore word learning is facilitated (Barsalou, 1999). However, this may not necessarily be the case, as studies examining the effects of manipulatives on learning have demonstrated that physically interacting with perceptually rich stimuli when the sensorimotor information gained through physical manipulation is not directly related to the object name can hinder, rather than facilitate, learning an object name (McNeil et al., 2009). Embodied learning experiences may be more beneficial when the sensorimotor information obtained relates to the information learned. When this is not the case, the embodied experiences may in fact alter what is required to complete the task, and thus facilitatory embodied effects are not observed. For example, in the Tare et al. (2010) study described above, the manipulations that were performed by the children did not provide sensorimotor experience that would help the children obtain knowledge about the animals. Attractive, attention-getting stimuli may not help children to learn the intended meaning of an abstract concept or a printed word, if the appealing element needs to be represented as a symbol for something else (Uttal et al., 2009).

Object labels (nouns) are typically the first part of speech that children learn (Waxman et al., 2013), and objects tend to be perceptually rich, with numerous affordances. As such, there may be circumstances where there is no incremental benefit to providing children with additional sensorimotor experience when teaching object labels. Evidence indicates that certain embodied instructional methods may only be beneficial for certain types of information. For example, de Nooijer et al. (2013) found children's knowledge for verbs was improved when they imitated a model by gesturing during encoding or during later retrieval; however, this gesture method was only beneficial for verbs that involved some sort of object manipulation with the hands. No beneficial effect of gesturing was observed for locomotion verbs or abstract verbs. It seems likely that in order for sensorimotor experience to be

beneficial for learning, this experience needs to be appropriate and relevant to the material to be learned (Kiefer and Trumpp, 2012). Of course, defining what it means for experience to be “appropriate” to word learning is something that has not yet been achieved.

A recent trend in the embodiment literature has been to emphasize the ways in which technology can be used to facilitate learning, and results suggest that the benefits of computer interaction may depend on the information to be learned. In recent research both children and adults demonstrated better letter recognition after hand writing new letters than after typing new letters (Kiefer and Trumpp, 2012). As another example, Smeets and Bus (2012) used computer storybooks to teach 4- and 5-year-old children new words. All children saw the story scenes presented, and heard the story narration, on the computer. Children either had the story read to them with certain key words repeated, had the story read to them and interacted using the mouse to find word “hotspots” in the story, or had the story read to them and at certain points they were presented with a multiple choice question about an object, with feedback. Children who responded to multiple-choice questions learned new words more accurately than those who interacted with the story to find the “hotspots” (Smeets and Bus, 2012). It seems that while there are some applications of technology that can provide embodied experience for letter and word learning, these experiences need to correspond to the information being learned.

## THE INFLUENCE OF SENSORIMOTOR EXPERIENCE ON LANGUAGE PROCESSING

Recent research has demonstrated that embodied effects can also be observed in children's early reading comprehension. Specifically, children's acquisition of conceptual knowledge is enhanced when they represent story information by interacting with physical objects or manipulating objects on a computer to represent story information (Glenberg et al., 2011). In one study, 6- and 7-year-old children with low reading skills read stories about a series of events (e.g., on a farm, at the zoo; Marley et al., 2010). Children were assigned to one of three conditions: children in one condition read story sentences and at certain points used toys to act out the story action from the previous sentence. Children in the second condition read story sentences and then watched the experimenter manipulate the toys to correspond with the sentences. Finally, children in the third condition simply reread each sentence a second time. Children in the first condition, who actively manipulated the toys themselves, and children in the second condition, who observed the experimenter manipulate the toys, had more accurate recall for story events in a subsequent comprehension task than did children in the third condition.

This embodied approach to reading development was later termed “moved by reading” (Glenberg, 2011), and was extended in a further study to examine the influence of interacting with technology on reading comprehension. Glenberg et al. (2011) showed that the facilitatory effect of interaction was observed even when children manipulated story objects on a computer screen by clicking and dragging with a mouse. In some instances, computer

manipulation was actually more beneficial than physical manipulation. This may be because understanding the components of the story does not require information gained from direct manipulation of physical objects; that is, haptic information such as weight or information on how to manipulate specific objects was not required in order to comprehend the stories.

These findings indicate that embodied experiences with real objects manipulated by either the self or others, as well as object manipulations on a computer, can facilitate children's language comprehension by helping them to situate the concepts from the story in experience. These manipulation activities ground the semantic and syntactic information in the sentence in action and experience, either with the physical objects, the computer objects, or through imagining. The aim of this reading program is to make reading comprehension fast and automatic, by linking written words to sensorimotor experience (Glenberg et al., 2013).

The reading studies described above examined the influence of sensorimotor experience during language comprehension; this kind of direct effect of sensorimotor interaction is often referred to as an online effect. In contrast, offline effects occur in the absence of direct interactions, and in this vein research has also demonstrated that *previous* sensorimotor experience can influence children's language processing. For instance, developmental studies have explored offline effects of sensorimotor experience on children's language processing during passive listening (James and Maouene, 2009), word naming (Wellsby and Pexman, in press) and sentence/picture verification tasks (Engelen et al., 2011).

James and Maouene (2009) presented 4- and 5-year-old children with auditory lists of verbs and adjectives while they were in the MRI scanner. The results indicated that areas of the brain associated with motor processing were activated when the children listened to verbs, but not when the children listened to adjectives. These results suggest that in the developing brain there is a link between sensorimotor experience and language processing, as words that are associated with action elicit activation in the corresponding motor areas of the brain.

Wellsby and Pexman (in press) examined the influence of previous sensorimotor experience on language processing in slightly older children, using a word naming task. They assessed prior sensorimotor experience using the body-object interaction (BOI) variable. BOI captures how easily a human body can interact with a word's referent (Siakaluk et al., 2008). This variable indexes previous sensorimotor experience, and in adult word recognition studies responses tend to be faster and more accurate to words that are high in BOI (e.g., belt) than words that are low in BOI (e.g., ship); this is termed the BOI effect (e.g., Siakaluk et al., 2008; Tillotson et al., 2008). In the Wellsby and Pexman study, 6- to 10-year-old children completed a word naming task in which high and low BOI words were presented one at a time on a computer screen and children were instructed to read the words out loud. The BOI effect in children's naming behavior was assessed using a composite measure obtained from children's response latency and accuracy data. Results showed that younger children (aged 6- to 8-years) did not show a BOI effect, but older children (aged 8- to 10-years) showed a facilitatory BOI effect for word naming. Wellsby and Pexman proposed that for the older children, the high BOI words activated richer

semantic representations based on previous sensorimotor experiences (either personal experience or observed experience with the words' referents). These richer representations, in the context of the older children's relatively more proficient reading systems, led to a facilitatory BOI effect. Therefore, once children have developed reasonably efficient lexical systems and sufficient sensorimotor experience with words' referents, they are able to use previous sensorimotor experience to facilitate word reading.

There is also evidence that older children (aged 7- to 13-years-old) construct sensory simulations of the objects and situations implied in sentences (Engelen et al., 2011). In this study, children either listened to sentences (Experiment 1) or read sentences (Experiment 2). Following each sentence children viewed a picture of an object that either matched or mismatched the visual orientation implied in the sentence. Children had to determine whether the object in each picture had been mentioned in the sentence. The results for both experiments indicated that children were faster to make this judgment when the picture matched the orientation implied in the sentence and were slower for mismatching pictures. Engelen et al. suggested that their results provide support for embodied theories of language comprehension, and the findings indicate that even in children's developing language processing systems, simulations are constructed of the objects and events described in each sentence. Thus, the Engelen et al. (2011) and Wellsby and Pexman (in press) studies both demonstrate that language processing in older children is grounded in sensorimotor information, even when processing is offline, and separated from direct sensorimotor engagement.

## ISSUES TO BE ADDRESSED

As reviewed above, some progress has been made in understanding the role of sensorimotor experience in children's conceptual and language learning. At the same time, there are numerous issues left to be resolved, and we highlight some of these issues here.

### WHAT SPECIFIC KINDS OF SENSORIMOTOR EXPERIENCE ARE MOST RELEVANT TO CHILDREN'S CONCEPTUAL AND LANGUAGE PROCESSING?

In this review we have discussed the role of sensorimotor experience in children's language and conceptual development. However, our construal of what constitutes "sensorimotor experience" is quite broad and included a range of sensory experiences, for instance, moving objects in space (e.g., Smith, 2005), performing actions on objects (e.g., James and Swain, 2011), getting tactile information from touching objects (O'Neill et al., 2002), general motor exploration of the environment (Bornstein et al., 2013) and visual experience watching someone else manipulate objects (e.g., Marley et al., 2010). While the term sensorimotor experience is generally used to refer to that which results from some sort of action, it can also be primarily visual or proprioceptive. As such, further research should aim to more precisely determine the specific kinds of sensorimotor experience that are beneficial for children's learning. As mentioned, the correspondence between sensorimotor experience and the concept to be learned is likely important. In addition, examining the exact nature of children's experience in a task, and

analyzing what it is they were trying to learn may help to determine the underlying mechanisms involved (see Wilson and Golonka, 2013, for extensive discussion of the need for task analysis).

As reviewed above, there are suggestions that the type of sensorimotor experience most relevant to word learning will depend on word class. It has also been argued that the extent to which embodied effects (attributed to sensorimotor experience) emerge in language processing depends on the degree to which that form of language comprehension is embedded in the environment (Zwaan, 2014). While there is now some research with adults on the context sensitivity of embodied language processing (e.g., van Dam et al., 2010; Tousignant and Pexman, 2012) this principle has not yet been tested in children. That is, we do not yet understand development of context sensitivity, and this seems a critical element in our understanding of the developmental pathway from the sensorimotor infant to the literate older child.

### ARE THERE OTHER ASPECTS OF EMBODIMENT THAT NEED TO BE CONSIDERED?

In the embodied literature, there has been a tendency to focus on effects of overt, goal-directed actions performed by the body; there has been less emphasis on passive sensations associated with having a body in the world when we are not directly interacting with objects (Borghi and Cimatti, 2010; Sidhu et al., 2014). That is, there has primarily been a focus on bodies acting in the environment, with limited examination of *sensing* bodies. This tendency has also been evident in the child literature. In order to fully understand development of EC, future research needs to examine the mechanisms involved in children developing a sense of their body, grounded in sensation, action, and language (Borghi and Cimatti, 2010).

A related issue is the manner in which we tend to characterize language itself. According to Borghi and Cimatti (2010) language can be conceived of as a tool that allows us to interact with our environment and, by using language, we can develop a sense of our body removed from direct actions with objects. Rather than focusing simply on how words are represented in the brain as a result of sensorimotor experiences, future research needs to examine how words can be used as tools to extend our body and interact with others (Borghi et al., 2013).

### HOW DO CHILDREN LEARN ABSTRACT CONCEPTS?

The majority of the literature reviewed has focused on children's learning about concrete concepts and language through direct sensorimotor experience interacting with objects. A major debate in the literature concerns the extent to which EC theories can explain the processing of abstract concepts (e.g., Barsalou, 2008; Mahon and Caramazza, 2008). Therefore, an examination of whether embodied experiences can help children learn abstract concepts could suggest a developmental trajectory for the acquisition of abstract concepts. One mechanism through which abstract concepts might be embodied is the semantic association to emotional states (Pulvermuller, 2013; Zdravilova and Pexman, 2013). Abstract words can be understood and become grounded through their associated emotional and physiological experiences, which



are also considered forms of embodiment (Kousta et al., 2011). Through the experience of various emotional states and situations, the meaning of abstract concepts can become grounded in embodied experience.

There is extensive research on how adults respond to emotion words in language processing tasks (e.g., Kousta et al., 2011). This work has helped clarify how the meanings of abstract words, in particular, may be grounded through emotion. To our knowledge, no parallel research has been conducted with children. Such studies could help identify, for instance, when children begin to show effects of valence in word recognition, and how this is related (or not) to understanding abstract concepts.

Of course, emotion is not the only means by which children could learn the meanings of abstract words. Borghi et al. (2011) described a training study in which adults learned the meanings of novel concrete and abstract concepts. The study tested the notion that children learn abstract meanings through verbal explanation and relationships between perceivable objects, while they learn concrete meanings through perception and action with manipulable objects. Results were consistent with these claims, and with claims about grounding of abstract meaning in language and perception (e.g., Barsalou et al., 2008; Dove, 2011). It will be important to further evaluate this proposal in future studies with children.

### THE IMPACT OF SENSORIMOTOR DEFICITS ON LANGUAGE AND CONCEPTUAL DEVELOPMENT

As mentioned, one debate in the EC literature is focused on whether sensorimotor information is essential for conceptual processing, or if it is information that is activated epiphenomenally, as a result of spreading activation. It seems likely that this debate could be constrained by additional developmental studies on the connection between children's sensorimotor abilities and their acquisition of language and concepts. We know that advanced motor skills and exploratory behavior in infancy are related to increased academic outcomes later in life (Bornstein et al., 2013), and a link between children's fine motor skills and vocabulary level has also been observed (Dellatolas et al., 2003), indicating that early embodied experiences have positive influences on children's conceptual and language learning.

While the focus has tended to be on these positive associations, additional inferences could be drawn from work on the relationship between early motor skill deficits and impairments in language and conceptual processing. For instance, developmental coordination disorder (DCD) is characterized by a general impairment in motor coordination (Visser, 2003). Many children diagnosed with DCD also show problems in other sensory domains such as vision and perception, and in cognitive domains such as attention, concentration, and language. In addition, children with specific language impairment (SLI), which is characterized by atypical language development, often show fine and gross motor skill deficits (Hill, 2001). There are numerous hypotheses as to why language and motor deficits are related, including a general slowing in processing speed (Kail, 1994), a deficit in the ability to automate skills (Fawcett et al., 1996), or an abnormality in certain brain structures (Hill, 2001). A lifespan perspective of EC could help

to clarify the relationship between language and motor difficulties, and by unpacking the nature of this relationship we could provide new insight on the issue of whether sensorimotor experience is necessary for conceptual and language processing.

### CONCLUSION

In order for theories of EC to fully describe conceptual and linguistic processing across the lifespan, several issues will need to be addressed, and we have outlined some of those here. Further studies need to be conducted to examine how sensorimotor processes interact with the developing linguistic and conceptual systems in order to map out the full developmental trajectory of EC. A lifespan approach to EC will involve mapping the developmental pathways (as Smith, 2013, recommends), through which sensorimotor experiences influence the acquisition of conceptual and linguistic knowledge. This kind of integration will be a challenge, but in tackling it we believe that theories of EC can be further refined.

### REFERENCES

- Antonucci, S. M., and Alt, M. (2011). A lifespan perspective on semantic processing of concrete concepts: does a sensory/motor model have the potential to bridge the gap? *Cogn. Affect. Behav. Neurosci.* 11, 551–572. doi: 10.3758/s13415-011-0053-y
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/S0140525X99002149
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Barsalou, L. W., Santos, A., Simmons, W. K., and Wilson, C. D. (2008). "Language and simulation in conceptual processing," in *Symbols, Embodiment, and Meaning*, eds M. De Vega, A. M. Glenberg, and A. Graesser (Oxford: Oxford University Press), 245–284. doi: 10.1093/acprof:oso/9780199217274.003.0013
- Borghi, A. M., and Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia* 48, 763–773. doi: 10.1016/j.neuropsychologia.2009.10.029
- Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., and Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Front. Psychol.* 2:15. doi: 10.3389/fpsyg.2011.00015
- Borghi, A. M., Scorolli, C., Caligiore, D., Baldassarre, G., and Tummolini, L. (2013). The embodied mind extended: using words as social tools. *Front. Psychol.* 4:214. doi: 10.3389/fpsyg.2013.00214
- Bornstein, M. H., Hahn, C.-S., and Suwalsky, J. T. D. (2013). Physically developed and exploratory young infants contribute to their own long-term academic achievement. *Psychol. Sci.* 24, 1906–1917. doi: 10.1177/0956797613479974
- Dellatolas, G., De Agostini, M., Curt, F., Kremin, H., Letierce, A., Maccario, J., et al. (2003). Manual skill, hand skill asymmetry, and cognitive performances in young children. *Laterality* 8, 317–338. doi: 10.1080/13576500342000121
- de Nooijer, J. A., van Gog, T., Paas, F., and Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychol.* 144, 173–179. doi: 10.1016/j.actpsy.2013.05.013
- Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Front. Psychol.* 1:242. doi: 10.3389/fpsyg.2010.00242
- Engel, A. K., Maye, A., Kurthen, M., and Konig, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends Cogn. Sci.* 17, 202–209. doi: 10.1016/j.tics.2013.03.006
- Engelen, J. A. A., Bouwmeester, S., de Bruin, A. B. H., and Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *J. Exp. Child Psychol.* 110, 659–675. doi: 10.1016/j.jecp.2011.06.009
- Fawcett, A. J., Nicolson, R. I., and Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tasks. *Ann. Dyslexia* 46, 259–283. doi: 10.1007/BF02648179
- Foglia, L., and Wilson, R. A. (2013). Embodied cognition. *WIREs Cogn. Sci.* 4, 319–325. doi: 10.1002/wcs.1226

- Gabbard, C. (2013). The role of mental simulation in embodied cognition. *Early Child Dev. Care* 183, 643–650. doi: 10.1080/03004430.2012.681649
- Gennari, S. P. (2012). Representing motion in language comprehension: lessons from neuroimaging. *Lang. Linguist. Compass* 6, 67–84. doi: 10.1002/lnc3.317
- Gibbs, R. W. (2006). *Embodiment and Cognitive Science*. New York: Cambridge University Press.
- Glenberg, A. M. (2011). How reading comprehension is embodied and why that matters. *Int. Electron. J. Elem. Educ.* 4, 5–18.
- Glenberg, A. M., and Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex* 48, 905–922. doi: 10.1016/j.cortex.2011.04.010
- Glenberg, A. M., Goldberg, A. B., and Zhu, X. (2011). Improving early reading comprehension using embodied CAL. *Instruct. Sci.* 39, 27–39. doi: 10.1007/s11251-009-9096-7
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565. doi: 10.3758/BF03196313
- Glenberg, A. M., Witt, J. K., and Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspect. Psychol. Sci.* 8, 573–585. doi: 10.1177/1745691613498098
- Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *Int. J. Lang. Commun. Disord.* 36, 149–171. doi: 10.1080/13682820010019874
- James, K. H., and Bose, P. (2011). Self-generated actions during learning objects and sounds create sensori-motor systems in the developing brain. *Cogn. Brain Behav.* 15, 485–503.
- James, K. H., and Maouene, J. (2009). Auditory verb perception recruits motor systems in the developing brain: an fMRI investigation. *Dev. Sci.* 12, F26–F34. doi: 10.1111/j.1467-7687.2009.00919.x
- James, K. H., and Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Dev. Sci.* 14, 673–678. doi: 10.1111/j.1467-7687.2010.01011.x
- Kail, R. (1994). A method for studying the generalized slowing hypothesis in children with specific language impairment. *J. Speech Hear. Res.* 37, 418–421.
- Kiefer, M., and Trumpp, N. M. (2012). Embodiment theory and education: the foundations of cognition in perception and action. *Trends Neurosci. Educ.* 1, 15–20. doi: 10.1016/j.tine.2012.07.002
- Kontra, C., Goldin-Meadow, S., and Beilock, S. L. (2012). Embodied learning across the lifespan. *Top. Cogn. Sci.* 4, 731–739. doi: 10.1111/j.1756-8765.2012.01221.x
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34. doi: 10.1037/a0021446
- Laakso, A. (2011). Embodiment and development in cognitive science. *Cogn. Brain Behav.* 15, 409–425.
- Louwerse, M. M., and Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition* 114, 96–104. doi: 10.1016/j.cognition.2009.09.002
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004
- Marley, S. C., Levin, J. R., and Glenberg, A. M. (2010). What cognitive benefits does an activity-based reading strategy afford young Native American readers? *J. Exp. Educ.* 78, 395–417. doi: 10.1080/00220970903548061
- McNeil, N. M., Uttal, D. H., Jarvin, L., and Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learn. Instruct.* 19, 171–184. doi: 10.1016/j.learninstruc.2008.03.005
- Meltzoff, A. N. (1990). Towards a developmental cognitive Science – The implications of cross-modal matching and imitation for the development of representation and memory in infancy. *Ann. N. Y. Acad. Sci.* 608, 1–37. doi: 10.1111/j.1749-6632.1990.tb48889.x
- Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. doi: 10.1016/j.cortex.2010.11.002
- Morse, A. F., Belpaeme, T., Cangelosi, A., and Smith, L. B. (2010). “Thinking with your body: modelling spatial biases in categorization using a real humanoid robot,” in *Proceedings of the 2010 Annual Meeting of the Cognitive Science Society*, Portland, 1362–1368.
- Myung, J.-Y., Blumstein, S. E., and Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: manipulation knowledge of objects. *Cognition* 98, 223–243. doi: 10.1016/j.cognition.2004.11.010
- O'Neill, D. K., Topolovec, J., and Stern-Cavalcante, M. (2002). Feeling sponginess: the importance of descriptive gestures in 2- and 3-year-old children's acquisition of adjectives. *J. Cogn. Dev.* 3, 243–277. doi: 10.1207/S15327647JC.D0303\_1
- Paivio, A. (1990). *Mental Representations: A Dual Coding Approach*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195066661.001.0001
- Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: Norton & Co. doi: 10.1037/11494-000
- Pulvermuller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbol semantics. *Trends Cogn. Sci.* 17, 458–470. doi: 10.1016/j.tics.2013.06.004
- Pulvermuller, F., and Garagnani, M. (2014). From sensorimotor learning to memory cells in prefrontal and temporal association cortex: a neurocomputational study of disembodiment. *Cortex* 57C, 1–21. doi: 10.1016/j.cortex.2014.02.015
- Scofield, J., Hernandez-Reif, M., and Keith, A. B. (2009). Preschool children's multimodal word learning. *J. Cogn. Dev.* 10, 306–333. doi: 10.1080/15248370903417662
- Siakaluk, P. D., Pexman, P. M., Aguilera, L., Owen, W. J., and Sears, C. R. (2008). Evidence for the activation of sensorimotor information during visual word recognition: the body-object interaction effect. *Cognition* 106, 433–443. doi: 10.1016/j.cognition.2006.12.011
- Sidhu, D. M., Kwan, R., Pexman, P. M., and Siakaluk, P. D. (2014). Effects of relative embodiment in lexical and semantic processing of verbs. *Acta Psychol.* 149, 32–39. doi: 10.1016/j.actpsy.2014.02.009
- Smeets, D. J. H., and Bus, A. G. (2012). Interactive electronic storybooks for kindergartners to promote vocabulary growth. *J. Exp. Child Psychol.* 112, 36–55. doi: 10.1016/j.jecp.2011.12.003
- Smith, L. B. (2005). Action alters shape categories. *Cogn. Sci.* 29, 665–679. doi: 10.1207/s15516709cog0000\_13
- Smith, L. B. (2013). It's all connected: pathways in visual object recognition and early noun learning. *Am. Psychol.* 68, 618–629. doi: 10.1037/a0034185
- Smith, L. B., Maouene, J., and Hidaka, S. (2007). “The body and children's word learning,” in *The Emerging Spatial Mind*, eds J. M. Plumert and J. P. Spencer (New York: Oxford University Press), 168–192. doi: 10.1093/acprof:oso/9780195189223.003.0008
- Smith, L. B., and Samuelson, L. (2010). “Objects in space and mind: from reaching to words,” in *Thinking Through Space: Spatial Foundations of Language and Cognition*, eds K. Mix, L. B. Smith, and M. Gasser (Oxford: Oxford University Press).
- Stanfield, R. A., and Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* 12, 153–156. doi: 10.1111/1467-9280.00326
- Tare, M., Chiong, C., Ganea, P., and DeLoache, J. (2010). Less is more: how manipulative features affect children's learning from picture books. *J. Appl. Dev. Psychol.* 31, 395–400. doi: 10.1016/j.appdev.2010.06.005
- Thelen, E. (2008). “Grounded in the world: developmental origins of the embodied mind,” in *Developmental Perspectives on Embodiment and Consciousness*, eds W. F. Overton, U. Muller, and J. L. Newman (New York: Taylor & Francis Group), 99–130.
- Tillotson, S. M., Siakaluk, P. D., and Pexman, P. M. (2008). Body-object interaction ratings for 1,618 monosyllabic nouns. *Behav. Res. Methods* 40, 1075–1078. doi: 10.3758/BRM.40.4.1075
- Tousignant, C., and Pexman, P. M. (2012). Flexible recruitment of semantic richness: context modulates body-object interaction effects in lexical-semantic processing. *Front. Hum. Neurosci.* 6:53. doi: 10.3389/fnhum.2012.00053
- Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., and DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Dev. Perspect.* 3, 156–159. doi: 10.1111/j.1750-8606.2009.00097.x
- van Dam, W. O., Rueschemeyer, S., Lindemann, O., and Bekkering, H. (2010). Context effects in embodied lexical-semantic processing. *Front. Cogn.* 1:150. doi: 10.3389/fpsyg.2010.00150
- Visser, J. (2003). Developmental coordination disorder: a review of research on subtypes and comorbidities. *Hum. Mov. Sci.* 22, 479–493. doi: 10.1016/j.humov.2003.09.005

- Warrington, E. K., and Shallice, T. (1984). Category-specific semantic impairments. *Brain* 107, 829–853. doi: 10.1093/brain/107.3.829
- Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., and Song, H.-J. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Dev. Perspect.* 7, 155–159. doi: 10.1111/cdep.12032
- Wellsby, M., and Pexman, P. M. (in press). The influence of bodily experience on children's language processing. *Top. Cogn. Sci.*
- Willems, R. M., and Francken, J. C. (2012). Embodied cognition: taking the next step. *Front. Psychol.* 3:582. doi: 10.3389/fpsyg.2012.00582
- Wilson, A. D., and Golonka, S. (2013). Embodied cognition is not what you think it is. *Front. Psychol.* 4:58. doi: 10.3389/fpsyg.2013.00058
- Yuruten, O., Sahin, E., and Kalkan, S. (2013). The learning of adjectives and nouns from affordance and appearance features. *Adapt. Behav.* 21, 437–451. doi: 10.1177/1059712313497976
- Zdravilova, L., and Pexman, P. M. (2013). Grasping the invisible: semantic processing of abstract words. *Psychon. Bull. Rev.* 20, 1312–1318. doi: 10.3758/s13423-013-0452-x
- Zwaan, R. A. (2014). Embodiment and language comprehension: reframing the discussion. *Trends Cogn. Sci.* 18, 229–234. doi: 10.1016/j.tics.2014.02.008
- Zwaan, R. A., and Kaschak, M. P. (2009). "Language in the brain, body, and world," in *The Cambridge Handbook of Situated Cognition*, eds P. Robbins and M. Aydede (New York: Cambridge University Press), 368–381.

**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.

Received: 31 January 2014; accepted: 08 May 2014; published online: 28 May 2014.

Citation: Wellsby M and Pexman PM (2014) Developing embodied cognition: insights from children's concepts and language processing. *Front. Psychol.* 5:506. doi: 10.3389/fpsyg.2014.00506

This article was submitted to Cognitive Science, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Wellsby and Pexman. 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.



# A perceptual account of symbolic reasoning

David Landy<sup>1\*</sup>, Colin Allen<sup>2\*</sup> and Carlos Zednik<sup>3</sup>

<sup>1</sup> Psychological and Brain Science/Cognitive Science, Indiana University, Bloomington, IN, USA

<sup>2</sup> History and Philosophy of Science/Cognitive Science, Indiana University, Bloomington, IN, USA

<sup>3</sup> Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Guy Dove, University of Louisville, USA

Robert Douglas Rupert, University of Colorado at Boulder, USA

## \*Correspondence:

David Landy, Psychological and Brain Science/Cognitive Science, Indiana University, 107 S Indiana Ave., Bloomington, IN 47405, USA  
e-mail: dlandy@indiana.edu;

Colin Allen, History and Philosophy of Science/Cognitive Science, Indiana University, 107 S Indiana Ave., Bloomington, IN 47405, USA  
e-mail: colallen@indiana.edu

People can be taught to manipulate symbols according to formal mathematical and logical rules. Cognitive scientists have traditionally viewed this capacity—the capacity for *symbolic reasoning*—as grounded in the ability to internally represent numbers, logical relationships, and mathematical rules in an abstract, amodal fashion. We present an alternative view, portraying symbolic reasoning as a special kind of embodied reasoning in which arithmetic and logical formulae, externally represented as notations, serve as targets for powerful perceptual and sensorimotor systems. Although symbolic reasoning often conforms to abstract mathematical principles, it is typically implemented by perceptual and sensorimotor engagement with concrete environmental structures.

**Keywords:** human reasoning, formal logic, mathematics, embodied cognition, perception

## INTRODUCTION

How do people reason arithmetically, algebraically, and logically? One well-known answer to this question holds that the human mind trades in inner symbols that amodally represent abstract arithmetic, algebraic, and logical propositions, and manipulates these symbols according to internally represented mathematical and logical rules. On this traditional view, the “inner” takes precedence over the “outer”: notations on paper, computer screens, and classroom blackboards are involved in mathematical problem-solving only insofar as they are “translated” into corresponding mental structures and processes.

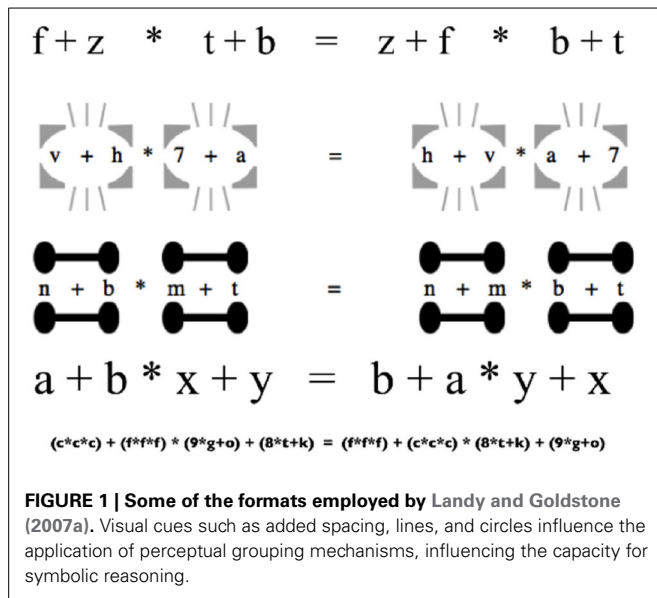
Suppose you hold such a traditional view, but then learn that stray marks and subtle changes in spacing can lead otherwise competent students of algebra to “forget” a basic rule such as operator precedence. Several recent experiments have demonstrated just this sort of influence of visual structure on algebraic performance. One example comes from Landy and Goldstone (2007a), who gave college undergraduates simple algebraic forms, such as “ $a + b * c + d = c + d * a + b$ ,” and asked them to decide whether or not the given symbols described a valid equation (see **Figure 1**). Because the expressions contained both additions and multiplications, determining their validity required respecting the order of operations, which stipulates that multiplications precede additions. By creating artificial visual groups (e.g., by manipulating the physical spacing of equations, or by introducing shapes into the surrounding context as depicted in **Figure 1**), participants’ performance could be predictably manipulated: validity-judgments were more likely to be correct if visual groupings were in line with valid operator precedence. Nor is this pattern restricted to algebraic validity. Related research has indicated that spatial layout impacts application of the order of

operations rules when calculating (Kirshner, 1989; Landy and Goldstone, 2010), when creating story problems (Jiang et al., in press), and when working in programming languages such as Python (Hansen et al., unpublished manuscript).

How might you interpret this sort of behavioral pattern? You could chalk failure to respect operator precedence, for example, up to performance error, and remain committed to the thesis that the underlying mathematical competence is largely independent of the way notational structures are perceived and physically manipulated. Alternatively, you could wonder whether competence with operator precedence depends non-trivially on the perceptual and sensorimotor mechanisms that target those external notations. To what extent might these mechanisms be responsible not just for our mathematical mistakes, but also for our successes?

The ability to follow operator-precedence rules is just one manifestation of the capacity for *symbolic reasoning*: the capacity to manipulate arbitrary symbolic tokens according to abstract mathematical and logical rules. In what follows, we propose an account of symbolic reasoning according to which perception, manipulation, and perceptual imagination lie at the heart of mathematical and logical competence. Rather than rely on amodally represented rules, symbolic reasoners make their mathematical judgments using perceptual processes that have no obvious link to the following of formal mathematical rules. Instead, we identify the capacity for symbolic reasoning with the ability to perceptually group, detect symmetry in, and otherwise perceptually organize symbolic notations as they are experienced in the environment. On this view, the kinds of behavioral patterns described above are typical: not only does written format impact the legibility of symbols, it also impacts the application





of well-known rules. When notational expressions afford active manipulation, symbolic reasoning is often accomplished by physically interacting with those notations. In contrast, when notations do not afford physical manipulation or perceptual processing, symbolic reasoning may involve processes of visual, aural, and even tactile imagination. Although symbolic reasoning can therefore become “internalized,” it remains rooted in mechanisms close to the sensorimotor periphery.

Although we will emphasize the kinds of algebra, arithmetic, and logic that are typically learned in high school, our view also potentially explains the activities of advanced mathematicians—especially those that involve representational structures like graphs and diagrams. Our major goal, therefore, is to provide a novel and unified account of both successful and unsuccessful episodes of symbolic reasoning, with an eye toward providing an account of mathematical reasoning in general. Before turning to our own account, however, we begin with a brief outline of some more traditional views.

## EXTANT ACCOUNTS OF SYMBOLIC REASONING

### COMPUTATIONALISM AND SEMANTIC PROCESSING: TRANSLATIONAL ACCOUNTS OF SYMBOLIC REASONING

Two prominent accounts of symbolic reasoning can be introduced via an analogy from the classroom. Consider the different ways in which students might be taught to think about the following syllogism:

All dogs are mammals;  
All mammals are animals;  
Therefore, all dogs are animals.

On one hand, students can think about such problems *syntactically*, as a specific instance of the more general logical form “All Xs are Ys; All Ys are Zs; Therefore, all Xs are Zs.” On the other hand, they might think about them *semantically*—as relations between subsets, for example. In an analogous fashion, two prominent

scientific attempts to explain how students are able to solve symbolic reasoning problems can be distinguished according to their emphasis on syntactic or semantic properties.

Analogous to the syntactic approach above, *computationalism* holds that the capacity for symbolic reasoning is carried out by mental processes of syntactic rule-based symbol-manipulation. In its canonical form, these processes take place in a general-purpose “central reasoning system” that is functionally encapsulated from dedicated and modality-specific sensorimotor “modules” (Fodor, 1983; Sloman, 1996; Pylyshyn, 1999; Anderson, 2007). Although other versions of computationalism do not posit a strict distinction between central and sensorimotor processing, they do generally assume that sensorimotor processing can be safely “abstracted away” (e.g., Kemp et al., 2008; Perfors et al., 2011). On all computationalist accounts, when an individual is confronted with a symbolic reasoning task such as a natural-language “word problem” or a formal reasoning problem expressed in the notational formalisms of algebra, calculus, and logic, the perception of notations in the environment causes a tokening of equivalent symbols and expressions of “Mentalese” (Fodor, 1975). These mental symbols and expressions are then operated on by syntactic rules that instantiate mathematical and logical principles, and that are typically assumed to take the form of productions, laws, or probabilistic causal structures (Newell and Simon, 1976; Sloman, 1996; Anderson, 2007). Once a solution is computed, it is converted back into a publicly observable (i.e., written or spoken) linguistic or notational formalism.

An influential alternative to computationalism is analogous to the semantic approach to the syllogism above: the heterogeneous family of *semantic processing* accounts, according to which symbolic reasoning is carried out by systems that interpret and represent meaningful mathematical and logical relations. Accounts of this type differ according to the particular representational formats they posit, ranging from amodal or generically spatial “mental models” (Johnson-Laird et al., 1992), to rich perceptual and sensorimotor “simulations” of specific objects and scenes (Barsalou, 1999), and even to indirect “conceptual metaphors” that drive people’s intuitions and conclusions about a specific mathematical problem (Lakoff and Núñez, 2000). What distinguishes these accounts from computationalism is the idea that symbolic reasoning occurs not on the basis of syntactic rules, but on the basis of meaningful interpretations of a particular mathematical or logical task domain. For example, Lakoff and Núñez argue that real-number concepts are derived from experiences with physical lengths, and that the capacity for simple arithmetic arises from an innate ability to estimate and compare such lengths. On Johnson-Laird’s “mental models” account, symbolic reasoning problems are solved by “inspecting” a mental model of the problem: the validity of “ $a \& b \therefore b$ ” can be determined by recognizing that “ $b$ ” is a component of the model for “ $a \& b$ .” In much the same way, Barsalou’s “perceptual symbol systems” account suggests that logical expressions are interpreted by mentally simulating concrete scenarios to which the expression applies: a scene that includes both an apple and an orange includes an orange.

Despite their differences, computationalist, and semantic processing accounts share the assumption that processes of

perception and action play a relatively limited role in the process of symbolic reasoning. Although both accounts acknowledge that the perception of notations is important for the construction of internal representations, they also assume that once such representations have been constructed, the physical notations that express the original mathematical or logical problem may be ignored or altogether discarded until a solution is communicated. Notably, this even applies to accounts which, like Barsalou's, posit a special role for sensorimotor representations in general, yet attribute a curiously limited role to sensorimotor representations of the notations that are actually perceived while a symbolic reasoning task is being performed. In general, computationalist and semantic processing accounts are alike in being essentially *translational*: they suppose that processes of perception and action do little other than mediate between notational structures in the external environment and the internal structures and processes in which symbolic reasoning *really* occurs.

It is worth elaborating on this translational aspect. The capacity for symbolic reasoning is expressed behaviorally by converting an input representation of a mathematics or logic problem into an output representation of a corresponding solution. Initially, the problem is represented in a public language, either as a natural-language "word problem", or in the special notational systems designed for algebra, calculus, and logic. Eventually, this problem representation is converted into a written or spoken solution. But exactly how does this conversion occur? Like many other kinds of problem solving, the process of symbolic reasoning can be seen as a chain of transformations that links input and output representations, each of which changes its format and/or semantic structure. Some transformations, such as " $a$  and  $b$ " to " $a \& b$ ," involve a change in format without a change in semantic structure. In contrast, transformations such as " $\sim(\sim a \vee \sim b) \therefore b$ " to " $a \& b \therefore b$ " involve changes in format *and* semantic structure: the resulting representation is a simplification of the original problem.

Computationalist and semantic processing accounts of symbolic reasoning are equally translational because they both assume that problem representations are passed from a perceptual apparatus to an internal processing system in a form that is no simpler than the external (notational or linguistic) problem representation. That is, they assume that all transformations that involve changes in semantic structure take place "internally," over Mentalese expressions, mental models, metaphors or simulations, and that sensorimotor interactions with physical notations involve (at most) a change in representational format. On these accounts, when a subject is asked to evaluate a formal expression such as " $\sim(\sim a \vee \sim b) \therefore b$ ," a mental representation of that expression must be constructed before it can be simplified to " $a \& b \therefore b$ ." Similarly, notational variants of one-and-the-same proposition—e.g., " $All\ Fs\ are\ Gs$ ," " $(x)(Fx \rightarrow Gx)$ ," and " $\forall x[Fx \supset Gx]$ " will be converted into one-and-the-same Mentalese expression, mental model, metaphor or simulation. In general, therefore, computationalist and semantic processing accounts of symbolic reasoning rely equally on the assumption that the principal role of sensorimotor processes—the processes that govern the perception of and physical interaction with public symbols and expressions—is simply to provide inputs to and carry outputs from those internal structures and processes that

are ultimately responsible for performing all substantial steps in a mathematical or logical problem solving chain.

### TOWARD A CONSTITUTIVE ACCOUNT: THE CYBORG VIEW

Translational accounts of symbolic reasoning can be distinguished from *constitutive* accounts, in which sensorimotor mechanisms are not merely part of the causal chain that links external notations to internal representations, but are crucially involved in transforming the problem representation into one that has a simplified semantic structure. Recall that on the translationist view, mental resources can be divided into those that "translate" the outer situation into a generally isomorphic inner representation, and those that act on that representation to solve the problem. On a constitutive account, sensorimotor mechanisms not only translate the problem, they are involved in the transformations that substantively solve it. One prominent view that can be associated with such a constitutive approach might, to borrow Andy Clark's terminology, be called the *cyborg view* of symbolic reasoning (Clark, 2003). Grounded on recent work in the area of "situated cognition," the cyborg view holds that notations constitute external technological artifacts that "scaffold" the biological processes involved in symbolic reasoning (Clark, 1997, 1998, 2006; Menary, 2007; Sutton, 2010). This "scaffolding" is typically achieved by notations that permit the extraneural storing, inspection, deletion and manipulation of information in a way that facilitates the execution of symbolic reasoning tasks, and has positive effects on the speed and accuracy with which these tasks can be performed as well as their potential complexity. To cite a well-known example, "carrying" a digit during a complex multiplication task by writing it on a piece of paper, adding it to the result and then crossing it out obviates the need to store and manipulate that digit in biological memory, thereby freeing up valuable cognitive resources, minimizing possible error from misremembering, and permitting the multiplication of extremely large values. One way of explaining the cognitive benefit of such "scaffolding" is to view notations as constitutive parts of integrated, boundary-crossing symbolic reasoning systems: When computing " $123 \times 89$ ," "carrying" the tens digit of the temporary product " $3 \times 9$ " and adding it to the units digit of " $2 \times 9$ " transforms the original complex multiplication problem into a series of simpler multiplication and addition problems that can easily be done in the head. Thus, the active manipulation of physical notations plays the role of "guiding" the human biological machinery through an abstract mathematical problem space—one that may far exceed the space of otherwise solvable problems.

While emphasizing the ways in which notations are acted upon, however, proponents of the cyborg view rarely consider how such notations are perceived. Sometimes, this neglect is intentional, as when the utility of cognitive artifacts is explained by stating that they become assimilated into a "body schema" in which "sensorimotor capacities function without... the necessity of perceptual monitoring" (Gallagher, 2005, p. 25). At other times, this neglect seems to be unintended, however, and subject to corrective elaboration. For example, although Andy Clark (1998, p. 168) argues that the human ability to deploy and manipulate notations in symbolic reasoning tasks "involves the use of the same old (essentially pattern-completing) resources

to model the special kinds of behavior observed in the public [notational] world,” it remains unclear exactly *which* pattern-completing resources are in play, and what kinds of patterns they complete. In general, therefore, although cyborg theorists have shown quite successfully that notations can be constitutively involved in symbolic reasoning, and have made great strides in cataloguing the kinds of bodily interactions that lead to cognitive success, few specific details have emerged regarding the relevant perceptual processes that facilitate these interactions, as well as the physical characteristics that determine when and why a particular notation is cognitively beneficial.

Consider how such details might explain the influence of visual structure on algorithmic reasoning discussed earlier. Order of operations behavior need not be implemented in a set of high-level productions or in a collection of explicit memorized rules, but also need not be determined by active manipulations of physical notations. Instead, such behavior might largely depend on visual processes that segment the scene into parts, wholes, and groups. One possibility is that because the algebraic system tends to align spatial structure and precedence rules, perceptual grouping processes acquire biases compatible with those rules (Kirshner and Awtry, 2004); another is that because proofs tend to maintain tightly bound structures, leading to increased statistical regularity in high precedence operations, experience with algebraic derivations modifies perceptual organization. Other regular cultural cues have long been known to impact grouping (Wertheimer, 1923/1938). By extending the cyborg view’s emphasis on environmental interaction with a detailed understanding of perceptual processing, a theoretical framework might be developed that accounts for the effect of aligning visual grouping and syntactic binding discussed earlier (see **Figure 1**), but that may also explain many other episodes of formally correct and incorrect symbolic reasoning.

In what follows, we articulate a constitutive account of symbolic reasoning, *Perceptual Manipulations Theory*, that seeks to elaborate on the cyborg view in exactly this way. While accommodating the cyborg view’s emphasis on the active manipulation of physical notations, Perceptual Manipulations Theory additionally emphasizes the perceptual processes that facilitate and govern such manipulations, as well as the physical characteristics of particularly successful (and unsuccessful) notational formalisms. On our view, the way in which physical notations are perceived is at least as important as the way in which they are actively manipulated.

## PERCEPTUAL MANIPULATIONS THEORY

### THE THEORY

Perceptual Manipulations Theory (PMT) goes further than the cyborg account in emphasizing the perceptual nature of symbolic reasoning. External symbolic notations need not be translated into internal representational structures, but neither does all mathematical reasoning occur by manipulating perceived notations on paper. Rather, complex visual and auditory processes such as affordance learning, perceptual pattern-matching and perceptual grouping of notational structures produce simplified representations of the mathematical problem, simplifying the task faced by the rest of the symbolic reasoning system. Perceptual

processes exploit the typically well-designed features of physical notations to automatically reduce and simplify difficult, routine formal chores, and so are themselves constitutively involved in the capacity for symbolic reasoning. Moreover, if a particular symbolic reasoning problem cannot be solved by perceptual processing and active manipulation of physical notations alone, subjects often invoke detail-rich sensorimotor representations that closely resemble the physical notations in which that problem was originally encountered. On our view, therefore, much of the capacity for symbolic reasoning is implemented as the perception, manipulation and modal and cross-modal representation of externally perceived notations.

The neural processes that PMT takes to be involved in symbolic reasoning almost never have as their primary function the implementation of amodally represented rules or models. Instead, they include sensorimotor systems for visual grouping and perceptual organization, object recognition, object tracking and symmetry detection, among others. Although skills such as object-recognition may appear quintessentially “cognitive” to some, we treat them as sensorimotor capacities to highlight the fact that, rather than apply to abstract mathematical or logical entities, they apply directly to the physical properties of notations in the environment such as shape, relative spacing and position. Indeed, insofar as most mathematical and logical notations are well-designed, these properties are frequently suggestive of how they ought to be manipulated, thus promoting formally valid “symbol-pushing”. For example, the fact that the multiplicands in “ $xy + z$ ” are closer to one another than to the additive term can be understood as a manifestation of the order-of-operations rule that multiplication is to be performed before addition—a manifestation that is immediately recognized by mechanisms of perceptual grouping (see section Evidence for Perceptual Manipulations Theory). Notably, such sensorimotor competences are often more robust than the formal systems to which they are applied: while a formula such as “ $((P \rightarrow ((Q \& R)))$ ” would be rejected by a machine following strict well-formedness rules, even beginning logic students interpret it as a conditional, and must be explicitly trained by pedagogues with ulterior motives to focus on a narrower set of structural elements. As we discuss in greater detail below, a wide range of (correct *and* incorrect) mathematical behavior can be attributed to the way the perceived details of formal notations “interlock” with domain-general sensorimotor capacities.

Perceptual Manipulations Theory suggests that most symbolic reasoning emerges from the ways in which notational formalisms are perceived and manipulated. Nevertheless, direct sensorimotor processing of physical stimuli is augmented by the capacity to imagine and manipulate mental representations of notational markings. Faculties of spatial reasoning, mental transformation, referential symbolism and a rich set of capacities for acquiring and imagining physical behaviors such as walking, pointing, writing, and erasing can all be used to internally reproduce the actual perceived details of physical notations and to mentally manipulate them in ways that resemble physical actions. Insofar as our account emphasizes perceptual representations of formal notations and imagined notation-manipulations, it can be contrasted with Barsalou’s perceptual symbol systems account, in which

“people often construct non-formal simulations to solve formal problems” (Barsalou, 1999, 606). Moreover, our emphasis differs from standard “conceptual metaphor” accounts, which suggest that formal reasoners rely on a “semantic backdrop” of embodied experiences and sensorimotor capacities to interpret abstract mathematical concepts. Our account is probably closest to one articulated by Dörfler (2002), who like us emphasizes the importance of treating elements of notational systems as physical objects rather than as meaning-carrying symbols.

Although there are clear differences between PMT and other accounts of symbolic reasoning, our view incorporates elements from many of them—albeit with a greater emphasis on perception. For illustration, consider a student already competent in logic now learning set theory. The perceivable physical similarities of  $\cap$  and  $\cup$  to  $\wedge$  and  $\vee$ , including the up-down symmetry between each pair, serve as a *perceptual*, rather than conceptual, metaphor. To see how this metaphor may be applied, consider the duality principle that

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

which bears a striking visual similarity to De Morgan’s law,

$$\overline{P \vee Q} \equiv \bar{P} \wedge \bar{Q}$$

This visual similarity is partially a result of common symbolology, including the use of capital letters for elements, the use of horizontal lines for equality, the use of bars for negation, and the above-mentioned use of similar shapes for basic operations. Partially, though, the similarity results from the arrangement of these parts—if one is written in prefix notation, for instance, the similarity is markedly decreased (it is beyond the scope of this work to attempt a general definition of similarity; for a review, see Goldstone and Son, 2005). For a student learning a new formal system, these notational similarities ground the transformations typical to set theory by mapping them onto the more familiar domain of logic, facilitating the application of similar principles and ideas, and licensing particular manipulations, sometimes even prior to obtaining a rich understanding of the conceptual issues involved. To the degree that these inferences are licensed, learning may be facilitated. Although the relevant perceptual and sensorimotor processes are modality-specific, when mathematical notations are well-designed, human mathematical competence can be incredibly flexible: radically different mathematical and logical propositions can be treated in similar formal ways because of similarities in the way in which they are physically manifested as notations. Of course, it is not always or often the case that capturing visual and semantic regularities across domains is the explicit goal of mathematicians introducing notation (though see Smaill, 2012, for one apparent case). We predict, however, that when there are significant visual similarities in notations used across domains, people will tend to import assumptions from a well-understood domain into a novel one.

Perceptual Manipulations Theory also posits a novel psychological role for much-discussed magnitude- and quantity-detection systems. Visual quantity (e.g., the number of blocks, dots, or sheep presented in a drawing or on a computer screen)

is often thought to be directly represented by an evolved “number system” dedicated to amodal magnitude representation (Gelman and Gallistel, 1978; Barth et al., 2003; Dehaene et al., 2004; Machery, 2007). It has been argued that such quantity-sensitive mechanisms provide the basic representational vehicles over which formal mathematical reasoning occurs (Gallistel et al., 2005; Spelke, 2005; Carey, 2009), but PMT holds a more textured view. Quantity-sensitive mechanisms certainly sometimes represent numbers. In symbolic reasoning tasks, however, a primary function of magnitude and quantity-detection systems is to enable reasoners to track magnitude and quantity properties of notational formalisms. For example, when dealing with large numbers such as “3,000,000,” magnitude-detection plays a role in keeping track of the number of digits (Hinrichs et al., 1982). Similarly, when teaching a rule such as the product rule captured by “ $a^5 a^3 = a^8$ ,” a teacher may write something like “(aaaaa)  $\times$  (aaa) = (aaaaaaaa)” and let magnitude-detection (and explicit counting) systems do the rest. Thus, a significant portion of the verification process may be implemented by perceptual and sensorimotor skills and quantity-detection systems that process the notational formalism itself, without necessarily interpreting the notation’s meaning.

The emphasis that PMT places on domain-general systems for perceptual processing and bodily interaction with physical systems of notations underscores the importance of the historical development of a common set of well-designed mathematical notations. Although historically the development of visual commonalities across notations may have been largely accidental, this development has served mathematics well, providing visual cues that allow the human perceptual and motor systems to effectively operate over them. One prediction of PMT is that when notations align perceptual and structural similarities, learning will be facilitated. Of course, when they misalign, as they sometimes do, learning is predicted to be impaired (Marquis, 1988 discusses several such cases). Still, better notation systems could yet be constructed in all branches of formal reasoning to take full advantage of visual cues that automatically “steer” the reasoner in the direction of formally valid solutions. In this way, the human capacity for symbolic reasoning winds up being ordinary, bodily situatedness in novel, artifactual sensorimotor space: the space of (well-designed!) notations.

## EVIDENCE FOR PERCEPTUAL MANIPULATIONS THEORY

Most of the existing literature on symbolic reasoning has been developed using an implicitly or explicitly translational perspective. Although we do not believe that the current evidence is enough to completely dislodge this perspective, it does show that sensorimotor processing influences the capacity for symbolic reasoning in a number of interesting and surprising ways. The translational view easily accounts for cases in which individual symbols are more readily perceived based on external format. For example, blurring symbols will make them harder to perceive. Perceptual Manipulations Theory also predicts this sort of impact, but further predicts that perceived structures will affect the application of rules—since rules are presumed to be implemented via systems involved in perceiving that structure. In this section, we will review several empirical



sources of evidence for the impact of visual structure on the implementation of formal rules. Although translational accounts may eventually be elaborated to accommodate this evidence, it is far more easily and naturally accommodated by accounts which, like PMT, attribute a constitutive role to perceptual processing.

Perceptual Manipulations Theory holds that skill with symbol systems is implemented in alignments between elements of external notations and perceptual and motor systems. Therefore, it predicts that the physical appearance of notations should strongly influence formal behavior. For example, it should be difficult to differentially respond to two similar-looking notational forms even if they are conceptually dissimilar. Substantial evidence suggests that this prediction holds. For example, Kirshner and Awtry (2004) show that the common mistake of confusing the valid rule regarding multiplication of two like terms by adding their exponents ( $a^n \times a^m = a^{n+m}$ ) with the visually similar but invalid rule regarding added terms ( $a^n + a^m = a^{n+m}$ ) can be avoided by teaching students a linguistic notation in which these equations no longer resemble one another. In the same way, common mistakes such as

$$\frac{a}{x} + \frac{b}{y} = \frac{a+b}{x+y}$$

can be prevented just by changing the notational format in which they are learned (see Marquis, 1988 for several examples of visual patterns in algebra). The frequency of these mistakes—as well as the fact that they can be prevented by switching notational formats—are hard to explain from a translational perspective in which perceived problems are converted into inner propositions or models, and in which formal dissimilarity ought to trump visual similarity. In contrast, they are quite easily explained from a perspective that attributes a constitutive role to perceptual processing. What appears to be happening is that students apply a very general maxim of perceptual pattern learning: if two things look similar, similar things can probably be done with them, and if they look different, they require different actions. Although this is not a formally valid way of reasoning over symbol systems (and indeed, often leads to the mistakes reported above), this general strategy may lead to correct solutions whenever visual similarity *does* mirror formal similarity (see also Cohen Kadosh, 2009). Indeed, such mirroring is widespread, and appears to be regularly exploited by reasoners. Consider the way algebraic notation aligns formal structure with perceptual grouping in the expression

$$\frac{a+b}{a+bc}.$$

Here, formal structure is mirrored in the visual grouping structure created both by the spacing ( $b$  and  $c$  are multiplied, then added to  $a$ ) and by the physical demarcation of the horizontal line. Instead of applying abstract mathematical rules to process such expressions, Landy and Goldstone (2007a,b see also Kirshner, 1989) propose that reasoners leverage visual grouping strategies to directly segment such equations into multi-symbol visual chunks. To test this hypothesis, they investigated the way

manipulations of visual groups affect participants' application of operator precedence rules. Maruyama et al. (2012) argue on the basis of fMRI and MEG evidence that mathematical expressions like these are parsed quickly by visual cortex, using mechanisms that are shared with non-mathematical spatial perception tasks.

Interestingly, perceptual processes play a role not only in the way notations are perceived, but also in the way they are created. By studying beginning logic students' physical arrangement of logical formulae in an online natural deduction tutoring system (Allen and Menzel, 2007), Landy and Goldstone (2007b) found statistically significant patterns of space-insertion consistent with the hypothesis that spaces are used to aid visual grouping within logical formulae. That is, reasoners not only exploit visual groups that are already present in the physical representation of a symbolic reasoning task, but also actively and endogenously reproduce such groups when they make it easier to find a solution. But why do reasoners insert such formally irrelevant features to their written notational formalisms? From a translational perspective, this question is difficult to answer: once a solution to a symbolic reasoning problem is computed, it merely needs to be translated into a public language, one in which the observed space-insertion patterns are formally irrelevant. From the perspective of PMT, however, it seems likely that such patterns either derive from the possibility that mathematical and logical equations are internally encoded in a perceptually-rich format in which details about spacing is retained, or from the utility of such patterns in computing intermediate solutions on paper by applying the same visual object-segmentation systems that were initially used to interpret the problem. Supporting the possibility that spatial structure plays a crucial role in the process of interpretation of equations, Jiang et al. (in press) report that subjects inventing story problems match the physical structure of provided equations.

The visual system is well-known to be particularly responsive to dynamic stimuli such as motion. This is reflected in the apparent relevance of motion and transformation in algebraic understanding of proofs. Nogueira de Lima and Tall (2007) documented that schoolchildren learning algebra often treat transformations such as

$$x + b = y - m$$

$$x = y - m - b$$

not as the repeated application of formal Euclidean axioms, but as “magic motion,” in which a term moves to the other side of the equation and “flips” sign. Landy and Goldstone (2009) suggest that this reference to motion is no mere metaphor. Subjects with significant training in calculus found it easier to solve problems of this form when an irrelevant field of background dots moved in the same direction as the variables, than when the dots moved in the contrary direction.

One suggestion of PMT is that mathematical concepts may be encoded using multiple strategies, and that perceptual-motor strategies may emerge over the process of using a symbol system. As an example, Varma and Schwartz (2011) examine



the case of negative number acquisition, and in particular the acquisition of processes allowing the comparison of positive and negative numbers. Initially, learners are faster at comparing numbers that are close together when one is positive and the other negative—a reversal of the usual *distance effect* that holds with positive numbers (Moyer and Landauer, 1967)—but one that is consistent with a rule-based strategy involving comparing signs. More expert learners show a typical size effect, so that numbers that are ‘far apart’ are discriminated more quickly. The authors suggest that negative numbers are initially processed by children using rules, but that “symbolic manipulation can transform an existing magnitude representation so that it incorporates additional perceptual-motor structure.”

In summary, PMT suggests that learning how to perceptually and physically engage notations is critical to the capacity for reasoning in accordance with their mathematical meanings. To be successful, learners must discover which aspects of a notation are relevant and meaningfully aligned with mathematical rules and concepts, and must then acquire an appropriately “rigged up” sensorimotor system (see also: Goldstone et al., 2010). Although the sensorimotor skillset required for sophisticated symbolic reasoning is likely to be highly developed and available to learners only after some struggle (Piaget, 1953; Bednarz et al., 1996), Kellman et al. (2008) have already found that training students to recognize algebraic expressions using standard perceptual learning techniques leads to lasting gains both in equation reading and comprehension, as well as in algebraic problem-solving. Indeed, substantial evidence indicates that notation systems that align with computationally useful processes are relatively easy to acquire across a variety of domains including arithmetic and algebra (Kirshner and Awtry, 2004; Landy and Goldstone, 2007c), electric circuit design (Cheng, 1999), and sequence and grammar learning (Pothos et al., 2006; Endress et al., 2007). Our account expects such results because appropriate alignment between the formal and the perceptual significantly simplifies the search for correct solutions. Although we will not speculate extensively about possible implications for mathematics education, results such as these also suggest that the PMT approach can be a productive way to think about new pedagogical approaches to designing and reasoning with formal notations. In particular, it seems likely that the most effective and easily-learned notations and rule-systems are the ones that have greatest alignment with preexisting or easily learned perceptual and sensorimotor routines. On our view, one principal virtue of well-structured notation systems is that they leverage automatic sensorimotor operations by making their products formally useful, and the better the alignment between the formal and the sensorimotor, the more useful those products will be.

## THEORETICAL IMPLICATIONS

### IS THERE A “FUNDAMENTAL” MATHEMATICAL REASONING SYSTEM?

A contribution of PMT is that it provides a novel account of how to bring mathematical and logical reasoning into the fold of embodied cognition more generally. Although PMT accommodates the cyborg view and its emphasis of the environment, it adds

a detailed conception of the constitutive role of perceptual processing in symbolic reasoning: perception is at least as important as physical manipulation. One consequence of this view is that mathematical and logical reasoning need not be rooted in single, special-purpose cognitive mechanisms. Although we do not deny the existence of amodal numerosity or magnitude detection systems, our account does not assign those systems a uniquely fundamental role in the development of mathematical reasoning capacities. Instead, on our view symbolic reasoning is carried out by a wide variety of perceptual and motor skills, including fast numerosity and magnitude evaluation; repeatable actions like pointing, counting, and stacking; object segmentation and grouping; motion detection and visualization; writing and reading; and many other sensorimotor skills. Additionally, it seems reasonable to assume that the same sensorimotor skillset may also play a pivotal role in other mathematical domains such as geometry and category theory, the elementary portions of which both of which rely considerably on diagrams and other iconic notations. More controversially perhaps, since all areas of mathematics and symbolic reasoning involve—at some point—the learning of rules and abstract principles via notational systems, it may even be the case that the same perceptual and motor processes that implement the capacity for symbolic reasoning also play different but equally fundamental roles in implementing various kinds of abstract reasoning in mathematics and beyond. Whether this leaves any significant role for amodal systems remains to be seen, but see Dove (in press) for an argument for representational pluralism.

A corollary of the claim that symbolic and other forms of mathematical and logical reasoning are grounded in a wide variety of sensorimotor skills is that symbolic reasoning is likely to be both idiosyncratic and context-specific. For one, different individuals may rely on different embodied strategies, depending on their particular history of experience and engagement with particular notational systems. For another, even a single individual may rely on different strategies in different situations, depending on the particular notations being employed at the time. Some of the relevant strategies may cross modalities, and be applicable in various mathematical domains; others may exist only within a single modality and within a limited formal context. For example, consider the fact that there is significant potential for error when a successful strategy in one domain is exported to another domain—as, for example, when beginning logic students make the mistake of distributing a negation across a conjunction, going from  $\sim(X \& Y)$  to  $(\sim X \& \sim Y)$ , because they perceive a similarity to the algebraically legal manipulation of  $-(x + y)$  to  $(-x + y)$ . Although in this particular case such cross-domain mapping leads to a formal error, it need not always be mistaken—as when understanding that “ $\sim\sim X$ ” is equivalent to “ $X$ ,” just as “ $- -x$ ” is equal to “ $x$ .” In some contexts, such perceptual strategies lead to mathematical success. In other contexts, however, the same strategies lead to mathematical failure.

If the capacity for symbolic reasoning is in fact idiosyncratic and context-dependent in the way suggested here, what are the implications for scientific psychology? PMT implies that the “deep” facts about human mathematical, algebraic, logical,

and other mathematical abilities are unlikely to be facts about inner computations and models, but are instead facts about how humans manage to exploit perceptual and sensorimotor strategies in appropriate, context-specific ways—and about how they fall prey to these strategies when applying them inappropriately. The reason that mathematicians have the intuition that people who are merely “pushing symbols” are failing to grasp fundamental mathematical meanings is that they are indeed failing to do so—though this failure may be more widespread, and indeed more powerful, than mathematicians and psychologists have previously assumed. Being more specific than this, however, seems difficult. Therefore, the key to understanding the human capacity for symbolic reasoning in general will be to characterize *typical* sensorimotor strategies, and to understand the particular conditions in which those strategies are successful or unsuccessful.

### WHAT IS MATHEMATICAL RULE-FOLLOWING AND WHO IS THE MATHEMATICAL RULE-FOLLOWER?

Perceptual Manipulations Theory claims that symbolic reasoning is implemented over interactions between perceptual and motor processes with real or imagined notational environments. Since symbolic reasoning involves manipulating symbols and expressions according to mathematical and logical rules, this view implies that the human ability to follow *abstract* mathematical and logical rules is carried out by sensorimotor processes that apply to *concrete*—i.e., readily perceivable and physically manipulatable—notations. But how is it that “primitive” sensorimotor processes can give rise to some of the most sophisticated mathematical behaviors? Unlike many traditional accounts, PMT does not presuppose that mathematical and logical rules must be internally represented in order to be followed. Rather, overt rule-following emerges from the fine-tuned interactions between the perceptual and sensorimotor systems with well-designed physical notations—symbolic reasoning is a form of sophisticated “symbol pushing” that *happens* to adhere to the formal rules of mathematics and logic, due to a lengthy process of cultural adaptation and pedagogical scaffolding.

Like interlocking puzzle pieces that together form a larger image, sensorimotor mechanisms and physical notations “interlock” to produce sophisticated mathematical behaviors. Insofar as mathematical rule-following emerges from active engagement with physical notations, the mathematical rule-follower is a distributed system that spans the boundaries between brain, body, and environment. For this interlocking to promote mathematically appropriate behavior, however, the relevant perceptual and sensorimotor mechanisms must be just as well-trained as the physical notations must be well-designed. Thus, on one hand, the development of symbolic reasoning abilities in an individual subject will depend on the development of a sophisticated sensorimotor skillset in the way outlined above. On the other hand, the development of symbolic reasoning abilities within a society will depend on the availability of notational formalisms that promote formally valid “symbol-pushing.” Indeed, the development of mathematical expertise is often historically contemporaneous with the development of powerful, efficient, and easily learned

systems of formal mathematical and logical notation (Dantzig, 1954; Stedall, 2007).

### CONCLUSION

We have described an approach to symbolic reasoning which closely ties it to the perceptual and sensorimotor mechanisms that engage physical notations. We argued for this approach on the basis of empirical evidence that shows algebraic and mathematical knowledge to be surprisingly fragile in the face of minor perceivable differences, and on the basis of evidence that suggests that competent symbolic reasoners typically rely on semantically irrelevant properties of notational formulae in order to quickly and accurately—but also sometimes inaccurately—solve symbolic reasoning problems. With respect to this evidence, PMT compares favorably to traditional “translational” accounts of symbolic reasoning.

Nevertheless, there is probably no uniquely correct answer to the question of how people do mathematics. Indeed, it is important to consider the relative merits of all competing accounts and to incorporate the best elements of each. Just as the particular sensorimotor strategies being invoked are likely to differ across individuals and situations, it is also likely that different episodes of symbolic reasoning require different explanations—be they in terms of comparisons based on conceptual metaphors, situated interactions with notations, or even conscious applications of formal rules. Although we believe that most of our mathematical abilities are rooted in our past experience and engagement with notations, we do not depend on these notations at all times. Moreover, even when we do engage with physical notations, there is a place for semantic metaphors and conscious mathematical rule following. Therefore, although it seems likely that abstract mathematical ability relies heavily on personal histories of active engagement with notational formalisms, this is unlikely to be the story as a whole. It is also why non-human animals, despite in some cases having similar perceptual systems, fail to develop significant mathematical competence even when immersed in a human symbolic environment. Although some animals have been taught to order a small subset of the numerals (less than 10) and carry out simple numerosity tasks within that range, they fail to generalize the patterns required for the indefinite counting that children are capable of mastering, albeit with much time and effort. If we consider the working memory requirements for noticing that the pattern \_\_\_-ty one, \_\_\_-ty two, \_\_\_-ty three, etc. repeats after “twen-,” “thir-,” “for-,” and so on, then it may not seem so unlikely that only a species with a rather large brain could even notice let alone generalize the pattern. And without that basis for understanding the domain and range of symbols to which arithmetical operations can be applied, there is no basis for further development of mathematical competence.

Although we have not accounted for forms of mathematical reasoning beyond symbolic reasoning except in passing, the account of mathematical rule-following suggested here points toward the possibility that processes of perception, visualization, and interaction may play a crucial constitutive role in mathematical and logical reasoning in general. Unlike more established views, many of which acknowledge the utility of

mathematical notations as concise representations of abstract mathematical meanings but then go on to downplay their importance for symbolic reasoning proper, PMT suggests that notations and the sensorimotor processes that engage them are often at the very heart of high-level mathematical and logical cognition. In this vein, since many forms of advanced mathematical reasoning rely on graphical representations and geometric principles, it would be surprising to find that perceptual and sensorimotor processes are *not* involved in a constitutive way. Therefore, by accounting for symbolic reasoning—perhaps the most abstract of all forms of mathematical reasoning—in perceptual and sensorimotor terms, we have attempted to lay the groundwork for an account of mathematical and logical reasoning more generally. The potential for a satisfying unification of the successes and failures of human symbolic and other forms of mathematical reasoning under a common set of mechanisms provides us with the confidence to claim that this is a topic worthy of further investigation, both empirical and philosophical.

## REFERENCES

- Allen, C., and Menzel, C. P. (2007). *The Logic Daemon and Quizmaster*. Available online at: <http://logic.tamu.edu/>. (Accessed April 9, 2014)
- Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195324259.001.0001
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.
- Barth, H., Kanwisher, N., and Spelke, E. S. (2003). The construction of large number representations in adults. *Cognition* 86, 201–221. doi: 10.1016/S0010-0277(02)00178-6
- Bednarz, N., Kieran, C., and Lee, L. (1996). *Approaches to Algebra: Perspectives for Research and Teaching*. Dordrecht: Kluwer Academic Publishers. doi: 10.1007/978-94-009-1732-3
- Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195367638.001.0001
- Cheng, P. C. (1999). Interactive law encoding diagrams for learning and instruction. *Learn. Instr.* 9, 309–325. doi: 10.1016/S0959-4752(98)00049-8
- Clark, A. (1997). *Being There*. Cambridge, MA: MIT Press.
- Clark, A. (1998). “Magic words: how language augments human computation,” in *Language and Thought: Interdisciplinary Themes*, eds P. Carruthers and J. Boucher (Cambridge: Cambridge University Press), 162–183.
- Clark, A. (2003). *Natural Born Cyborgs*. Oxford: Oxford University Press.
- Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends Cogn. Sci.* 10, 370–374. doi: 10.1016/j.tics.2006.06.012
- Cohen Kadosh, R. (2009). Numerical representation in the parietal lobes: abstract or not abstract? *Behav. Brain Sci.* 32, 313–328. doi: 10.1017/S0140525X09990938
- Dantzig, T. (1954). *Number: The Language of Science*. 4th Edn. Garden City, NY: Doubleday Anchor Books.
- Dehaene, S., Molko, N., Cohen, L. and Wilson, A. J. (2004). Arithmetic and the brain. *Curr. Opin. Neurobiol.* 14, 218–224. doi: 10.1016/j.conb.2004.03.008
- Dörfler, W. (2002). Formation of mathematical objects as decision making. *Math. Think. Learn.* 4, 337–350. doi: 10.1207/S15327833MTL0404\_03
- Dove, G. (in press). Thinking in words: language as an embodied medium of thought. *Topics Cogn. Sci.*
- Endress, A. D., Dehaene-Lambertz, G., and Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition* 105, 577–614. doi: 10.1016/j.cognition.2006.12.014
- Fodor, J. A. (1975). *The Language of Thought*. New York, NY: Crowell.
- Fodor, J. A. (1983). *Modularity of Mind*. Cambridge, MA: MIT Press.
- Gallagher, S. (2005). *How the Body Shapes The Mind*. Oxford: Oxford University Press. doi: 10.1093/0199271941.001.0001
- Gallistel, C. R., Gelman, R. and Cordes, S. (2005). “The cultural and evolutionary history of the real numbers,” in *Evolution and Culture*, eds S. C. Levinson and P. Jaisson Cambridge, MA: MIT Press.
- Gelman, R. and Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.
- Goldstone, R. L., Landy, D., and Son, J. Y. (2010). The education of perception. *Topics Cogn. Sci.* 2, 265–284. doi: 10.1111/j.1756-8765.2009.01055.x
- Goldstone, R. L., and Son, J. Y. (2005). “Similarity,” in *Cambridge Handbook of Thinking and Reasoning*, eds K. J. Holyoak and R. G. Morrison (Cambridge: Cambridge University Press), 13–36.
- Hansen, M., Goldstone, R. L., and Lumsdaine, A. (unpublished manuscript). What makes code hard to understand? Preprint: arXiv:1304.5257v2.
- Hinrichs, J. V., Berie, J. L., and Mosell, M. K. (1982). Place information in multidigit number comparison. *Mem. Cogn.* 10, 487–495. doi: 10.3758/BF03197652
- Jiang, M. J., Cooper, J. L., and Alibali, M. W. (in press). Spatial factors influence arithmetic performance: the case of the minus sign. *Q. J. Exp. Psychol.*
- Johnson-Laird, P. N., Byrne, R. M. J., and Shaeken, W. (1992). Propositional reasoning by model. *Psychol. Rev.* 99, 418–439. doi: 10.1037/0033-295X.99.3.418
- Kellman, P. J., Massey, C. M., Roth, Z., Burke, T., Zucker, J., Saw, A., et al. (2008). Perceptual learning and the technology of expertise: fraction learning and algebra. *Cogn. Pragmatics Spec. Issue Cogn. Technol.* 16, 356–405. doi: 10.1075/p&c.16.2.07kel
- Kemp, C., Goodman, N. D., and Tenenbaum, J. B. (2008). “Theory acquisition and the language of thought,” in *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, eds B. C. Love, K. McRae, and V. M. Sloutsky (Austin, TX: Cognitive Science Society), 1606–1611.
- Kirshner, D. (1989). The visual syntax of algebra. *J. Res. Math. Educ.* 20, 274–287. doi: 10.2307/749516
- Kirshner, D., and Awtry, T. (2004). Visual salience of algebraic transformations. *J. Res. Math. Educ.* 35, 224–257. doi: 10.2307/30034809
- Lakoff, G., and Núñez, R. E. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York, NY: Basic Books.
- Landy, D., and Goldstone, R. L. (2007a). How abstract is symbolic thought? *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 720–733. doi: 10.1037/0278-7393.33.4.720
- Landy, D., and Goldstone, R. L. (2007b). Formal notations are diagrams: evidence from a production task. *Mem. Cogn.* 35, 2033–2040. doi: 10.3758/BF03192935
- Landy, D., and Goldstone, R. L. (2007c). “How space guides interpretation of a novel mathematical system,” in *The 29th Annual Conference of the Cognitive Science Society*. Nashville, TX.
- Landy, D., and Goldstone, R. L. (2009). “Pushing symbols,” *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Amsterdam.
- Landy, D., and Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Q. J. Exp. Psychol.* 63, 1953–1968. doi: 10.1080/17470218.2012.656664
- Machery, E. (2007). Concept empiricism: a methodological critique. *Cognition* 104, 19–46. doi: 10.1016/j.cognition.2006.05.002
- Marquis, J. (1988). “Common mistakes in algebra,” in *The Ideas of Algebra, K-12: 1988 Yearbook*, eds A. F. Coxford and A. P. Shulte (Reston, VA: National Council of Teachers of Mathematics), 204–205.
- Maruyama, M., Pallier, C., Jobert, A., Sigman, M., and Dehaene, S. (2012). The cortical representation of simple mathematical expressions. *Neuroimage* 61, 1444–1460. doi: 10.1016/j.neuroimage.2012.04.020
- Menary, R. (2007). Writing as thinking. *Lang. Sci.* 29, 621–632. doi: 10.1016/j.langsci.2007.01.005
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Newell, A., and Simon, H. A. (1976). Computer science as empirical inquiry: symbols and search. *Commun. ACM* 19, 113–126. doi: 10.1145/360018.360022
- Nogueira de Lima, R., and Tall, D. (2007). Procedural embodiment and magic in linear equations. *Educ. Stud. Math.* 67, 3–18. doi: 10.1007/s10649-007-9086-0
- Perfors, A., Tenenbaum, J. B., and Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition* 118, 306–338. doi: 10.1016/j.cognition.2010.11.001
- Piaget, J. (1953). *The Origins of Intelligence in Children*. London: Routledge and Kegan Paul.
- Pothos, E. M., Chater, N., and Ziori, E. (2006). Does stimulus appearance affect learning? *Am. J. Psychol.* 119, 277–301. doi: 10.2307/20445339
- Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav. Brain Sci.* 22, 341–423. doi: 10.1017/S0140525X99002022
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychol. Bull.* 119, 3–22. doi: 10.1037/0033-2909.119.1.3
- Smaill, A. (2012). “Mathematical notation and analogy,” in *Symposium on Mathematical Practice and Cognition II. Paper Presented at AISB/IACAP World*

- Congress (30-31), eds A. Pease, and B. Larvor (Birmingham: The Society for the Study of Artificial Intelligence and Simulation of Behavior).
- Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? *Am. Psychol.* 60, 950–958. doi: 10.1037/0003-066X.60.9.950
- Stedall, J. (2007). Symbolism, combinations, and visual imagery in the mathematics of Thomas Harriot. *Hist. Math.* 34, 380–401. doi: 10.1016/j.hm.2007.05.001
- Sutton, J. (2010). “Exograms and interdisciplinarity: history, the extended mind, and the civilizing process,” in *The Extended Mind*, ed R. Menary (Cambridge, MA: MIT Press), 189–226. doi: 10.7551/mitpress/9780262014038.003.0009
- Varma, S., and Schwartz, D. L. (2011). The mental representation of integers: an abstract-to-concrete shift in the understanding of mathematical concepts. *Cognition* 121, 363–385. doi: 10.1016/j.cognition.2011.08.005
- Wertheimer, M. (1923/1938). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung* 4, 301–350. Translation published in Ellis, W. (1938). *A Source Book of Gestalt Psychology* (London: Routledge & Kegan Paul), 71–88.

**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.

Received: 11 December 2013; accepted: 14 March 2014; published online: 21 April 2014.

Citation: Landy D, Allen C and Zednik C (2014) A perceptual account of symbolic reasoning. *Front. Psychol.* 5:275. doi: 10.3389/fpsyg.2014.00275

This article was submitted to Cognitive Science, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Landy, Allen and Zednik. 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.





# NIRS in motion—unraveling the neurocognitive underpinnings of embodied numerical cognition

**Julia Bahnmüller<sup>1,2\*</sup>, Thomas Dresler<sup>3,4</sup>, Ann-Christine Ehlis<sup>3,4</sup>, Ulrike Cress<sup>2,3</sup> and Hans-Christoph Nuerk<sup>1,2,3</sup>**

<sup>1</sup> Department of Psychology, University of Tuebingen, Tuebingen, Germany

<sup>2</sup> Knowledge Media Research Center, Tuebingen, Germany

<sup>3</sup> LEAD Graduate School, University of Tuebingen, Tuebingen, Germany

<sup>4</sup> Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany

\*Correspondence: julia.bahnmueller@uni-tuebingen.de

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

David Hamilton Landy, Indiana University, USA

Guy Dove, University of Louisville, USA

**Keywords: embodied cognition, functional near-infrared spectroscopy (fNIRS), numerical cognition, motion, imaging techniques**

The central representation of numerical cognition is commonly considered an abstract magnitude representation serving as one key precursor for higher mathematical thinking. However, recent research indicates that the representation might not be purely abstract. In fact, accumulating evidence suggests that numerical representations are rooted in and shaped by specific motor activities and sensory-bodily experiences and, therefore, are influenced by so-called embodied numerical representations. If we want to understand how numerical understanding develops, it is crucial to elucidate the basic cognitive tools with which we develop a sense of number. We argue that it is necessary to address this issue on both a behavioral and a neural level.

Contrasting the view of functional magnetic resonance imaging (fMRI) being the generally preferable neuroimaging technique, we argue that particularly in embodied cognition, restrictions and benefits of different imaging methods should guide the chosen research question. In our opinion, near-infrared spectroscopy (NIRS) is optimally suited to investigate embodied cognition paradigms that explicitly involve motion. In the following, recent research will be outlined showing that numerical cognition is not purely abstract, but influenced by embodied representations. NIRS will then be introduced as a feasible technique for the investigation of embodied cognitions. Since research in this domain is largely restricted to the

perception of embodied experiences, but fails to address motion itself, we will finally argue that NIRS offers a good opportunity to fill this research gap.

## EMBODIED NUMERICAL COGNITION: WHERE WE ARE

Embodied cognition refers to the idea that, throughout our lifespan, we consistently associate specific motor activities and sensory-bodily experiences with more abstract concepts such as words or numbers (Barsalou, 2008). In numerical cognition, a growing body of research indicates that number is a prime example of such embodied cognitions. To clearly separate embodied numerical cognition from other related concepts influencing the way we learn, represent and deal with numbers, Fischer (2012; Fischer and Brugger, 2011) distinguishes grounded, situated and embodied numerical cognition. Grounded numerical cognition means that universal laws of the physical world are reflected in our representation of numbers (i.e., small numbers are associated with lower space whereas large numbers are associated with upper space). Situated numerical cognition refers to the idea that situations (including external stimuli as well as our body posture) influence how we process numbers. In this vein, Loetscher et al. (2008) demonstrated that turning the head to the right resulted in the production of larger random numbers than turning it to the left. In contrast, embodied numerical cognition

relates to repeated, culturally dependent learning experiences directly associating representations of number with specific motor activities or other bodily-sensory experiences.

Several research branches in embodied numerical cognition have addressed questions of automaticity, directionality, functionality, developmental aspects as well as the generality of embodied numerical representations on behavior. A rather concrete link between number magnitude and embodied numerosity is investigated in the most prominent example of embodied numerical cognition: finger counting. There is behavioral evidence that the activation of this association is (i) automatic (e.g., Klein et al., 2011), (ii) already evident in childhood (Domahs et al., 2008), and persisting into adulthood (e.g., Di Luca et al., 2006; Klein et al., 2011). Furthermore, the association is (iii) culturally dependent (e.g., Domahs et al., 2010), and (iv) dependent on the spatial representation of numerical magnitude (cf. the mental number line; e.g., Fischer, 2008; Lindemann et al., 2011). A second branch supporting and generalizing findings from finger counting is grasping. Research on grasping shows that the association between the number magnitude representation and grasping actions is also (i) automatic (Andres et al., 2004, 2008; Lindemann et al., 2007; Ranzini et al., 2011) and additionally (ii) bidirectional, meaning both that the representation of number magnitude influences

grasping actions (e.g., Andres et al., 2004; Badets et al., 2007) and vice-versa (Badets and Pesenti, 2010; Badets et al., 2010). It is important to note that, with few exceptions, behavioral studies on finger counting and grasping do not include actual actions except for response giving. Rather, the association of numerical magnitude and fingers/hands is achieved by perceptual and mostly static cues (e.g., finger postures on a screen). This does not mean that these studies do not give important insights into the association of number magnitude and finger-/hand-based representations. However, questions of how and why the association develops, how and where it originates and whether or not it is influenceable remain open.

To investigate the functional relevance, generality and variability of associations of number magnitude and embodied representations, first correlation and training studies have been conducted. Concerning finger counting, finger-based representations are functionally relevant for numerical development: Noël (2005) showed that finger gnosis predicts future numerical skills. Moreover, a systematic training of finger gnosis ameliorated numerical performance (e.g., Gracia-Bafalluy and Noël, 2008).

Few studies investigating the neurocognitive underpinnings of embodied numerical representations are available. Usually, neural correlates (the “where”) of associations of finger- as well as hand-related and numerical magnitude representations have been studied in perceptual studies. fMRI studies demonstrated that cortical areas related to number magnitude processing are in close proximity to areas activated when fingers/hands are used. Additionally, it has been shown that applying transcranial magnetic stimulation (TMS) to the left angular gyrus resulted in a disruption of finger schema and number processing (e.g., Rusconi et al., 2005). Evidence for automatic activation of fingers when dealing with numbers comes from an fMRI study (Tschemtscher et al., 2012). They found an automatically coactivated right motor cortex when small numbers were passively viewed by subjects habitually starting finger counting with the left hand. Developmental changes were indicated by Kaufmann et al. (2008) showing that the coactivation of finger- and number-related

areas is more pronounced for children as compared to adults. Taken together, similar to behavioral results, neural evidence supports the existence and automaticity of an association of finger-/hand-based and numerical magnitude representations. However, mostly perceptual and static cues have been used, which prevents direct insight into questions of how and why the association is established, how it originates and whether or not we can or should promote this association to increase learning outcomes. Moreover, the assumption that findings from perceptual paradigms can be generalized to actually motion-involving paradigms still needs to be tested.

### **SPATIAL FULL-BODY MOVEMENTS SUBSERVING ABSTRACT SEMANTIC REPRESENTATIONS**

There are first studies showing that not only finger but also body movements corresponding to the spatial representation of numbers influence number processing. Here, it is important to keep in mind that every bodily movement is accompanied by a spatial processing component (e.g., whenever we turn our head in a certain direction, we not only have a bodily-sensory experience but also a spatial one). Following the idea that the coactivation of a corresponding embodied-spatial representation of number might promote the development of the numerical magnitude representation, training studies were designed aiming at enriched training conditions that use (full) body movements and are congruent with a left-to-right orientation of the mental number line. Indeed, first studies indicate that such embodied-spatial numerical trainings are more effective than non-embodied control trainings and even show specific transfer effects (Fischer et al., 2011; Link et al., 2013; for an overview, see Moeller et al., 2012). This suggests that embodied numerical cognitions are not restricted to fingers/hands but generalize to more complex bodily experiences that relate to the spatial representation of numerical magnitude. Furthermore, it indicates that we can and should use the association of number and embodied-spatial representations to support mathematical learning, meaning that we should use motion and bodily-sensory experiences that correspond to

the semantic and spatial representation of numerical magnitude. However, as motion can hardly be executed using the most prominent imaging methods, the neural mechanisms of how the association is established and of how learning is supported by such an enriched learning environment has not yet been addressed. Here, NIRS represents a good opportunity to begin filling this research gap.

### **IMAGING METHODS AND MOTION**

Several different neuroimaging tools are available, each with obvious benefits, but with specific shortcomings as well. Therefore, when choosing one of them, the respective research question needs to be considered. For instance, fMRI with its high spatial resolution enables detailed insights into the function of both superficial and deep neural structures. Electroencephalography (EEG) with its high sampling rate allows for a precise investigation of temporal processes. Both are prone to motion artifacts and rely on a rather restrictive measurement setting—not a major problem for most paradigms, but an enormous limitation for some. Therefore, the investigation of processes inherently relying on motion or body postures remains a challenge. To address research questions that involve motoric and sensory (embodied) processes underlying seemingly abstract numerical cognition, we argue that NIRS is a good alternative.

NIRS measures cortical oxygenation by emitting near-infrared light to the brain using optical properties of oxy- and deoxygenated hemoglobin, thus providing an optical blood-oxygen-level dependent (BOLD) signal (Obrig and Villringer, 1997; Ferrari and Quaresima, 2012). Light emitters and detectors are attached to the head in various arrangements resulting in a grid of measurement channels. NIRS as neuroimaging method has been introduced two decades ago and was applied in various settings since (e.g., Ehlis et al., 2014). It is non-stationary and applicable as bedside technique and in other real-world settings (e.g., class room, Dresler et al., 2009). NIRS allows the measurement of subjects in more natural positions and under several body postures. For instance, one well-known study compared apple peeling with NIRS to mock

apple peeling in an fMRI-fNIRS setting (Okamoto et al., 2004). It could be shown that the natural action resulted in a different activation pattern as compared to the mocked action, illustrating the feasibility of NIRS as an imaging device that can easily be used during daily activities as well as embodied cognition paradigms.

## EMBODIMENT AND NIRS: A NEXT STEP

Research in embodied numerical cognition has addressed different research questions to show that embodied numerical representations do influence number processing and numerical learning. Nevertheless, it mostly focused on paradigms that do not include actual and intended motion but rather static, perceptual cues. On a neural level, we argue that NIRS—despite its sub-optimal spatial resolution—is especially suited to help filling this gap. So far, studies successfully using NIRS have been conducted both for different motor tasks (for an overview see Leff et al., 2011) and in the field of numerical cognition investigating different paradigms with adults (e.g., Richter et al., 2009; Cutini et al., 2014) and children (e.g., Dresler et al., 2009; Hyde et al., 2010). Combining those research branches, NIRS offers the possibility for examining embodied numerical representations as it allows measuring brain activity during natural movements and in ecologically more valid settings. Integrating online measures of embodied numerical representations can add to a more elaborated picture of the neurocognitive underpinnings of the interplay, origin and development of abstract and embodied numerical representations.

We are aware of existing problems that need to be addressed in NIRS methodology. Although NIRS is less prone to movement artifacts as compared to EEG and fMRI, it does not mean that it is not affected at all. When the head is moved, small agitations of the sensors reduce data quality as direction of light flow is changed. Different analytic approaches have been suggested and are used in practice to deal with movement (see Cui et al., 2010; Brigadoi et al., 2014). Furthermore, in recent years, remote NIRS devices have been developed which have already been applied during bicycle riding

(Piper et al., 2014) and allow for greater freedom of motion than the common non-remote systems.

Additionally, NIRS has a low spatial resolution (both lateral resolution and penetration depth) depending on the distance between optodes. Therefore, studies asking for a fine-grained topological resolution of a particular region (e.g., Harvey et al., 2013) can currently not be investigated using NIRS. Unraveling the neuronal basis of higher cognitive processes such as (embodied) numerical cognition does, however, not solely rely on spatially high resolving devices and research questions. Nonetheless, ongoing methodological progress is made in overcoming shortcomings in terms of spatial resolution (for an overview see, e.g., Ferrari and Quaresima, 2012; Scholkmann et al., 2014). In terms of lateral resolution, already available high-density arrangements offer a much higher resolution when compared to broadly used continuous wave systems. Considering penetration depth, continuous wave systems only allow for measuring cortical structures located few centimeters under a respective emitter-detector-channel. Improved depth penetration can be achieved by frequency- and time-domain instrumentation allowing for a clearer separation between extra- and intracerebral oxygenation changes. Methodological research points in a promising direction and we are convinced that the availability of these higher-resolution devices will increase in the next years and will enlarge the feasibility of NIRS even further.

Considering new evidence in research in embodied numerical cognition as well as technological developments, we are convinced that NIRS will add to a more elaborated picture of neurocognitive underpinnings of embodied cognition and to a broader understanding of the basis of numerical cognition as well.

## ACKNOWLEDGMENTS

Julia Bahnmüller was supported by a graduate scholarship of the state of Baden-Wuerttemberg (Landesgraduiertenförderung). Thomas Dresler is funded by the LEAD Graduate School [GSC1028], a project of the Excellence Initiative of the German federal and state governments. This research was

supported in part by a project within the ScienceCampus (WissenschaftsCampus) Tuebingen (Cluster 8/TP 4) and by funding of the German Research Foundation (DFG; CR 110/8-1) to Ulrike Cress, Hans-Christoph Nuerk und Ann-Christine Ehlis.

## REFERENCES

- Andres, M., Davare, M., Pesenti, M., Olivier, E., and Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport* 15, 2773–2777.
- Andres, M., Ostry, D. J., Nicol, F., and Paus, T. (2008). Time course of number magnitude interference during grasping. *Cortex* 44, 414–419. doi: 10.1016/j.cortex.2007.08.007
- Badets, A., Andres, M., Di Luca, S., and Pesenti, M. (2007). Number magnitude potentiates action judgements. *Exp. Brain. Res.* 180, 525–534. doi: 10.1007/s00221-007-0870-y
- Badets, A., and Pesenti, M. (2010). Creating number semantics through finger movement perception. *Cognition* 115, 46–53. doi: 10.1016/j.cognition.2009.11.007
- Badets, A., Pesenti, M., and Olivier, E. (2010). Response-effect compatibility of finger-numeral configurations in arithmetical context. *Q. J. Exp. Psychol.* 63, 16–22. doi: 10.1080/17470210903134385
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Brigadoi, S., Ceccherini, L., Cutini, S., Scarpa, F., Scatturin, P., Selb, J., et al. (2014). Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *Neuroimage* 85, 181–191. doi: 10.1016/j.neuroimage.2013.04.082
- Cui, X., Bray, S., and Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* 49, 3039–3046. doi: 10.1016/j.neuroimage.2009.11.050
- Cutini, S., Scatturin, P., Basso Moro, S., and Zorzi, M. (2014). Are the neural correlates of subitizing and estimation dissociable? An fNIRS investigation. *Neuroimage* 85, 391–399. doi: 10.1016/j.neuroimage.2013.08.027
- Di Luca, S., Graná, A., Semenza, C., Seron, X., and Pesenti, M. (2006). Finger-digit compatibility in Arabic numerical processing. *Q. J. Exp. Psychol.* 59, 1648–1663. doi: 10.1080/17470210500256839
- Domahs, F., Krinzinger, H., and Willmes, K. (2008). Mind the gap between both hands: evidence for internal finger-based number representations in children's mental calculation. *Cortex* 44, 359–367. doi: 10.1016/j.cortex.2007.08.001
- Domahs, F., Moeller, K., Huber, S., Willmes, K., and Nuerk, H.-C. (2010). Embodied numerosity: implicit hand-based representations influence symbolic number processing across cultures. *Cognition* 116, 251–266. doi: 10.1016/j.cognition.2010.05.007
- Dresler, T., Obersteiner, A., Schecklmann, M., Vogel, A. C. M., Ehlis, A.-C., Richter, M. M., et al. (2009). Arithmetic tasks in different formats and their

- influence on behavior and brain oxygenation as assessed with near-infrared spectroscopy (NIRS): a study involving primary and secondary school children. *J. Neural Transm.* 116, 1689–1700. doi: 10.1007/s00702-009-0307-9
- Ehlis, A.-C., Schneider, S., Dresler, T., and Fallgatter, A. J. (2014). Application of functional near-infrared spectroscopy in psychiatry. *Neuroimage* 85, 478–488. doi: 10.1016/j.neuroimage.2013.03.067
- Ferrari, M., and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 63, 921–935. doi: 10.1016/j.neuroimage.2012.03.049
- Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex* 44, 386–392. doi: 10.1016/j.cortex.2007.08.004
- Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cogn. Process.* 13, 161–164. doi: 10.1007/s10339-012-0477-5
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260
- Fischer, U., Moeller, K., Bientzle, M., Cress, U., and Nuerk, H.-C. (2011). Sensori-motor spatial training of number magnitude representation. *Psychon. Bull. Rev.* 18, 177–183. doi: 10.3758/s13423-010-0031-3
- Gracia-Bafalluy, M., and Noël, M. P. (2008). Does finger training increase young children's numerical performance? *Cortex* 44, 368–375. doi: 10.1016/j.cortex.2007.08.020
- Harvey, B. M., Klein, B. P., Petridou, N., and Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science* 341, 1123–1126. doi: 10.1126/science.1239052
- Hyde, D., Boas, D., Blair, C., and Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *Neuroimage* 53, 647–652. doi: 10.1016/j.neuroimage.2010.06.030
- Kaufmann, L., Vogel, S. E., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L. B., et al. (2008). A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex* 44, 376–385. doi: 10.1016/j.cortex.2007.08.003
- Klein, E., Moeller, K., Willmes, K., Nuerk, H. C., and Domahs, F. (2011). The influence of implicit hand-based representations on mental arithmetic. *Front. Psychol.* 2:197. doi: 10.3389/fpsyg.2011.00197
- Leff, D. R., Orihuela-Espina, F., Elwell, C. E., Athanasiou, T., Delpy, D. T., Darzi, A. W., et al. (2011). Assessment of the cerebral cortex during motor task behaviours in adults: a systematic review of functional near infrared spectroscopy (fNIRS) studies. *Neuroimage* 54, 2922–2936. doi: 10.1016/j.neuroimage.2010.10.058
- Lindemann, O., Abolafia, J. M., Girardi, G., and Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1400–1409. doi: 10.1037/0096-1523.33.6.1400
- Lindemann, O., Alipour, A., and Fischer, M. H. (2011). Finger counting habits in middle eastern and western individuals: an online survey. *J. Cross Cult. Psychol.* 42, 566–578. doi: 10.1177/0022022111406254
- Link, T., Moeller, K., Huber, S., Fischer, U., and Nuerk, H.-C. (2013). Walk the number line – an embodied training of numerical concepts. *Trends Neurosci. Educ.* 2, 74–84. doi: 10.1016/j.tine.2013.06.005
- Loetscher, T., Schwarz, U., Schubiger, M., and Brugger, P. (2008). Head turns bias the brain's random number generator. *Curr. Biol.* 18, R60–R62. doi: 10.1016/j.cub.2007.11.015
- Moeller, K., Fischer, U., Link, T., Wasner, M., Huber, S., Cress, U., et al. (2012). Learning and development of embodied numerosity. *Cog. Process.* 13, 271–274. doi: 10.1007/s10339-012-0457-9
- Noël, M. P. (2005). Finger gnosis: a predictor of numerical abilities in children? *Child Neuropsychol.* 11, 413–430. doi: 10.1080/09297040590951550
- Obrig, H., and Villringer, A. (1997). Near-infrared spectroscopy in functional activation studies. Can NIRS demonstrate cortical activation? *Adv. Exp. Med. Biol.* 413, 113–127. doi: 10.1007/978-1-4899-0056-2\_13
- Okamoto, M., Dan, H., Shimizu, K., Takeo, K., Amita, T., Oda, I., et al. (2004). Multimodal assessment of cortical activation during apple peeling by NIRS and fMRI. *Neuroimage* 21, 1275–1288. doi: 10.1016/j.neuroimage.2003.12.003
- Piper, S. K., Krueger, A., Koch, S. P., Mehnert, J., Habermehl, C., Steinbrink, J., et al. (2014). A wearable multi-channel fNIRS system for brain imaging in freely moving subjects. *Neuroimage* 85, 64–71. doi: 10.1016/j.neuroimage.2013.06.062
- Ranzini, M., Lugli, L., Anelli, F., Carbone, R., Nicoletti, R., and Borghi, A. M. (2011). Grasable objects shape number processing. *Front. Hum. Neurosci.* 5:147. doi: 10.3389/fnhum.2011.00147
- Richter, M. M., Zierhut, K. C., Dresler, T., Plichta, M. M., Ehlis, A. C., Reiss, K., et al. (2009). Changes in cortical blood oxygenation during arithmetical tasks measured by near-infrared spectroscopy. *J. Neural Transm.* 116, 267–273. doi: 10.1007/s00702-008-0168-7
- Rusconi, E., Walsh, V., and Butterworth, B. (2005). Dexterity with numbers: rTMS over left angular gyrus disrupts finger gnosis and number processing. *Neuropsychologia* 43, 1609–1624. doi: 10.1016/j.neuropsychologia.2005.01.009
- Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., et al. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 85, 6–27. doi: 10.1016/j.neuroimage.2013.05.004
- Tschentscher, N., Hauk, O., Fischer, M. H., and Pulvermüller, F. (2012). You can count on the motor cortex: finger counting habits modulate motor cortex activation evoked by numbers. *Neuroimage* 59, 3139–3148. doi: 10.1016/j.neuroimage.2011.11.037

**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.

Received: 25 February 2014; accepted: 25 June 2014; published online: 18 July 2014.

Citation: Bahnmüller J, Dresler T, Ehlis A-C, Cress U and Nuerk H-C (2014) NIRS in motion—unraveling the neurocognitive underpinnings of embodied numerical cognition. *Front. Psychol.* 5:743. doi: 10.3389/fpsyg.2014.00743

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Bahnmüller, Dresler, Ehlis, Cress and Nuerk. 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.





# The specificity of action knowledge in sensory and motor systems

Christine E. Watson<sup>1,2\*</sup>, Eileen R. Cardillo<sup>2</sup>, Bianca Bromberger<sup>2</sup> and Anjan Chatterjee<sup>2</sup>

<sup>1</sup> Moss Rehabilitation Research Institute, Einstein Healthcare Network, Elkins Park, PA, USA

<sup>2</sup> Department of Neurology and Center for Cognitive Neuroscience, University of Pennsylvania, Philadelphia, PA, USA

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Emiliano Ricciardi, University of Pisa, Italy

Anna M. Borghi, University of Bologna, Italy

## \*Correspondence:

Christine E. Watson, Moss Rehabilitation Research Institute, 50 Township Line Rd., Elkins Park, PA 19027, USA  
e-mail: watsonch@einstein.edu

Neuroimaging studies have found that sensorimotor systems are engaged when participants observe actions or comprehend action language. However, most of these studies have asked the binary question of whether action concepts are embodied or not, rather than whether sensory and motor areas of the brain contain graded amounts of information during putative action simulations. To address this question, we used repetition suppression (RS) functional magnetic resonance imaging to determine if functionally-localized motor movement and visual motion regions-of-interest (ROI) and two anatomical ROIs (inferior frontal gyrus, IFG; left posterior middle temporal gyrus, PMTG) were sensitive to changes in the exemplar (e.g., two different people “kicking”) or representational format (e.g., photograph or schematic drawing of someone “kicking”) within pairs of action images. We also investigated whether concrete versus more symbolic depictions of actions (i.e., photographs or schematic drawings) yielded different patterns of activation throughout the brain. We found that during a conceptual task, sensory and motor systems represent actions at different levels of specificity. While the visual motion ROI did not exhibit RS to different exemplars of the same action or to the same action depicted by different formats, the motor movement ROI did. These effects are consistent with “person-specific” action simulations: if the motor system is recruited for action understanding, it does so by activating one’s own motor program for an action. We also observed significant repetition enhancement within the IFG ROI to different exemplars or formats of the same action, a result that may indicate additional cognitive processing on these trials. Finally, we found that the recruitment of posterior brain regions by action concepts depends on the format of the input: left lateral occipital cortex and right supramarginal gyrus responded more strongly to symbolic depictions of actions than concrete ones.

**Keywords:** actions, functional magnetic resonance imaging (fMRI), motor system, semantic memory, visual motion

## INTRODUCTION

A growing body of research suggests that our knowledge about the world is tightly intertwined with the brain’s systems for perception and action (Barsalou, 1999; Gallese and Lakoff, 2005; Decety and Grèzes, 2006; see Barsalou, 2008 for a review). On these “embodied” accounts of semantic memory, sensory and motor states from real-world experiences are re-activated, or simulated, when we understand the meaning of words or other symbols (Barsalou, 1999, 2003; Gallese and Lakoff, 2005). In part because of the discovery of neurons in monkeys that fire both during action execution and observation (e.g., Di Pellegrino et al., 1992), researchers have been particularly interested in understanding the way in which the meanings of human actions and events are represented within the semantic system (Pulvermüller, 1999; Vigliocco et al., 2004; Gallese and Lakoff, 2005; Aziz-Zadeh and Damasio, 2008; Gallese and Sinigaglia, 2011). The extant evidence indicates that when we comprehend language referring to actions or think about the actions depicted in photographs or drawings, we engage, at least in part, sensory and motor systems

in the brain (e.g., Kable et al., 2002; Hauk et al., 2004; Assmus et al., 2007; Raposo et al., 2009; Saygin et al., 2010). For example, reading words referring to actions performed with different body parts (e.g., “pick,” “lick,” “kick”) activates primary motor and premotor cortex in a somatotopic way (Hauk et al., 2004; see also Boulenger et al., 2009). Similarly, when participants view or make semantic decisions about actions in drawings or photographs (Kable et al., 2002; Assmus et al., 2007), or comprehend sentences describing motion events (Pirog Revell et al., 2008; Saygin et al., 2010 see Gennari, 2012 for a review), activation is observed within area MT+, a part of the visual system specialized for processing motion (Huk et al., 2002). Thus, action concepts may be represented within the same areas of the brain involved in actually executing and perceiving dynamic actions (see Watson et al., 2013 for a meta-analysis of this literature). (Throughout the manuscript, we will use “action concepts” as shorthand for “the semantic representations of actions”).

However, most studies on the neural basis of action concepts have asked the binary question of whether action concepts are

embodied or not, rather than whether action concepts contain graded amounts of sensory and motor information during putative action simulations (see Chatterjee, 2010; Willems and Francken, 2012 for similar critiques). One possible scenario is that action concepts typically evoke the same simulation: different exemplars of an action (e.g., different photographs of someone diving) or different representational formats (e.g., photographs, drawings, or words) produce the same response within sensory and motor systems. Alternatively, neural activity in sensory and motor systems may differ each time an action concept is engaged, preserving details specific to the particular exemplar of an action or format of the input.

In the present study, we addressed this question by examining neural responses to action concepts evoked by different exemplars of actions and by distinct visual formats. First, we used a repetition suppression (RS) paradigm (Grill-Spector and Malach, 2001; Maccotta and Buckner, 2004; Grill-Spector et al., 2006) to determine whether functionally-localized motor movement and visual motion (area MT+) regions-of-interest (ROIs) were sensitive to changes in the exemplar (different people performing the same action) or format (perceptually-rich photographs vs. pared-down, schematic drawings) between pairs of action images. If visual motion or motor areas exhibit decreases in activation (RS) to pairs of images depicting different exemplars of the same action or the same action in different formats, relative to pairs of different action images, it would suggest that an action concept (e.g., running) always evokes the same embodied response. On the other hand, an absence of RS for changes in exemplar or format would be consistent with the hypothesis that sensory and motor simulations preserve instance-specific details about actions.

In addition to these functional ROIs, we also looked for RS within left posterior middle temporal gyrus (pMTG) and bilateral inferior frontal gyri (IFG), two areas of the brain consistently implicated in the representation of semantic knowledge of actions (e.g., Kilner et al., 2009; Kalénine et al., 2010). The proximity of pMTG and IFG to visual motion and motor systems, respectively, enabled us to test the claim that areas of the brain adjacent to modality-specific regions may represent more abstract information derived from those modalities (Plaut, 2002; Thompson-Schill, 2003; Chatterjee, 2008, 2010).

Examining RS within these ROIs allowed us to determine the specificity of action knowledge represented in sensory and motor systems. Additionally, we tested whether photographs of actions and schematic drawings of actions elicited different patterns of activation throughout the brain; we refer to these two types of visual depictions of actions as different “representational formats.” In contrast to perceptually-rich photographs, schematic drawings preserve the fundamental analog structure of the things they represent while eliminating specific perceptual details (Peirce, 1955; Deacon, 1997). As a result, schematic drawings represent meaning more symbolically than photographs, but less symbolically than words. Consequently, schematic drawings may also engage more abstract mental representations than those engaged by concrete percepts, and less abstract representations than those engaged by purely-symbolic language (Chatterjee, 2001). Recent evidence from stroke patients (Amorapanth et al.,

2012; Kranjec et al., 2013) implicates the right supramarginal gyrus as harboring such pared-down schematic visual representations.

Additionally, on a graded view of conceptual representation in the brain (Thompson-Schill, 2003; Chatterjee, 2008, 2010), more abstract representations of knowledge are located adjacent to primary sensory and motor cortices. Given that schematic drawings are a more symbolic representational format than photographs, we predict that they will activate brain regions adjacent to those activated by more concrete photographs. Alternatively, areas of the brain involved in representing action concepts may not distinguish between these different representational formats.

## MATERIALS AND METHODS

### PARTICIPANTS

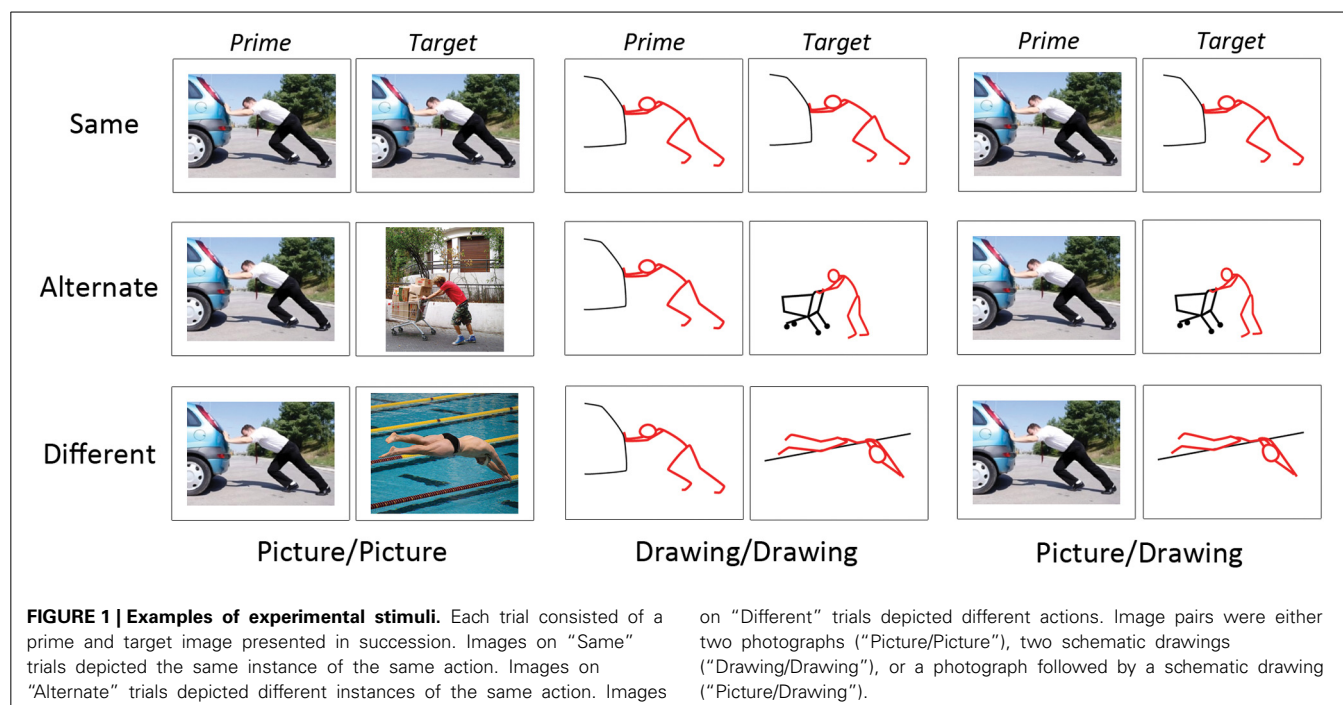
Sixteen participants (7 male;  $M_{\text{age}} = 25.3$  years, range: 20–34 years) participated in the study. All participants were right-handed, native speakers of English with normal or corrected-to-normal vision and no history of neurologic or psychiatric illness. All participants gave informed consent in accordance with the procedures of the University of Pennsylvania Institutional Review Board and were paid \$20/h for their participation. One participant was excluded from the study for having average task accuracy less than 2.5 standard deviations from the group’s mean accuracy.

### STIMULI

Stimuli were 30 photographs (hereafter, “pictures”) and 30 schematic drawings (hereafter, “drawings”) of humans performing common transitive or intransitive actions. We created schematic drawings by tracing with a thick red line the configuration of the actor’s body in each picture. Drawings of transitive actions contained a simple black shape or line representing the recipient object; drawings of intransitive actions contained a black line representing the ground or other relevant background indicator. To ensure that pictures and drawings were equally recognizable, we collected name agreement measurements from 20 pilot participants. The two image formats did not differ on average name agreement [ $M_{\text{pictures}} = 97.9\%$ ,  $SD_{\text{pictures}} = 2.5$ ;  $M_{\text{drawings}} = 97.7\%$ ,  $SD_{\text{drawings}} = 2.9$ ;  $t_{(29)} = 0.43$ ,  $p > 0.8$ ].

Pictures and drawings depicted six unique actions: three transitive actions (“kick”, “pull”, “push”) and three intransitive actions (“stretch”, “dive”, “walk”). Each action was represented in the stimulus set by five pictures and five corresponding drawings showing different exemplars of the action (e.g., five different people diving).

Each experimental trial contained a prime image and a target image. We paired the 30 pictures and 30 drawings in different ways to form the two conditions of interest (Figure 1). First, we manipulated the representational format of the prime and target (“format type”). The prime and target could both be pictures (Picture/Picture), both drawings (Drawing/Drawing), or the prime could be a picture and the target, a drawing (Picture/Drawing). Critically, we did not examine statistically the fourth combination of format types, Drawing/Picture trials; these trials served as filler trials. We adopted this approach to avoid unnecessarily testing conditions with no unique hypotheses. By examining Picture/Drawing trials, we could assess whether RS



occurred between format types. If we used Drawing/Picture trials to address the same question a second time, we would increase the likelihood of a finding a false positive result.

Second, we manipulated the perceptual and/or conceptual similarity between the prime and target (“action similarity”), where “conceptual similarity” refers to the same action (e.g., “kicking”). On “Same” trials, the prime and target depicted the *same exemplar of the same action*; thus, prime and target were similar perceptually and conceptually. On “Alternate” trials, the prime and target depicted *different exemplars of the same action*; thus, the prime and target were similar conceptually but not perceptually. On “Different” trials, the prime and target depicted *different actions* and so were unrelated both perceptually and conceptually. Note that although prime and target were always perceptually similar on Same trials, the degree of this perceptual similarity was greater for Picture/Picture and Drawing/Drawing trials (i.e., the identical picture or drawing as prime and target) relative to Picture/Drawing trials (i.e., the picture and the schematic drawing derived from it as prime and target).

In sum, we manipulated the format type (3) and action similarity (3) of the image pairs. Each cell of our design contained 30 behavioral trials, yielding 270 trials of interest. Given our initial set of 30 pictures and 30 drawings, only 30 prime-target pairings were possible for Same trials of each format type (Picture/Picture, Drawing/Drawing, Picture/Drawing). To create Alternate and Different trials, we selected randomly 30 prime-target pairs from all possible pairings at each level of format type and action similarity. We used these same procedures to select Drawing/Picture filler trials.

## PROCEDURE

During the experiment, participants decided if the prime and target images depicted “the same or different actions” at a

conceptual level. The correct response for Same and Alternate trials was “yes” (e.g., prime and target both depict the same exemplar, or different exemplars, of “diving”). The correct response for Different trials was “no” (e.g., prime and target depict “diving” and “kicking”). Prior to entering the scanner, participants completed 5 min of practice trials to ensure that they understood the task. To prevent participants from exploiting low-level visual cues to make their decisions (e.g., correspondences between the image boundaries of prime and target), prime and target images were presented at different random locations on the screen.

On each trial, participants viewed the prime image for 1000 ms, followed by a 250 ms fixation cross. Then, the target image appeared for 1750 ms, during which the participant made his or her response. In total, each trial lasted 3000 ms. On null trials, participants viewed a fixation cross for 3000 ms. Trials were separated by a 500 ms blank screen. The experiment was presented using E-Prime software (Psychology Software Tools, Pittsburgh, PA) on a computer connected to a projector. Manual responses and reaction times (RTs) were recorded with a button box held by participants with both hands. “Yes” or “no” responses were made by pressing a button with the left or right thumb. Half of the participants indicated “yes” responses with a right button press and “no” responses with a left button press; the other half of participants were assigned the reverse pattern. While in the scanner, participants completed 270 trials of interest, 90 filler trials, and 90 null trials. Trials were presented in five scanning runs of 5.4 min each. Each run began with 9 s of introductory screens. Following these “ready screens,” experimental, filler, and null trials occurred randomly within and across runs for each participant.

After the experimental trials, participants completed two functional localizer scans. During the visual motion (area MT+) localizer, participants passively viewed four 32.5-s blocks each of

moving (flow fields) or stationary white dots on a black background (Bavelier et al., 2001; Saygin et al., 2010). During the motor movement localizer, participants were instructed via computer screen to move the right hand, left hand, right foot, and left foot continuously for 20 s, or to rest for 20 s (Hauk et al., 2004; Boulenger et al., 2009; Raposo et al., 2009). Each type of block was presented 4 times.

### DATA ACQUISITION

We collected structural and functional data on a 3.0 Tesla Siemens Trio scanner using an eight-channel head coil. We acquired high-resolution T1-weighted structural images using a MP-RAGE pulse sequence and near-isotropic voxels ( $0.98 \times 0.98 \times 1$  mm). T2\*-weighted echo-planar images were collected during the five experimental scanning runs (104 volumes each), the MT+ localizer (91 volumes), and the motor localizer (102 volumes) (repetition time = 3 s; echo time = 30 ms; flip angle =  $90^\circ$ ; field of view = 220 mm; slice thickness = 3 mm; matrix size =  $64 \times 64$ ; voxel size =  $3.4 \times 3.4 \times 3$  mm). Each functional volume consisted of 50 axial slices that covered the whole cerebral cortex.

### fMRI DATA PREPROCESSING

Imaging data was preprocessed and analyzed using the FMRIB Software Library (FSL version 4.1; <http://www.fmrib.ox.ac.uk/fsl>). The first three volumes of each functional run were discarded to allow for steady state magnetization. Functional data were slice timing corrected using sinc interpolation, motion corrected, and high-pass filtered (0.01 Hz). For each participant, functional data from each run were registered to a participant's high-resolution structural image using FMRIB's Linear Registration Tool with  $7^\circ$  of freedom. One set of functional data for use in region-of-interest analyses was kept in each participant's native space and smoothed with a Gaussian kernel of 4 mm (full-width at half-maximum). A second copy of functional data for use in group-level analyses was registered to Montreal Neurological Institute standard space (MNI-152) using linear registration with  $12^\circ$  of freedom and smoothed with a Gaussian kernel of 8 mm.

### FIRST-LEVEL ANALYSES

We first modeled each functional scanning run separately for each participant with FMRIB's FEAT (fMRI Expert Analysis Tool). We used an event-related model in which the events of interest began with the onset of the prime image and ended with the offset of the target image. Events were modeled as single impulses convolved with FSL's double-gamma hemodynamic response function (HRF), along with the event's temporal derivative. Regressors were created for each format type/action similarity combination [e.g., Picture/Picture(Same), Picture/Picture(Alternate), etc.], and for filler trials and null trials. Contrasts of interest were computed at the first level using linear combinations of these regressors.

### HIGHER-LEVEL ANALYSES

For each participant, contrasts between conditions modeled within a run were combined at the second-level using a fixed effects model within FMRIB's Local Analysis of Mixed Effects (FLAME). Finally, contrasts intended for third-level, group analyses were combined across participants using a mixed effects

model (FLAME1+2). Resulting group-level maps of  $z$ -statistics were thresholded at  $z > 2.3$  with a corrected cluster significance threshold of  $p < 0.05$  (Worsley et al., 1992). In order to compare the location of the visual motion ROI with our group-level results, we also computed the location of the visual motion ROI at the group level. To more precisely determine the anatomical location of this region, we thresholded this analysis using voxel-based, rather than cluster-based, thresholding (GRF-theory-based maximum height thresholding with  $p < 0.05$ , corrected) (Worsley et al., 1992).

### REGION-OF-INTEREST ANALYSES

For region-of-interest (ROI) analyses, we used FMRIB's Featquery tool to compute, for each participant, the mean contrast of parameter estimates in each ROI for each condition [i.e., Picture/Picture (Same), Picture/Picture (Alternate), etc.] minus null (fixation) trials. With this data, within-subject RS effects were evaluated using SPSS software. We looked for RS within each ROI by looking for effects of action similarity (Same, Alternate, Different) and format type (Picture/Picture, Drawing/Drawing, Picture/Drawing) using a two-way repeated measures ANOVA. When we observed an interaction between action similarity and format type,  $p$ -values from tests of simple effects were corrected for multiple comparisons using the Holm-Sidak method.

Our two ROIs of primary interest were defined functionally for each participant. Visual motion ROIs were defined by contrasting blocks in which participants perceived moving vs. stationary dots (see above). The resulting map of  $z$ -values for this contrast was thresholded first at a False Discovery Rate (FDR) (Nichols and Holmes, 2002) of  $q = 0.000001$ . (Here, we used the FDR method given that it controls the family-wise error rate without being overly conservative for low smoothness data with few degrees of freedom, Nichols and Hayasaka, 2003.) We then selected the largest cluster in each hemisphere that survived this threshold and fell within lateral occipital cortex. This anatomical constraint was applied rarely and excluded clusters that emerged in the occipital poles. Using this procedure, visual motion ROIs were localized for 10 participants. For 2 participants, no voxels survived at this threshold, so we used a more lenient threshold of  $q = 0.05$ . We note that using a more lenient threshold to identify ROIs in some participants does not bias us to find differences between the experimental conditions. On the contrary, by using voxels that respond less strongly to visual motion, we may have increased noise in our analyses, making it more difficult to detect effects. For 3 participants, no visual-motion-preferring voxels were detected even at a relaxed threshold. The average visual motion ROI had a volume of  $7995 \text{ mm}^3$  ( $SD = 5420$ ).

Motor movement ROIs were defined in each participant by contrasting the movement of each effector (left hand, right hand, left foot, right foot) with rest (see above). Resulting  $z$ -maps for each of these contrasts were thresholded with the same general procedure described for the visual motion ROI. For each effector, we selected the largest cluster that survived the threshold. Clusters for each of the four effectors were then combined to form a participant's entire motor movement ROI. In 10 participants, a motor ROI was identified at  $q = 0.0000001$ ; for 2 other participants, the threshold was relaxed to  $q = 0.05$ . We were unable to identify



a motor movement ROI in 3 participants. The average motor movement ROI had a volume of 22813 mm<sup>3</sup> ( $SD = 11111$ ).

**Figure 2** depicts the overlap of participants' visual motion and motor movement ROIs transformed into MNI-152 standard space. The location of visual motion ROIs within lateral temporo-occipital cortex agrees with previous localizations of area MT+ (e.g., Dumoulin et al., 2000). Motor movement ROIs primarily covered lateral and medial pre- and post-central gyri.

To ensure that RS within the motor movement ROI could not be attributed to lower-level processes, we made a further adjustment to analyses performed within each participant's motor movement ROI. In the experimental task, trials on which a participant responds "yes" (i.e., Same and Alternate trials) occurred more frequently than "no" trials (i.e., Different trials). Since participants used one hand more often throughout the experiment, it is possible that we could observe a decrease in neural activity for Same/Alternate trials relative to Different trials within motor regions due to manual response priming (i.e., repeated use of one hand for responding). Therefore, we calculated the effects of Same/Alternate trials (relative to null trials) and Different trials (relative to null trials) *only* within the hemisphere ipsilateral to the manual response for each condition. In other words, for analyses within the motor movement ROI, we only considered activation within the hemisphere not responsible for a participant's button press. For participants who responded "yes" with the right hand (to Same/Alternate trials), mean contrast of parameter estimates for Same and Alternate trials relative to null were computed only within the right hemisphere motor movement ROI; mean contrast of parameter estimates for Different trials ("no" responses made with the left hand) were computed only within the left hemisphere motor movement ROI. In using this procedure, we ensured that RS effects observed within motor regions could be attributable only to the experimental manipulations rather than priming of manual responses.

In addition to these two functionally-defined ROIs, we created two anatomical ROIs: bilateral IFG and left pMTG. Each area was taken from the Harvard-Oxford Cortical Atlas that is registered to MNI-152 standard space and included in the FSL distribution. ROIs in standard space were transformed into each participant's native space using linear registration (FLIRT). For each ROI, we excluded any voxels that were also included in a participant's

functionally-defined visual motion and motor movement ROIs to ensure that observations within the ROIs were independent of each other. Similarly, participants for whom visual motion ( $n = 3$ ) and motor movement ( $n = 3$ ) ROIs could not be located were excluded from IFG and pMTG ROI analyses given that we could not rule out overlap between functionally-responsive and anatomically-localized areas in these participants. Finally, given the contribution of IFG to action execution (e.g., Caspers et al., 2010; Press et al., 2012), we analyzed activation with the IFG ROI in the same manner as the motor movement ROI (see above).

## RESULTS

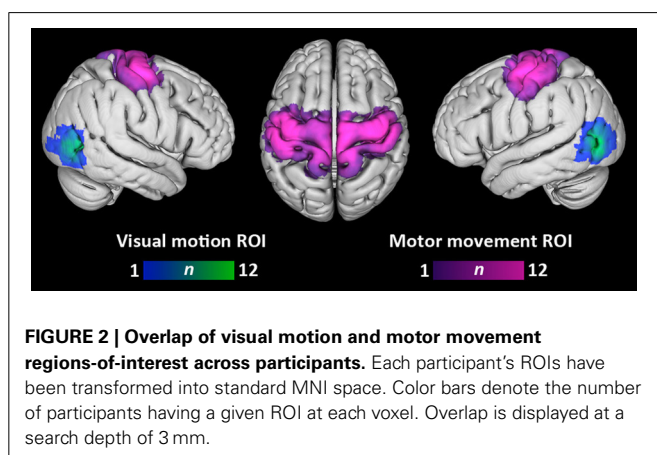
### BEHAVIORAL ANALYSES

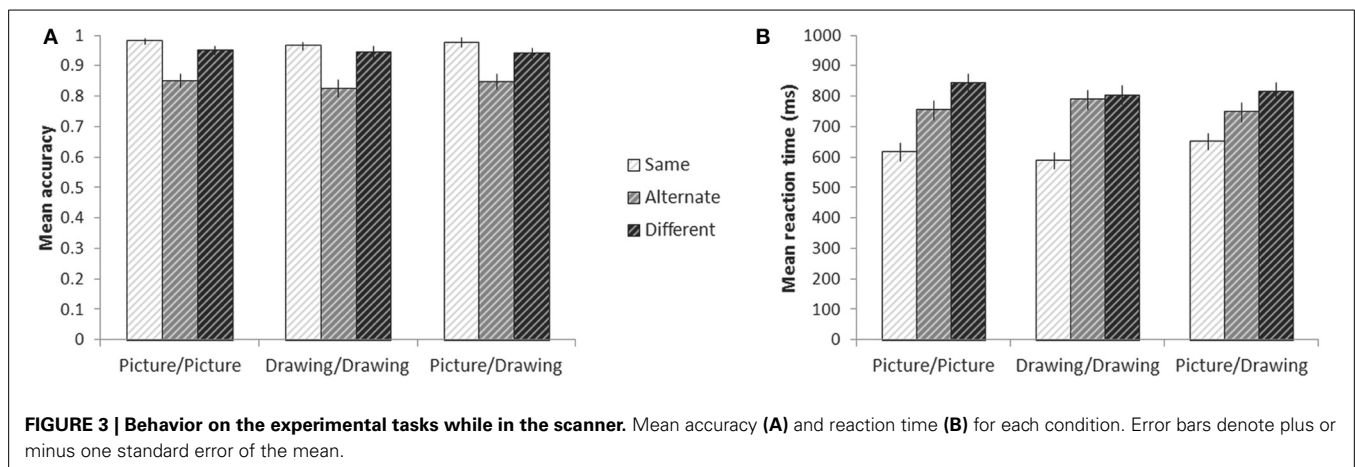
We used a two-way repeated measures ANOVA to look for effects of action similarity (Same, Alternate, and Different) and format type (Picture/Picture, Drawing/Drawing, and Picture/Drawing) on accuracy. We found a significant effect of action similarity [ $F(2, 28) = 28.7, p < 0.001$ ] and a marginal effect of format type [ $F(2, 28) = 2.7, p = 0.08$ ] (**Figure 3A**). The interaction between action similarity and format type was not significant. Pairwise comparisons revealed that participants were significantly less accurate on Alternate trials relative to Different ( $p = 0.02$ ) and Same ( $p = 0.02$ ) trials, and significantly less accurate on Different trials than Same trials ( $p = 0.01$ ). Pairwise comparisons between format types showed that participants were significantly less accurate on Drawing/Drawing trials than Picture/Picture trials ( $p = 0.03$ ); however, the mean difference in accuracy between these conditions was very small (1.6%). No other pairwise differences between format types reached significant.

Reaction time analyses were conducted only for correct trials. There was a significant effect of action similarity on participants' RTs [ $F(2, 28) = 67.2, p < 0.001$ ] and a significant interaction between action similarity and format type [ $F(4, 56) = 30.0, p < 0.001$ ] (**Figure 3B**). The effect of format type was not significant. To explore the interaction, we calculated simple effects between levels of action similarity for each format type. For every format type, participants responded to Same trials significantly faster than either Alternate trials (all  $p < 0.001$ ) or Different trials (all  $p < 0.001$ ). For Picture/Picture trials, participants also responded more quickly to Alternate trials than Different trials ( $p = 0.005$ ). For Drawing/Drawing and Picture/Drawing trials, however, there was no significant difference between RTs to Alternate and Different trials. When jointly considering participants' RTs and accuracy, we note that participants' lower accuracy on Alternate trials may not reflect errors, *per se*, but individual differences in whether a participant believed the two images indeed depicted the same action. On the other hand, reaction time analyses were only carried out on trials in which participants accepted identical and alternate exemplars and rejected images of different actions as depicting the same action; RTs thus reflect the time to accumulate sufficient information to make each type of decision (e.g., Ratcliff, 1978).

### ROI ANALYSES

Visual motion and motor movement ROIs were functionally-localized for each participant. For each participant, we calculated the mean contrast of parameter estimates between each





condition and null (fixation) trials within these regions. Then, we looked for effects of the action similarity (Same, Alternate, Different) and format type (Picture/Picture, Drawing/Drawing, and Picture/Drawing) of the prime and target images using a two-way repeated measures ANOVA. Within the visual motion ROI, there were significant effects of action similarity [ $F_{(2, 22)} = 8.3$ ,  $p = 0.002$ ] and format type [ $F_{(2, 22)} = 7.0$ ,  $p = 0.005$ ], and a marginally significant interaction between the two [ $F_{(4, 44)} = 2.2$ ,  $p = 0.08$ ] (Figure 4A). Simple effects between levels of action similarity for each format type showed significant suppression for Same trials relative to Different ( $p = 0.03$ ) and relative to Alternate ( $p = 0.003$ ) trials only for the Picture/Picture condition. No other pairwise comparisons were significant. Thus, the visual motion ROI exhibited RS only when the prime and target images were identical, perceptually-rich photographs of actions.

We evaluated RS effects within the motor movement ROI only within the hemisphere ipsilateral to each condition's expected manual response (see Materials and Methods). We observed a significant effect of action similarity [ $F_{(2, 22)} = 8.4$ ,  $p = 0.002$ ] but no effect of format type or interaction between the two (Figure 4B). Planned comparisons between each level of action similarity showed significant suppression for Same trials relative to Different ( $p = 0.006$ ) and Alternate trials ( $p = 0.01$ ). Suppression for Alternate trials relative to Different trials was not significant but showed a trend in that direction ( $p = 0.09$ ). However, the main effect of action similarity was significantly fit by a linear contrast between Same, Alternate, and Different levels [ $F_{(1, 22)} = 11.5$ ,  $p = 0.006$ ], suggesting that RS occurred in the motor movement ROI when the prime and target images referred to the same basic action, even if different exemplars or representational formats.

Next, we looked for effects of action similarity and format type within areas of the brain near to functionally-localized visual motion and motor movement ROIs. Within left pMTG, we observed significant effects of format type [ $F_{(2, 22)} = 9.5$ ,  $p = 0.001$ ] and action similarity [ $F_{(2, 22)} = 3.8$ ,  $p = 0.04$ ], but no significant interaction between the two (Figure 4C). Planned comparisons between each level of action similarity revealed significant suppression for Same trials relative to Alternate trials ( $p = 0.03$ ) and marginally significant suppression for Same

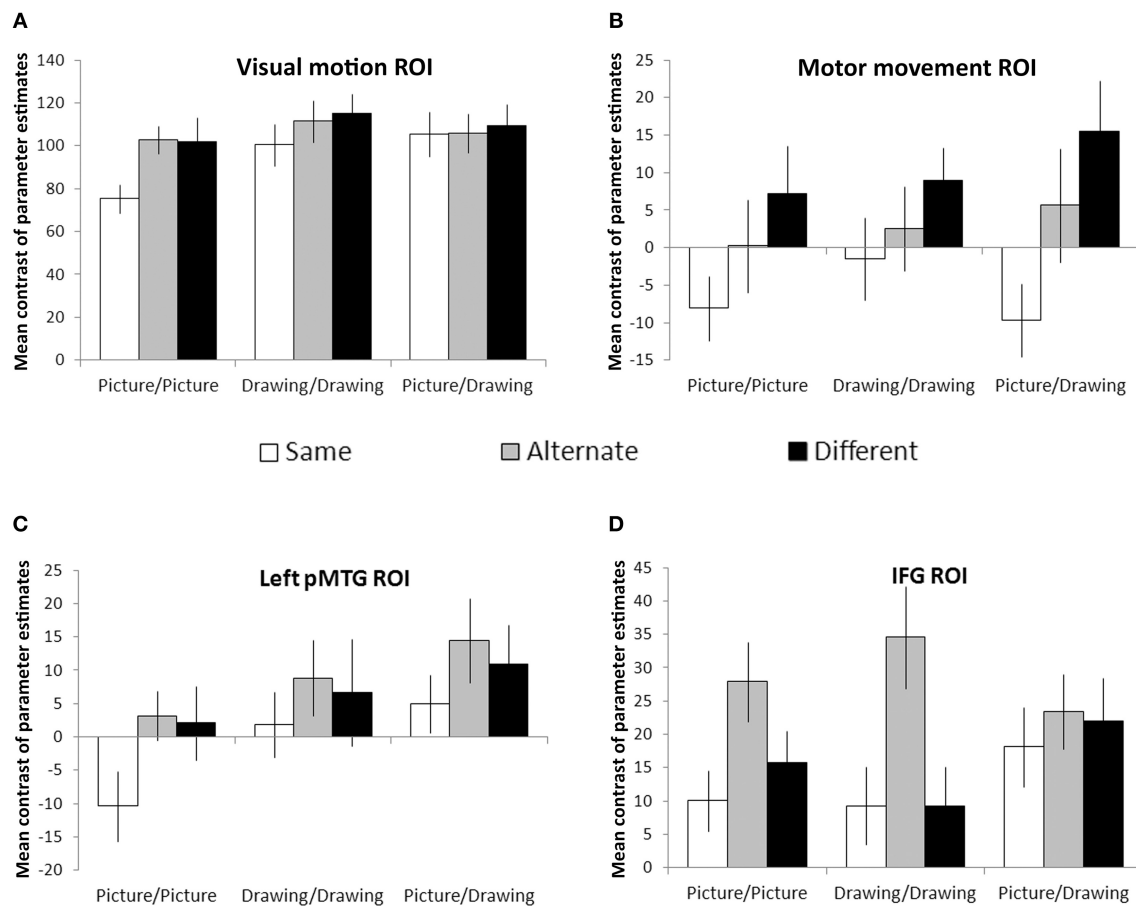
trials relative to Different trials ( $p = 0.08$ ). There was no difference between Alternate and Different trials. Planned comparisons between each format type indicated significantly less activation within left pMTG for Picture/Picture trials relative to Drawing/Drawing ( $p = 0.01$ ) or Picture/Drawing trials ( $p = 0.001$ ), and Drawing/Drawing and Picture/Drawing trials were not significantly different from one another. Thus, left pMTG exhibited suppression when the prime and target were identical but not when they were merely different exemplars of the same action. And, this area of the brain was more strongly activated overall when the prime or target image was a schematic drawing of an action.

Finally, we examined RS effects within the IFG. As with the motor movement ROI, we analyzed activation within the hemisphere ipsilateral to each condition's expected manual response (see Materials and Methods). Within IFG, we found a significant effect of action similarity [ $F_{(2, 22)} = 8.1$ ,  $p = 0.002$ ]. There was no effect of format type or interaction (Figure 4D). Planned comparisons between levels of action similarity revealed no difference between activation on Same and Different trials ( $p = 0.53$ ). Surprisingly, we also observed significant enhancement (i.e., an increase) for Alternate trials relative to both Different ( $p = 0.02$ ) and Same ( $p < 0.001$ ) trials. This result indicates that IFG exhibited not suppression, but *increased* activity when the images depicted different exemplars of the same action.

Although these analyses examined the *patterns* of RS effects between conditions, we note that the overall magnitude of values within each ROI reflects the degree to which an ROI was more active during the task than fixation. For example, large mean contrasts of parameter estimates within the visual motion ROI likely reflect the richer visual input present on experimental trials relative to fixation crosses.

## WHOLE-BRAIN ANALYSES

To determine if concrete and more symbolic representations of actions activate distinct areas throughout the brain, we also used a whole-brain, group-level analysis to compare activation for perceptually-rich photographs of actions (Picture/Picture trials) with activation for schematic drawings of actions (Drawing/Drawing trials). Because Same and Alternate trials were



**FIGURE 4 | Region-of-interest analyses.** Visual motion (A) and motor movement (B) areas were functionally-localized in each participant. Left pMTG (C) and bilateral IFG (D) were defined anatomically using the

Harvard-Oxford cortical atlas. Bars reflect the mean contrast of parameter estimates between each condition and null (fixation) trials. Error bars denote plus or minus one standard error of the mean.

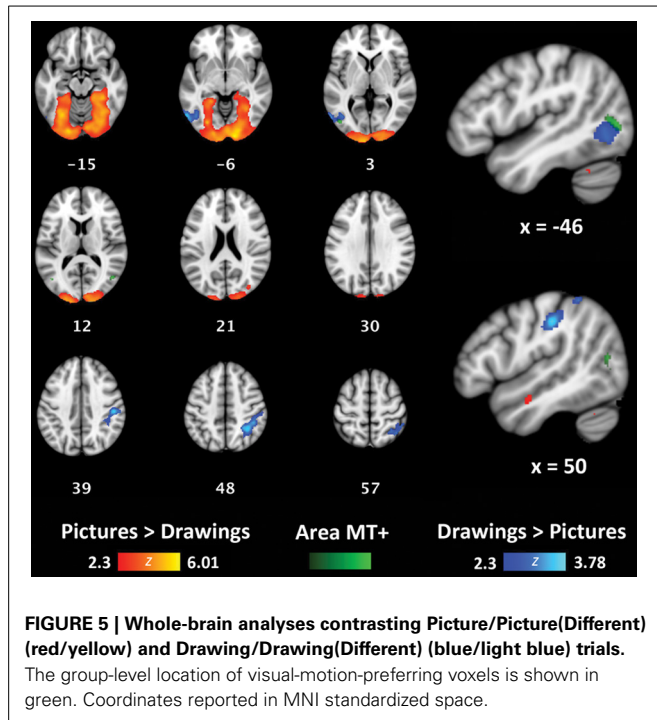
hypothesized to exhibit RS effects, we only compared Different trials for each of these two formats. Relative to Drawings, Pictures activated a large, bilateral cluster that began in the occipital poles and extended into the fusiform gyri in both hemispheres (volume = 32710 mm<sup>3</sup>; maximum  $z$ -value = 6.01; MNI coordinates of maximum:  $x = 16$ ,  $y = -96$ ,  $z = -8$ ) (Figure 5, red/yellow). Relative to Pictures, Drawings activated a cluster in the right supramarginal gyrus and superior parietal lobule (volume = 3096 mm<sup>3</sup>; maximum  $z$ -value = 3.78; MNI coordinates of maximum:  $x = 32$ ,  $y = -52$ ,  $z = 52$ ) (Figure 5, light blue/dark blue). Drawings also activated a smaller cluster within left lateral occipital cortex (volume = 1782 mm<sup>3</sup>; maximum  $z$ -value = 3.51; coordinates of maximum:  $x = -58$ ,  $y = -66$ ,  $z = -6$ ). The majority of voxels in this cluster were located anterior to the typical location of area MT+, as reported in other studies (Dumoulin et al., 2000) and within our own participant group (Figure 5, group-level visual-motion-preferring voxels shown in light green/dark green).

## DISCUSSION

In the present study, we used RS fMRI to determine the specificity of information carried by sensory and motor systems during

conceptual processing of actions. Of primary interest was whether brain regions involved in performing movements and perceiving visual motion, two areas of the brain often engaged by action concepts (Hauk et al., 2004; e.g., Kable et al., 2002), were sensitive to changes in the exemplar or representational format of pairs of action images.

Our results reveal strikingly different response patterns between these two brain areas: while the visual motion ROI exhibited RS only for identical photographs of actions, suppression occurred in the motor movement ROI for repetitions of the same *and* alternate exemplars of an action, irrespective of the format. This result suggests that neural activity within these sensorimotor regions during semantic tasks represents information about actions at different levels of specificity. On the one hand, during comprehension of static depictions of actions, voxels that respond strongly to visual motion appear to encode information highly specific to a particular exemplar of an action or particular representational format: only when the prime and target images were identical *and* conveyed many perceptual details about the actor or action context did we observe RS within the visual motion ROI. Because this region was strongly active for all conditions, it cannot be the case that some conditions merely failed to activate visual



motion areas at all. Instead, neural responses to action concepts within this area preserve detailed information about the specific instance of an action; different actors and/or representational formats activate different neural representations. Furthermore, we did not observe RS when prime and target images were identical schematic drawings. Thus, the absence of perceptually-rich details in schematic drawings may result in a more variable response within areas specialized for visual motion, even across repeated instances of the same schematic drawing.

Although we focused on the activation of visual motion areas by conceptual processing of *static* action images, our results accord with other studies on the response of area MT+ to different types of visual motion. In particular, this area is sensitive to changes in the speed, direction, and velocity of low-level visual motion (Wall et al., 2008; Lingnau et al., 2009; Cardin et al., 2012; Weigelt et al., 2012). Thus, to the extent that different exemplars of an action or different representational formats convey actions performed at different speeds, in different directions, etc., the response within visual motion regions may differ.

Yet, our results are at odds with two prior studies investigating RS between pairs of *dynamic* action stimuli (i.e., videos) using a semantic task (Kable and Chatterjee, 2006; Wiggett and Downing, 2010; but see Grossman et al., 2010). In both of these studies, area MT+ was insensitive to changes in the actor and thus responded similarly as long as the same action was repeated (e.g., “kicking”). Given that both of these studies used stimuli that contained actual visual motion, an alternative explanation of the present results is that area MT+ exhibits a narrower range of responses to static images than dynamic action stimuli. Although static images engage this area, they may do less strongly and with less variability than dynamic depictions of actions. If so, then the absence of RS to alternate exemplars within the visual motion ROI in

the current study may reflect insufficient physiological power to detect differences between all conditions in this area.

In contrast to the highly-specific effects we observed within the visual motion ROI, the motor movement ROI exhibited RS between pairs of images that depicted identical actions *and* pairs that depicted alternate exemplars of the same action. This response occurred both when the prime and target were the same format (Picture/Picture, Drawing/Drawing) or different formats (Picture/Drawing). This result suggests that a similar representation is evoked within the motor system irrespective of the way in which an action concept is accessed; the same motor simulation is produced in response to different exemplars of the same action or to actions presented in different formats.

One way in which this result could arise is if motor simulations are grounded in person-specific motor programs for actions. In other words, no matter who I perceive doing an action (e.g., Jack kicking, Jane kicking) or the format of the input (e.g., a photograph or schematic drawing of “kicking”), my motor simulation will reflect the way in which *I* am inclined to kick. Indeed, there is prior evidence that the involvement of motor regions in representing action concepts depends on an individual’s particular physical experiences (Calvo-Merino et al., 2005, 2006; Beilock et al., 2008). For example, Calvo-Merino et al. (2005) found that the degree to which expert ballet and capoeira dancers recruited motor regions during action observation differed when watching their own style of dance versus the other; the authors conclude that “...action observation evokes individual, acquired motor representations...” (p. 1247). Similarly, participants’ ability to recall actions depends on their motor expertise with those actions (Pezzulo et al., 2010). The present results extend these findings by suggesting that an action evokes the same person-specific motor simulation irrespective of the way in which an action concept is accessed.

However, we note that the degree to which the motor system participates in representing action concepts *at all* is also modulated by physical experience (described above) and task demands (Van Dam et al., 2012). Our recent meta-analysis of neuroimaging studies using action words and action images did not find consistent involvement of premotor or primary motor cortex in conceptual processing of these stimuli (Watson et al., 2013). In the current study, we used a small set of very familiar actions, and we functionally-localized areas involved in performing movements within each participant. Therefore, we may have been more likely than other studies to generate and detect effects within the motor system during conceptual processing of actions.

Even though participants made manual responses on each trial, our study design makes it unlikely that the RS we observed within the motor movement ROI reflects manual response priming. First, for each participant, we only analyzed activation within the hemisphere that was ipsilateral to each condition’s expected response. Thus, results from the motor movement ROI reflect activation within the hemisphere not responsible for the button press. Second, the RS effects were not entirely determined by activation within hand-preferring parts of the motor system: we also functionally-localized areas active when performing foot movements. Finally, we observed significantly different levels of activation within the motor movement ROI for Same and



Alternate trials. If manual response priming was driving suppression effects, then we would expect *no* difference between conditions responded to with the same hand.

We used functionally-defined visual motion and motor movement ROIs rather than ROIs defined anatomically or from group-level results. However, since the tasks used to define these ROIs did not require measurable behavioral responses, we cannot be certain that a given participant was paying attention or performing the localizer task; indeed, differences in task engagement may explain why visual motion and motor movement ROIs could not be identified, or required a more lenient threshold to be identified, in some participants. Yet, given the potentially variable functional brain organization of each participant, using ROIs defined in this way allowed us to more precisely test functionally-motivated hypotheses (see Saxe et al., 2006 for a similar argument), i.e., that voxels that participate in more basic cognitive tasks (processing visual motion, executing body movements) would encode information at different levels of specificity during a conceptual task.

We also examined RS effects in anatomically-defined ROIs. Within two brain areas neighboring visual motion and motor movement ROIs, we observed RS when the prime and target image depicted the same instance of the same action, but not different instances of the same action. Instead, within left pMTG, we observed no differentiation between Alternate and Different trials, and within IFG, we observed *enhancement* for Alternate relative to Different and Same trials. In some respects, these results are surprising: some researchers have suggested a “graded” view of embodiment in which more abstract representations of action meaning are represented in brain areas adjacent to modality-specific cortices (Thompson-Schill, 2003; Kable et al., 2005; Chatterjee, 2008, 2010). Therefore, we expected to observe RS for different exemplars of the same action within left pMTG and IFG. However, our pattern of results may be consistent with findings of “repetition enhancement” rather than “repetition suppression” (Raposo et al., 2006; Kuperberg et al., 2008; see Segaert et al., 2013 for a review). One hypothesis is that while suppression occurs when the same cognitive process is performed on a prime and target, enhancement occurs when the target requires additional processes, like explicit memory retrieval (Henson, 2003).

In the current study, we found significant enhancement for Alternate trials within IFG and non-significant but numerically higher activation for Alternate trials relative to Different trials for each format type within left pMTG. Alternate trials were also the most difficult for participants. Therefore, it is possible that verifying alternate exemplars of the same action (vs. the easier tasks of verifying an identical match or a complete mismatch) required additional cognitive processing—and neural activity—within IFG and left pMTG. IFG, in particular, has been shown to play a role in selecting among competing representations in memory (Thompson-Schill et al., 1997; Moss et al., 2005). When determining whether two images were different exemplars of the same action, participants may have had to exert more cognitive effort to find the link between two conceptually similar, but perceptually dissimilar, instances of an action. Lack of RS and numerical enhancement within pMTG may similarly reflect participants’ greater need

to retrieve explicit information about actions in the Alternate condition.

Finally, we investigated at the whole-brain level the degree to which the brain distinguishes between perceptually-rich photographs of actions and more symbolic schematic drawings of actions. Given that they contain more visual details than drawings, pictures unsurprisingly yielded greater activation throughout early visual cortex. The reverse comparison, however, yielded greater activation for schematic drawings in two areas of the brain. First, drawings more strongly engaged the right supramarginal gyrus and parts of the superior parietal lobe, a result in agreement with a recent voxel-based lesion-symptom mapping (VLSM) study from our lab. In this study, stroke patients with damage to the left or right hemisphere matched categorical spatial relations among objects (e.g., “above,” “below”) across different representational formats (i.e., pictures, schematic drawings, and words) (Amorapanth et al., 2012). Patients with damage to right supramarginal gyrus were particularly impaired matching spatial relation words to their corresponding schematic drawings relative to their corresponding pictures. A recent case study also supports the view that schematic drawings are processed differently than perceptually-rich photographs: a patient with simultagnosia, a condition in which patients are characteristically unable to perceive more than a single object at a time (Luria, 1959), was better able to comprehend spatial relations between objects (e.g., “above,” “below”) when they were depicted as schematic drawings rather than as photographs (Kranjec et al., 2013). Given the present results as well as neuroimaging evidence for the activation of right supramarginal gyrus during the naming of spatial relations between objects (e.g., Damasio et al., 2001), this part of the brain may be responsible for recognizing the schematic structure of these pared-down percepts.

We also found greater activation for schematic drawings of actions relative to photographs in left lateral occipital cortex; most voxels in this cluster were located anterior to visual motion-prefering areas, in lateral occipital cortex and the most posterior aspect of pMTG. This result is consistent with a graded view of conceptual representation (Chatterjee, 2008, 2010; Watson and Chatterjee, 2011). Action knowledge derived from visual motion area MT+ is represented along a temporal posterior-to-anterior axis in which increasingly abstract information is represented more anteriorly. Accordingly, a brain area anterior to area MT+ responded more strongly to pared-down, more symbolic schematic drawings than to perceptually-rich photographs of actions. We also observed greater overall activation of the left pMTG ROI for trials that included a schematic drawing (Picture/Drawing or Drawing/Drawing trials). Together, these results suggest that more abstract or symbolic depictions of actions recruit areas adjacent to modality-specific cortices. Consistent with this claim, we found using a meta-analysis approach that words referring to actions consistently activated an area within left middle temporal gyrus anterior to the area associated with visual depictions of actions (Watson et al., 2013). The implication of these findings for embodied accounts of semantic knowledge is that the recruitment of modality-specific—or other—regions depends on whether concepts are accessed by more or less symbolic means. More symbolic depictions may

additionally, or instead, recruit information that is abstracted from direct experience and represented adjacent to modality-specific areas.

Finally, we acknowledge that participants' did not *need* to access conceptual knowledge of actions on all trials. When the prime and target images were identical (Same trials), participants' decisions could be based solely on visual similarity. We note that the RS effects seen in the visual motion ROI suggest that some inference about the images is being made even when they are perceptually identical insofar as neural activity in an area sensitive to visual motion is influenced by static images. A visual similarity strategy would not work on the Alternate and Different trials: though prime and target stimuli were visually dissimilar for both, these trial types required different behavioral responses. Therefore, participants' needed to access the *meaning* of the actions depicted in these images in order to make a response. Furthermore, the pattern of results suggests that participants drew upon action concepts even on Same trials: it is not obvious why the repetition of visually similar images should yield decreased activation in the motor movement ROI. Instead, we suggest that the conceptual similarity of these images—and images in the Alternate condition—produces RS within the motor movement ROI.

Understanding the specificity of brain regions to different exemplars of actions and representational formats makes embodied accounts of the semantic system more precise. Here, we found that sensory and motor systems carried different amounts of information during conceptual processing of actions: while visual motion areas preserved exemplar- and format-specific details, regions involved in performing movements responded similarly as long as images referred to the same basic action (e.g., “kicking”). Thus, when the motor system participates in understanding an action, it may do so by activating one's own motor program for that particular action. Additionally, two brain regions (left lateral occipital cortex and right supramarginal gyrus) responded more strongly to more symbolic representations of actions (i.e., schematic drawings) than to concrete ones (i.e., photographs). For embodied accounts, these data indicate that even outside of area MT+, the recruitment of posterior brain regions by action concepts depends on the format of the input. Within lateral occipitotemporal cortex, in particular, more abstract representations of actions may be represented adjacent to modality-specific cortical areas.

## ACKNOWLEDGMENTS

We would like to thank Geoffrey Aguirre for his help with experimental design and data analysis and Matthew Lehet for his help running participants in pilot normative studies. This work was supported by the National Institutes of Health (grant numbers RO1 DC008779, RO1 DC012511 to Anjan Chatterjee, and T32-NS054575-04 to Christine E. Watson as a trainee).

## REFERENCES

Amorapanth, P., Kranjec, A., Bromberger, B., Lehet, M., Widick, P., Woods, A. J., et al. (2012). Language, perception, and the schematic representation of spatial relations. *Brain Lang.* 120, 226–236. doi: 10.1016/j.bandl.2011.09.007

Assmus, A., Giessing, C., Weiss, P. H., and Fink, G. R. (2007). Functional interactions during the retrieval of conceptual action knowledge: an fMRI study. *J. Cogn. Neurosci.* 19, 1004–1012. doi: 10.1162/jocn.2007.19.6.1004

Aziz-Zadeh, L., and Damasio, A. (2008). Embodied semantics for actions: findings from functional brain imaging. *J. Physiol.* 102, 35–39. doi: 10.1016/j.jphysparis.2008.03.012

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 1177–1187. doi: 10.1098/rstb.2003.1319

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Bavelier, D., Brozinsky, C., Tomann, A., Mitchell, T., Neville, H., and Liu, G. (2001). Impact of early deafness and early exposure to sign language on the cerebral organization for motion processing. *J. Neurosci.* 21, 8931–8942.

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., and Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13269–13273. doi: 10.1073/pnas.0803424105

Boulenger, V., Hauk, O., and Pulvermüller, F. (2009). Grasping Ideas with the motor system: semantic somatotopy in idiom comprehension. *Cereb. Cortex* 19, 1905–1914. doi: 10.1093/cercor/bhn217

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., and Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.* 16, 1905–1910. doi: 10.1016/j.cub.2006.07.065

Cardin, V., Hemsworth, L., and Smith, A. T. (2012). Adaptation to heading direction dissociates the roles of human MST and V6 in the processing of optic flow. *J. Neurophysiol.* 108, 794–801. doi: 10.1152/jn.00002.2012

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112

Chatterjee, A. (2001). Language and space: some interactions. *Trends Cogn. Sci.* 5, 55–61. doi: 10.1016/S1364-6613(00)01598-9

Chatterjee, A. (2008). The neural organization of spatial thought and language. *Semin. Speech Lang.* 29, 226–238. doi: 10.1055/s-0028-1082886

Chatterjee, A. (2010). Disembodying cognition. *Lang. Cogn.* 2, 79–116. doi: 10.1515/LANGCOG.2010.004

Damasio, H., Grabowski, T. J., Tranel, D., Ponto, L. L. B., Hichwa, R. D., and Damasio, A. R. (2001). Neural correlates of naming actions and of naming spatial relations. *Neuroimage* 13, 1053–1064. doi: 10.1006/nimg.2001.0775

Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Human Brain*. New York, NY: W. W. Norton & Company.

Decety, J., and Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Res.* 1079, 4–14. doi: 10.1016/j.brainres.2005.12.115

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180. doi: 10.1007/BF00230027

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Pike, G. B., et al. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb. Cortex* 10, 454–463. doi: 10.1093/cercor/10.5.454

Gallese, V., and Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310

Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cogn. Sci.* 15, 512–519. doi: 10.1016/j.tics.2011.09.003

Gennari, S. P. (2012). Representing motion in language comprehension: lessons from neuroimaging. *Lang. Linguist. Compass* 6, 67–84. doi: 10.1002/lnc.3.317

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006

Grill-Spector, K., and Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst.)* 107, 293–321. doi: 10.1016/S0001-6918(01)00019-1

Grossman, E. D., Jardine, N. L., and Pyles, J. A. (2010). fMR-adaptation reveals invariant coding of biological motion on the human STS. *Front. Hum. Neurosci.* 4:15. doi: 10.3389/neuro.09.015.2010

- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9
- Henson, R. N. A. (2003). Neuroimaging studies of priming. *Prog. Neurobiol.* 70, 53–81. doi: 10.1016/S0301-0082(03)00086-8
- Huk, A. C., Dougherty, R. F., and Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205.
- Kable, J. W., and Chatterjee, A. (2006). Specificity of action representations in the lateral occipitotemporal cortex. *J. Cogn. Neurosci.* 18, 1498–1517. doi: 10.1162/jocn.2006.18.9.1498
- Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., and Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *J. Cogn. Neurosci.* 17, 1855–1870. doi: 10.1162/089892905775008625
- Kable, J. W., Lease-Spellmeyer, J., and Chatterjee, A. (2002). Neural substrates of action event knowledge. *J. Cogn. Neurosci.* 14, 795–805. doi: 10.1162/08989290260138681
- Kalénine, S., Buxbaum, L. J., and Coslett, H. B. (2010). Critical brain regions for action recognition: lesion symptom mapping in left hemisphere stroke. *Brain* 133, 3269–3280. doi: 10.1093/brain/awq210
- Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., and Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *J. Neurosci.* 29, 10153–10159. doi: 10.1523/JNEUROSCI.2668-09.2009
- Kranjec, A., Ianni, G., and Chatterjee, A. (2013). Schemas reveal spatial relations to a patient with simultanagnosia. *Cortex* 49, 1983–1988. doi: 10.1016/j.cortex.2013.03.005
- Kuperberg, G. R., Lakshmanan, B. M., Greve, D. N., and West, W. C. (2008). Task and semantic relationship influence both the polarity and localization of hemodynamic modulation during lexico-semantic processing. *Hum. Brain Mapp.* 29, 544–561. doi: 10.1002/hbm.20419
- Lingnau, A., Ashida, H., Wall, M. B., and Smith, A. T. (2009). Speed encoding in human visual cortex revealed by fMRI adaptation. *J. Vis.* 9, 1–14. doi: 10.1167/9.13.3
- Luria, A. R. (1959). Disorders of “simultaneous perception” in a case of bilateral occipito-parietal brain injury. *Brain* 82, 437–449. doi: 10.1093/brain/82.3.437
- Maccotta, L., and Buckner, R. L. (2004). Evidence for neural effects of repetition that directly correlate with behavioral priming. *J. Cogn. Neurosci.* 16, 1625–1632. doi: 10.1162/0898929042568451
- Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., et al. (2005). Selecting among competing alternatives: selection and retrieval in the left inferior frontal gyrus. *Cereb. Cortex* 15, 1723–1735. doi: 10.1093/cercor/bhi049
- Nichols, T. E., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. doi: 10.1002/hbm.1058
- Nichols, T., and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* 12, 419–446. doi: 10.1191/0962280203sm341ra
- Peirce, C. S. (1955). *Philosophical Writings of Peirce*. New York, NY: Courier Dover Publications.
- Pezzulo, G., Barca, L., Bocconi, A. L., and Borghi, A. M. (2010). When affordances climb into your mind: advantages of motor simulation in a memory task performed by novice and expert rock climbers. *Brain Cogn.* 73, 68–73. doi: 10.1016/j.bandc.2010.03.002
- Pirog Revill, K., Aslin, R. N., Tanenhaus, M. K., and Bavelier, D. (2008). Neural correlates of partial lexical activation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13111–13115. doi: 10.1073/pnas.0807054105
- Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: a computational account of optic aphasia. *Cogn. Neuropsychol.* 19, 603–639. doi: 10.1080/02643290244000112
- Press, C., Weiskopf, N., and Kilner, J. M. (2012). Dissociable roles of human inferior frontal gyrus during action execution and observation. *Neuroimage* 60, 1671–1677. doi: 10.1016/j.neuroimage.2012.01.118
- Pulvermüller, F. (1999). Words in the brain’s language. *Behav. Brain Sci.* 22, 253–279.
- Raposo, A., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2006). Repetition suppression and semantic enhancement: an investigation of the neural correlates of priming. *Neuropsychologia* 44, 2284–2295. doi: 10.1016/j.neuropsychologia.2006.05.017
- Raposo, A., Moss, H. E., Stamatakis, E. A., and Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia* 47, 388–396. doi: 10.1016/j.neuropsychologia.2008.09.017
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59
- Saxe, R., Brett, M., and Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *Neuroimage* 30, 1088–1096. doi: 10.1016/j.neuroimage.2005.12.062
- Saygin, A. P., McCullough, S., Alac, M., and Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *J. Cogn. Neurosci.* 22, 2480–2490. doi: 10.1162/jocn.2009.21388
- Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., and Hagoort, P. (2013). The suppression of repetition enhancement: a review of fMRI studies. *Neuropsychologia* 51, 59–66. doi: 10.1016/j.neuropsychologia.2012.11.006
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring “how” from “where.” *Neuropsychologia* 41, 280–292. doi: 10.1016/S0028-3932(02)00161-6
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a re-evaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797.
- Van Dam, W. O., van Dijk, M., Bekkering, H., and Rüschemeyer, S.-A. (2012). Flexibility in embodied lexical-semantic representations. *Hum. Brain Mapp.* 33, 2322–2333. doi: 10.1002/hbm.21365
- Vigliocco, G., Vinson, D. P., Lewis, W., and Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cognit. Psychol.* 48, 422–488. doi: 10.1016/j.cogpsych.2003.09.001
- Wall, M. B., Lingnau, A., Ashida, H., and Smith, A. T. (2008). Selective visual responses to expansion and rotation in the human MT complex revealed by functional magnetic resonance imaging adaptation. *Eur. J. Neurosci.* 27, 2747–2757. doi: 10.1111/j.1460-9568.2008.06249.x
- Watson, C. E., Cardillo, E. R., Ianni, G. R., and Chatterjee, A. (2013). Action concepts in the brain: an activation likelihood estimation meta-analysis. *J. Cogn. Neurosci.* 25, 1191–1205. doi: 10.1162/jocn\_a\_00401
- Watson, C. E., and Chatterjee, A. (2011). The functional neuroanatomy of actions. *Neurology* 76, 1428–1434. doi: 10.1212/WNL.0b013e3182166e2c
- Weigelt, S., Singer, W., and Kohler, A. (2012). Feature-based attention affects direction-selective fMRI adaptation in hMT+. *Cereb. Cortex* 23, 2169–2178. doi: 10.1093/cercor/bhs192
- Wiggett, A. J., and Downing, P. E. (2010). Representation of action in occipito-temporal cortex. *J. Cogn. Neurosci.* 23, 1765–1780. doi: 10.1162/jocn.2010.21552
- Willems, R. M., and Francken, J. C. (2012). Embodied cognition: taking the next step. *Front. Psychol.* 3:582. doi: 10.3389/fpsyg.2012.00582
- Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918. doi: 10.1038/jcbfm.1992.127

**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.

Received: 28 January 2014; accepted: 06 May 2014; published online: 26 May 2014.  
 Citation: Watson CE, Cardillo ER, Bromberger B and Chatterjee A (2014) The specificity of action knowledge in sensory and motor systems. *Front. Psychol.* 5:494. doi: 10.3389/fpsyg.2014.00494  
 This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.  
 Copyright © 2014 Watson, Cardillo, Bromberger and Chatterjee. 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.



# Clustering, hierarchical organization, and the topography of abstract and concrete nouns

Joshua Troche<sup>1\*</sup>, Sebastian Crutch<sup>2</sup> and Jamie Reilly<sup>3,4</sup>

<sup>1</sup> Department of Speech, Language, and Hearing Sciences, University of Florida, Gainesville, FL, USA

<sup>2</sup> Department of Neurodegenerative Disease, Dementia Research Centre, Institute of Neurology, University College London, London, UK

<sup>3</sup> Eleanor Saffran Center for Cognitive Neuroscience, Temple University, Philadelphia, PA, USA

<sup>4</sup> Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, USA

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Serge Thill, University of Skövde, Sweden

David Vinson, University College London, UK

## \*Correspondence:

Joshua Troche, University of Florida, 336 Daur Hall, PO Box 100174, Gainesville, FL 32610, USA  
e-mail: jetgator@phhp.ufl.edu

The empirical study of language has historically relied heavily upon concrete word stimuli. By definition, concrete words evoke salient perceptual associations that fit well within feature-based, sensorimotor models of word meaning. In contrast, many theorists argue that abstract words are “disembodied” in that their meaning is mediated through language. We investigated word meaning as distributed in multidimensional space using hierarchical cluster analysis. Participants ( $N = 365$ ) rated target words ( $n = 400$  English nouns) across 12 cognitive dimensions (e.g., polarity, ease of teaching, emotional valence). Factor reduction revealed three latent factors, corresponding roughly to perceptual salience, affective association, and magnitude. We plotted the original 400 words for the three latent factors. Abstract and concrete words showed overlap in their topography but also differentiated themselves in semantic space. This topographic approach to word meaning offers a unique perspective to word concreteness.

**Keywords:** semantic memory, concreteness, abstract concepts, embodied cognition, emotion, magnitude

## INTRODUCTION

A narrow empirical focus on concrete words yields an incomplete picture of the mental lexicon. Today, substantial gaps persist in our knowledge of the cognitive and neural underpinnings of abstract words (e.g., love, truth). Readers of English encounter abstract and concrete words with comparable frequency (Reilly, 2005; Reilly and Kean, 2007). Thus, it is difficult to justify sidestepping the abstract half of the lexicon that poses an empirical challenge.

Despite lopsided attention to concrete words, cognitive science has shown longstanding interest in abstract words (Locke, 1685). Empirical work in abstract-concrete word differences advanced rapidly in the late 1960s when psycholinguists defined *concreteness* and devised a means of measuring its strength. Concreteness, the extent to which a word can be perceived through the senses, is typically measured as a continuous, ratio level variable anchored by a zero point, with zero indicating no evoked perception (Paivio et al., 1968). Psycholinguists have compiled concreteness ratings for many thousands of words across numerous languages with the aim of elucidating the word concreteness effect, a term that reflects the collective advantage for concrete words in a variety of domains, including recall accuracy (Walker and Hulme, 1999), age of acquisition (Gilhooly and Logie, 1980), word list memory (Allen and Hulme, 2006), naming latency (Bleasdale, 1987), word recognition (Schwanenflugel et al., 1988), and dissociations in performance associated with neurological injury (Warrington, 1975, 1981; Breedin et al., 1994; Franklin et al., 1995; Bonner et al., 2009; Jefferies et al., 2009).

It has proven exceptionally difficult to develop a comprehensive theory accounting for the word concreteness effect (Connell and Lynott, 2012). Abstract and concrete words differ on a variety

of non-semantic dimensions, including sound structure and morphological complexity (Reilly and Kean, 2007; Westbury and Moroschan, 2009; Reilly et al., 2012), polysemy and homonymy (Anderson and Nagy, 1991; Crutch and Jackson, 2011). Thus, when one observes a concreteness advantage in a particular task, it is not always clear where the locus of the effect lies (for an example see Kroll and Merves, 1986).

An intimate link between language and abstract word representation forms the backbone of today's dominant model of word concreteness. Paivio's (1991) Dual Coding Theory (DCT) offers a multiple semantics approach to word meaning based on the premise that verbal knowledge and visuoperceptual knowledge reflect two parallel but also highly interactive codes that support a word's meaning. Concrete words benefit from the support of both visual and verbal codes (i.e., they are dually coded), whereas abstract word meaning is mediated almost exclusively through a verbal code. DCT has proven its durability as a model that accounts for word concreteness effects in early childhood language learning and reading, as well as in neurological dissociations in adults (Franklin et al., 1994, 1995; Sadoski and Paivio, 2004; Sadoski, 2005).

Although DCT is compelling in scope, many psycholinguists now recognize the need for finer-grained specificity in delineating the topography of abstract and concrete words. Several approaches to concrete-abstract word representation have recently emerged to address this need. Gallese and Lakoff (2005) and Kousta et al. (2011) have proposed “embodied” approaches to abstract word representation that anchor abstract word meaning in somatic states such as emotion. These embodied approaches offer a radical departure from the dominant view that abstract words are mediated exclusively through symbolic, propositional



knowledge. In one such approach, Kousta et al. (2011) argue that emotion is a powerful latent factor (with somatic and perceptual underpinnings) that underlies the meaning of abstract words (Andrews et al., 2009; Kousta et al., 2009, 2011; Newcombe et al., 2012). Kousta et al. further argued that many past studies of concreteness have confounded the constructs of imageability (i.e., the ability to evoke a mental image) and context availability and that when such confounding factors are tightly controlled, the concreteness advantage either disappears or modestly reverses such that abstract words show a processing advantage (but see Paivio, 2013).

Other theorists attribute abstract-concrete differences to the rapid access to contextual information for concrete words (i.e., context availability) (Schwanenflugel and Shoben, 1983), a greater number of semantic units to support concrete concepts (Plaut and Shallice, 1993) or greater number of semantic predicates for concrete items (Jones, 1985). An alternative formulation has suggested that abstract words have a relatively greater reliance upon associative information, whilst concrete words have a relatively greater reliance upon semantic similarity information (Crutch and Warrington, 2005). The predictions of this “different representational frameworks” hypothesis have been confirmed by a number of recent studies (Duñabeitia et al., 2009), with semantic similarity and association demonstrated to exert a graded effect across the concreteness spectrum (Crutch and Jackson, 2011).

Language researchers have long recognized the role of taxonomic hierarchies in concrete word representation (Rosch, 1973; Lakoff, 1990). For example, *dog* is a basic level concept that has both superordinate (e.g., *animal*) and subordinate distinctions (e.g., *collie*). Much of our knowledge of lexical category structure is derived from studies where participants generate lists of features (e.g., *dog* → *tail*) or associations (e.g., *dog* → *leash*) for concrete target words (Garrard et al., 2001, 2005; Cree and McRae, 2003; Rogers et al., 2004; Cree et al., 2006; Dilkina and Lambon Ralph, 2012). These feature listings yield distance metrics that speak to the family resemblance among concrete words. While these feature listing methods have some utility when applied to abstract words there are inherent weaknesses to this approach for abstract words. Abstract concepts, by their nature, lack the taxonomic hierarchical organization and unambiguous contextual properties imbued within concrete concepts and which make a feature listing method ideal (But see Barsalou and Wiemer-Hastings, 2005; Wiemer-Hastings and Xu, 2005 for examples of feature listing approaches for abstract concepts).

Recently a novel abstract concept feature (ACF) rating approach has been used in combination with multi-dimensional scaling techniques to examine distance metrics and cohesion among abstract words. This approach, developed by Crutch et al. (2012a,b), asks participants to rate the importance of particular types of information for the meaning of a concept. Crutch et al. originally performed this procedure on a corpus of 50 abstract words, spanning nine cognitive dimensions, including emotion, magnitude, and spatial relations. Unlike standard measures of word concreteness, this unique clustering solution revealed that concepts such as VAPOR and ILLUSION aggregate closely within semantic space. Standard semantic distance metrics

gleaned through feature listing approaches or unidimensional ratings often fail to capture such similarities.

Here we performed the ACF in order to determine the clustering attributes of larger corpus of concrete and abstract concepts within a higher dimensional space than was originally employed by Crutch et al. (2012a,b). We measured each word's salience on 12 unique dimensions, including: Sensation, Action; Thought; Emotion; Social Interaction; Time; Space; Quantity; Polarity; Morality; Ease of Modifying; and Ease of Teaching.

Sensorimotor information has long been known to play an important role in the representation of concrete concepts, and a growing body of research has made the argument for the role of affective association in the representation of abstract concepts (Andrews et al., 2009; Kousta et al., 2009, 2011). We included metrics for *Sensation*, *Action*, *Emotion*, and *Polarity* based on the dominance of these variables in previous work. We also included a more nuanced set of dimensions linked to *Social Interaction* and *Thought*. Our rationale for the inclusion of these dimensions stems from the work of Borghi et al. (2011) and Barsalou (1999), who argue for the contributions of social interaction and introspection on abstract word acquisition and representation. We assessed the salience of *Time* in abstract and concrete word meaning due to its role in the temporal unfolding of event structure (Allman and Meck, 2012). We assessed the salience of Spatial information due to its roles both in the organization of geographical concepts, as well as more oblique contributions to metaphor (Zwaan and Yaxley, 2003; Lakoff and Johnson, 2008). We assessed *Quantity* with the aim of tapping the division between numerical and non-numerical semantics (e.g., mass-count distinctions) (Gathercole, 1985). The *Morality* dimension characterizes the social mores that govern behavior which have been hypothesized to reflect a cognitive emotional association complex which can be represented across the prefrontal cortex and limbic system (Moll et al., 2005). *Ease of teaching* reflects variety in both age of acquisition and learning style (e.g., experiential observation vs. explicit verbal instruction) that mark abstract and concrete words (Coltheart et al., 1988; Strain et al., 2002; Reilly et al., 2007). *Ease of Modifying* provides an index of the contextual availability of a word in terms of adjectival description (Schwanenflugel and Shoben, 1983; Schwanenflugel et al., 1992). It should be noted that this is not an exhaustive list of dimensions and that the inclusion of certain dimensions is more empirically/theoretically justified than others. It should also be noted that we were constrained by selecting dimensions that could be easily distinguished and comprehended by the lay participant.

## HYPOTHESES, AIMS, AND SIGNIFICANCE

The DCT is premised upon the interaction of two parallel semantic memory systems, one dedicated to sensory imagery and the other dedicated to language. We hypothesize that word concreteness might ultimately be better contextualized within one semantic system. One might specify such a system in terms of a high dimensional space where word meanings cluster along axes representing key cognitive dimensions (e.g., emotional salience, sensory salience). We hypothesize that this unitary space comprises a topography wherein the meanings of words (both concrete and abstract) are distributed. Here, we investigated the

clustering behaviors of a relatively large ( $N = 400$ ) set of abstract and concrete nouns within a semantic space bounded 12 dimensions, including: Sensation; Action; Thought; Emotional Valence; Social Interaction; Time; Space; Quantity; Polarity; Morality; Ease of Modifying; and Ease of Teaching.

We hypothesize that this topographic approach would produce regions of overlap, as well as distinct clusters corresponding to “concreteness” (e.g., abstract words cluster at the high end of emotional valence). Importantly, the presence of a unitary, multi-dimensional space would obviate the need for an artificial dichotomy such as concreteness by treating this and other psycholinguistic variables as continuous.

## METHODS

### OVERVIEW

We isolated a set of abstract ( $N = 200$ ) and concrete ( $N = 200$ ) English nouns and obtained Likert-scale ratings for each word on 12 variables (dimensions). We then employed factor reduction and hierarchical cluster analysis to model the topography of how these words scaled.

### PARTICIPANTS

Participants included native English speakers recruited through the online crowd-sourcing program, Mechanical Turk. Following trimming procedures aimed at eliminating spurious participants, we isolated a sample ( $N = 365$ ) with an age ranging from 17 to 83 years, (mean = 40.7). Education ranged from 9 to 20 years (mean = 15.4). Sex distribution was 68.2% female.

### MATERIALS AND PROCEDURE

Stimuli included English nouns ( $N = 400$ ) from the MRC Psycholinguistic Database (Coltheart, 1981). Stimuli were pure nouns in that we ensured they had no alternate grammatical class (e.g., *desk* but not *phone*). Target words were either abstract or concrete based on rated concreteness. The MRC database concreteness values reflect a 100–700 scale. In our sample, concrete words had an average concreteness rating of 589 ( $SD = 46.9$ ), whereas abstract words had a rated average of 304 ( $SD = 47.1$ ). There was no overlap in the distributions of abstract and concrete words, and their means were distant ( $Z_{\text{difference}} = 2.38$ ). The list of dimensions chosen for the analysis was not an exhaustive set of dimensions. In order to provide proof of concept that this clustering procedure could prove successful, we sampled words from the tails of the concreteness spectrum (high/low).

### SCALE DEVELOPMENT AND IMPLEMENTATION

Participants rated each of the target words on the following 12 dimensions using a 7-point Likert Scale: 1. Sensation; 2. Action; 3. Thought; 4. Emotional Valence; 5. Social Interaction; 6. Time; 7. Space; 8. Quantity; 9. Polarity; 10. Morality; 11. Ease of Modifying; 12. Ease of Teaching. **Table 1** reflects the wording given to participants.

Each stimulus appeared in randomized order within the context of separate surveys dedicated to each cognitive dimension. Participants were instructed to use the entire scale and to work quickly but carefully.

### DATA COLLECTION

Participants completed ratings via Amazon Mechanical Turk, an online pool of workers from around the globe who perform virtual tasks (Buhrmester et al., 2011). Participants logged into Mechanical Turk, electronically consented, and then completed up to 12 individual surveys, one for each dimension.

### DATA ANALYSES

We excluded participant data that corresponded to any of the following conditions: (1) Taking less than 10 min to complete the survey (less than 1.5 s per response), (2) Using less than half of the seven point scale (i.e., 3 numbers or less) which was considered not following our directions of using the entire scale, or (3) The presence of runs of more than 20 identical consecutive responses (2.5  $SD$  away from the average run mean;  $M = 3.2$ ,  $SD = 6.8$ ). We then performed intraclass correlational analyses in order to measure inter-rater reliability. We also ran correlation analyses between individual item standard deviations and concreteness in order to determine if concreteness led to greater variability in the rating of items.

We first pursued exploratory factor analysis with the goal of reducing the dimensionality and redundancy of the original set of 12 variables. We converted the original ratings into a series of factor scores using the Anderson-Rubin method (Anderson and Rubin, 1956). The factor analyses yielded three latent factors that subsequently define a three-dimensional space upon which distance metrics between any two words can be derived. We report the Euclidean squared coefficient as a metric of semantic distance (Danielsson, 1980).

Using the reduced dataset, we then conducted a hierarchical agglomerative cluster analysis using Ward’s method (Ward, 1963). This procedure iteratively clusters observations into groups in a bottom-up manner until only one large cluster remains. We determined the optimal clustering solution by comparing clusters

**Table 1 | Parameter description.**

Parameter	Definition
Polarity	I relate this word to positive or negative feelings in myself
Sensation	I relate this word to physical feelings like vision, hearing, smelling, etc
Action	I relate this word to actions, doing, performing, and influencing
Thought	I relate this word to mental activity, ideas, opinions, and judgments
Emotion	I relate this word with human emotion
Social interaction	I relate this word with relationships between people
Time	I relate this word with time, order, or duration
Space	I relate this word to position, place, or direction
Quantity	I relate this word to size, amount, or scope
Morality	I relate this word to morality, rules, or anything that governs my behavior
Ease of modifying	I can easily choose an adjective for this word (the ugly truth, whole truth, etc.)
Ease of teaching	This word could be easily taught to a person who does not speak English

from the hierarchical cluster analysis with clusters created by a partitional k-means iterative analysis using Cohen's Kappa (Aldenderfer and Blashfield, 1984). The cluster analysis allowed us to create an empirical metric of how items grouped in the semantic. In other words this allowed us to determine how items grouped on a smaller dimensions as compared to macro dimensions (i.e., Abstract-Concrete).

## RESULTS

### DATA TRIMMING

The first author and a blinded rater showed 99.3% inter-rater agreement on surveys to be excluded (see method for criteria). Of the original 545 surveys, 180 (33%) were eliminated, leaving 365 surveys for final analysis (See Supplementary Material for how many responses were removed per condition). Removal was comparable across all surveys. The intraclass correlation coefficient (ICC) was found to high throughout all 12 surveys with the lowest ICC being 0.991 (see Table 2). Table 3 displays the correlations between item standard deviations and concreteness for each survey dimension. Two of the dimensions showed greater variability for more concrete concepts, three showed no variability differences and seven showed greater variability for more abstract items.

### INDIVIDUAL RATINGS EMOTION

Figure 1 reflects scatterplots of ratings for each of the 12 original dimensions plotted against the a priori concreteness values for each target word. All of the bivariate correlations were significant ( $\alpha \leq 0.01$ ).

### EXPLORATORY FACTOR ANALYSIS

We extracted three latent factors (model fit,  $R^2 = 0.81$ ) from the original set of 12 dimensions (see Table 1). The reduced set of factors and the constituent variables they subsume were as follows: (1) Emotion, Polarity, Social, Morality, Action, Thought; (2) Ease of Teaching, Sensation, Ease of Modifying, Time; (3) Space, Quantity (see Table 4). In terms of nomenclature, we will refer to these latent constructs hereafter as: (1) Affective Association/Social Cognition; (2) Perceptual Salience; and (3) Magnitude.

**Table 2 | Inter-rater reliability.**

Parameter	ICC
Space	0.996
Morality	0.995
Quantity	0.992
Social interaction	0.995
Ease of teaching	0.994
Sensation	0.996
Time	0.944
Action	0.991
Ease of modifying	0.996
Thought	0.993
Emotion	0.997
Polarity	0.996

Table 5 represents relations between the three factors with other salient psycholinguistic variables (e.g., word frequency, age of acquisition). Figure 2 displays the spread between concrete and abstract words within the 3-dimensional space defined by the three factors.

### HIERARCHICAL CLUSTER ANALYSIS

A 12-cluster solution yielded an optimal model (Cohen's Kappa = 0.87). Figure 3 reflects a dendrogram corresponding to this optimal clustering solution. Table 6 reflects quantitative aspects of each cluster in terms of psycholinguistic attributes (e.g., lexical frequency).

The dendrogram shows that most concrete words are contained in the first four clusters (C1–C4), whereas abstract words are mostly found in latter clusters (C5–C12). Focusing on the clusters of abstract words, it is apparent that the level of affective association increases from left to right on the dendrogram. Cluster 8 is also of interest as it is a cluster of concrete words (e.g., chocolate, father) that are high in affective salience and nestled within many other abstract words describing social cognition.

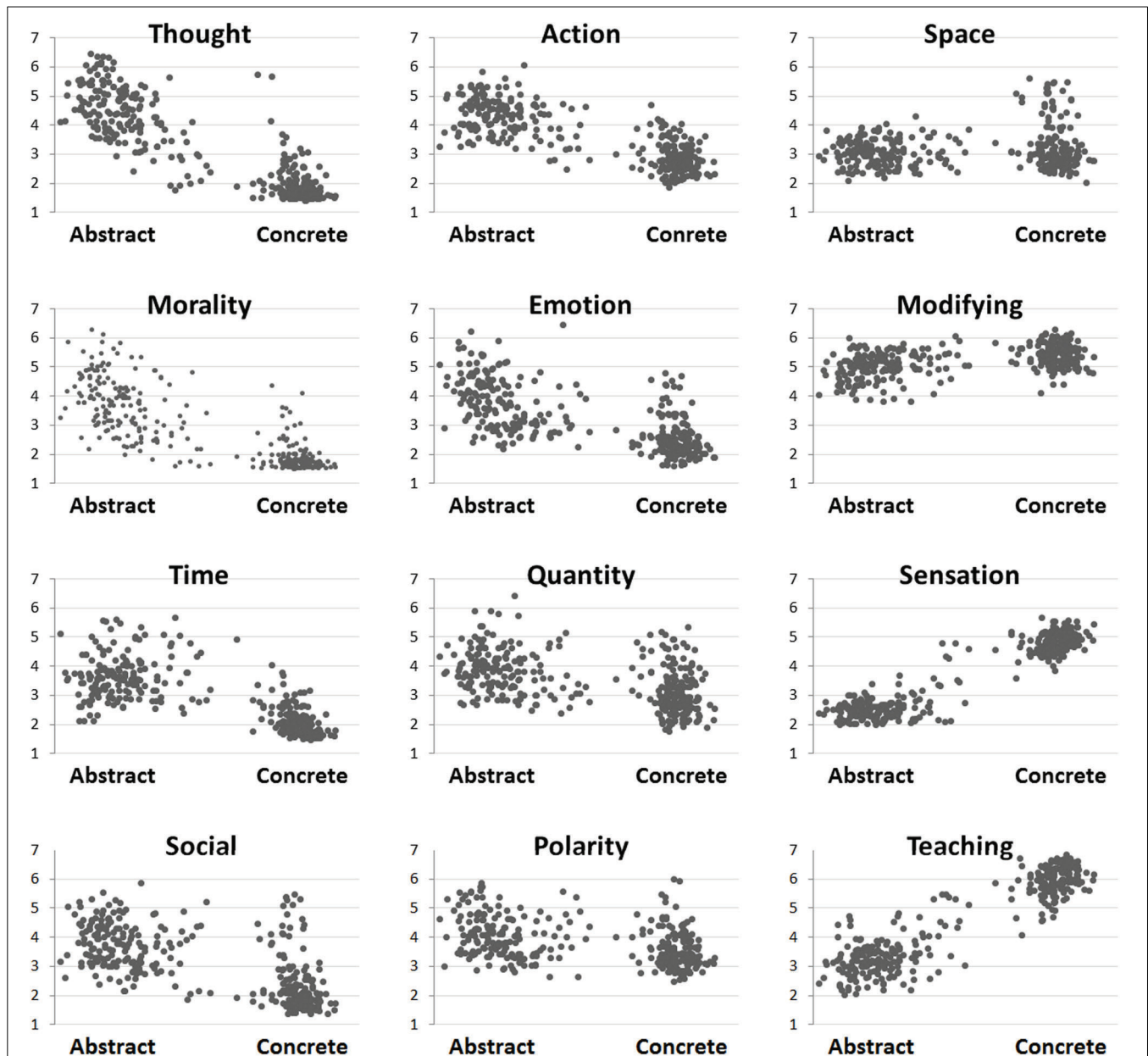
## DISCUSSION

Using hierarchical cluster analyses, we explored the topography of abstract and concrete nouns ( $N = 400$ ). We first defined a multi-dimensional semantic space that was composed of 12 individual predictors, each with precedence as a moderator of concreteness effects. Participants subsequently rated the original set of abstract and concrete nouns on all of the individual dimensions. We then used factor analysis to examine whether the original multi-dimensional semantic space could be reduced. This approach yielded three latent constructs, corresponding roughly to affective association/social cognition, perceptual salience, and magnitude. We then calculated distance metrics for the abstract and concrete words within the semantic space defined by this reduced set of predictors. Abstract and concrete words have both unique and common regions of overlap within semantic space. Moreover, factors such as affective association/social cognition and magnitude appear to play significant roles in delineating this space.

**Table 3 | Correlation of concreteness and dimension SD.**

Dimension	Correlation
Action	0.09
Ease of modifying	−0.55*
Ease of teaching	−0.31*
Emotion	−0.57*
Morality	−0.36*
Polarity	−0.61*
Quantity	0.08
Sensation	0.06
Social interaction	−0.43*
Space	−0.24*
Thought	0.36*
Time	0.23*

\* $\alpha < 0.01$ .



**FIGURE 1 |** Scatterplots of mean Likert scale ratings (1–7; y-axis) for each of the 12 rating dimensions for words from across the concreteness spectrum (x-axis).

There are two primary ways of visualizing these data. The first is at the level of the individual predictors, and the second is through a clustering analysis that considers the predictors together.

#### INDIVIDUAL PREDICTORS

**Figure 1** highlights the variability and weighting across the 12 unique dimensions in isolation prior to factor reduction. The bivariate correlations between concreteness and each predictor vary from strongly positive (e.g.,  $r = 0.94$  for *sensation*) to strongly negative (e.g.,  $r = -0.87$  for *thought*). In addition,

several predictors (e.g.,  $r = 0.10$  for *space*) had relatively flat slopes, indicating that these variables only weakly discriminated concrete from abstract words in isolation. With respect to concreteness, we observed the strongest positive bivariate correlations with sensation ( $r = 0.94$ ) and ease-of-teaching ( $r = 0.92$ ). Sensation, analogous to imageability, is a construct intimately related to concreteness ( $R^2 = 0.88$ ) but one that captures a wider range of somatosensory states. Ease-of-teaching has a close parallel to ease of learning. A vast body of literature investigating age-of-acquisition has shown that the earliest acquired words tend to be concrete (e.g., ball, mama). One common developmental



**Table 4 | Factor analysis/component matrix for dimensions.**

Predictor	Component		
	Factor 1	Factor 2	Factor 3
Emotion	0.905	0.229	−0.027
Polarity	0.880	−0.115	0.235
Social	0.855	0.280	0.090
Morality	0.794	0.479	0.057
Action	0.722	0.517	0.169
Thought	0.719	0.594	0.094
Ease of teaching	−0.376	−0.880	−0.040
Sensation	−0.447	−0.846	−0.026
Ease of modifying	0.104	−0.736	0.310
Time	0.350	0.685	0.319
Space	−0.006	−0.208	0.846
Quantity	0.273	0.412	0.691

The above component matrix was derived using SPSS-18's factor analysis algorithm employing a Varimax rotation with Kaiser normalization. The rotation converged after five iterations.

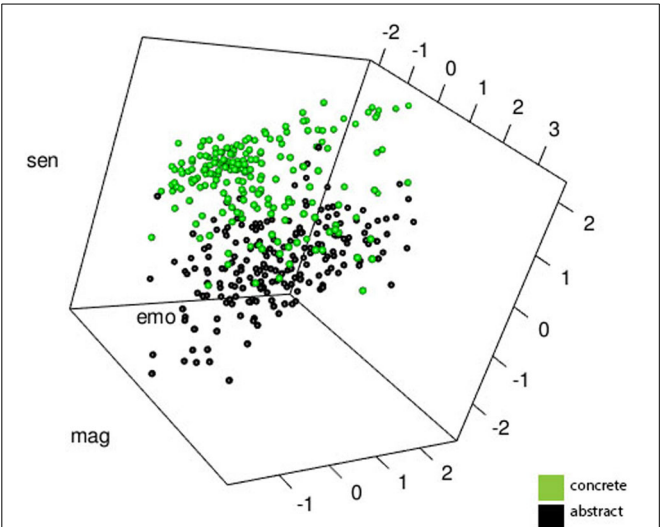
**Table 5 | Psycholinguistic and factor score correlation matrix.**

	Imag	AOA	Frqy	CNC	Fam	Emo	Cnc/Tch	Mag
Imag	1							
AOA	−0.86*	1						
Frqy	0.22*	−0.44*	1					
CNC	0.94*	−0.85*	0.21*	1				
Fam	0.29*	−0.56*	0.44*	0.27*	1			
Emo	−0.36*	0.23*	0.26*	−0.56*	0.16*	1		
Cnc/Tch	0.73*	−0.84*	0.40*	0.78*	0.51*	0	1	
Mag	−0.01	−0.03	0.09	−0.06	0.24*	0	0	1

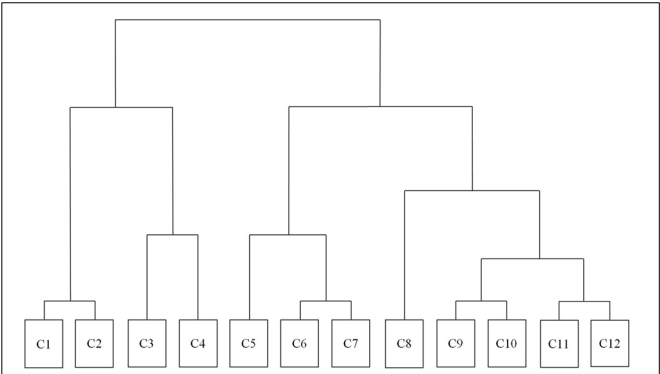
Imag, Imageability; AOA, Age of Acquisition; Frqy, Frequency; CNC, concreteness; Fam, Familiarity; Emo, Emotion/Social Cognition; Cnc/Tch, Concreteness/Ease of Teaching; Mag, Magnitude; \* $p < 0.01$ .

explanation is that the salience of a concrete word’s referent facilitates a fast and durable mapping (Gilhooly and Logie, 1980; Bloom, 1998). Abstract words, in contrast, have no physical referent and must therefore be learned through alternate means, often through nuanced experiences with concrete objects and emotions. For example, one must first learn “sad” before acquiring a more abstract state such as “melancholy.”

In addition to strong positive relationships with concreteness, we also observed several robust negative correlations, including *thought* ( $r = -0.87$ ) and *morality* ( $r = -0.81$ ). Participants rated *thought* according to the salience of ideas, opinions, judgments, and mental operations. Many words that are considered classically abstract are often defined as “the feeling of X.” Thus, the strong negative correlation between concreteness and *thought* reflects a logical property of abstract words (i.e., they tend to often denote unobservable mental states). *Morality* is similar to *thought* in that this construct often denotes phenomena that are not directly observable but instead reflect complex social mores that govern and denote behavior (e.g., truth, honesty).



**FIGURE 2 | Three Dimensional Scatterplot Representing Abstract and Concrete Word Meaning.** This view represents rotation about the axes/planes defined by the factors: Sens, sensation; Mag, magnitude; and Emo, emotion.



**FIGURE 3 | Dendrogram of hierarchical cluster analysis.** Each cluster has been given a cluster number (e.g., C1,C2). The words inside each cluster can be found in Supplementary Material.

MULTIDIMENSIONAL SOLUTION

The strength of this approach lies not within individual predictors but in a solution that considers all such variables simultaneously. This multi-dimensional solution yielded a dynamic structure whereby abstract and concrete words can be differentiated. We view two properties of the observed topography as particularly salient: (A) Abstract and concrete words have unique topographies within a multi-dimensional space defined by affective association/social cognition, magnitude, and perceptual salience; (B) The topography of abstract and concrete words also overlap within this space. For example, *father* and *love* load high on emotion and ultimately cluster together despite the fact that *father* is classically considered concrete and “love” as abstract. It should be noted that this clustering emerges despite all words being rate independently (i.e., there were no ratings of the direct association between any pairs of concepts).

**Table 6 | Psycholinguistic properties of clusters.**

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Imag	595.54	618.53	587.05	608.22	340.81	240.38	289.52	563.09	370.13	383.73	362.67	309.70
AOA	280.07	249.86	311.50	270.00	500.53	617.00	495.15	221.00	512.50	463.38	442.67	478.00
Frqy	14.05	53.98	38.85	59.71	6.14	1.48	9.37	223.05	7.40	23.50	45.46	22.97
CNC	598.67	605.60	585.65	588.21	326.73	285.67	301.10	574.48	272.25	303.00	312.53	310.64
Fam	519.33	564.40	534.70	557.06	485.19	424.43	505.00	584.95	465.63	529.27	557.59	528.48
Emo	-0.86	-0.15	-0.88	-0.65	-0.80	-0.34	0.33	1.23	1.34	1.67	1.62	0.60
Per	0.54	1.25	0.68	0.80	-1.59	-1.59	-0.95	1.69	-0.45	-0.02	0.05	-0.67
Mag	-0.85	-0.55	2.55	0.60	0.54	-1.48	-0.45	-0.30	-1.60	-0.17	0.99	0.65

*Imag*, Imageability; *AOA*, Age of Acquisition; *Frqy*, Frequency; *CNC*, concreteness; *Fam*, Familiarity; *Emo*, Emotion/Social Cognition; *Per*, Perceptual Salience; *Mag*, Magnitude.

### **The topographies of abstract and concrete words are unique**

While affective association/social cognition and concreteness/perceptual salience have been regularly indicated as dimensions that underlie the representations of concrete and abstract concepts, the role of magnitude is less clear.

The factor analysis identified a latent variable reflecting a combination of space and quantity. We interpreted this amalgamation as corresponding roughly to the construct of magnitude. Magnitude in this context reflects both the scalar features of concrete words (e.g., how large?, how hot?, how loud?) but also gradations of many abstract emotions (e.g., irritated < angry < infuriated). Walsh (2003) has argued that such a magnitude system detects and appreciate such gradations. Neurological damage to regions of the parietal lobes (e.g., cortical basal degeneration) results in deficits for estimating and appreciating many magnitude distinctions, including time, physical size, and affect (i.e., emotional blunting; Gibb et al., 1989; Crutch et al., 2012a,b).

Magnitude is a construct that has previously received attention in the psycholinguistic literature, particularly with respect to spatial metaphor comprehension (Lakoff, 1990, 2012; Barsalou and Wiemer-Hastings, 2005; Jefferies et al., 2009; Connell and Lynott, 2012). During semantic relatedness tasks (e.g., match two related pictures from a field of three), both healthy adults and patients with neurological disorders (e.g., stroke aphasia) tend to take longer to match items that are more geographically distant (e.g., London:New York vs. London:Manchester; Crutch and Warrington, 2003), or items that appear in reverse-lexical order (e.g., basement:attic vs. attic:basement; Zwaan and Yaxley, 2003). Similar findings have been reported for the directionality and congruency of spatial metaphors with respect to one's own body (Zwaan and Taylor, 2006). Thus, our scaling results confirm a place of prominence and a dimension of discrimination for magnitude and related variables (e.g., polarity, valence) in supporting the meanings of both abstract and concrete words.

### **The topographies of abstract and concrete words also overlap**

The scatterplot in Figure 2 demonstrates several regions of significant overlap in the topographies of abstract and concrete words. The area of highest overlap was apparent for words at the high end of the affective association/social cognition dimension. Concrete

words that loaded high on the affective association/social cognition factor (e.g., father, chocolate) were closer via distance metrics in semantic space to abstract words (e.g., love, justice) than they were to other concrete or abstract words lacking an affective association/social cognition component (e.g., aspect, paradigm, fisherman, and banana). This underscores the importance of emotional valence in word meaning. Altarriba et al. (1999) have argued that emotional valence can be viewed as orthogonal to concreteness and should accordingly be viewed as an independent dimension of word meaning (i.e., there are abstract, concrete, and emotion words). More recently Kousta et al. (2011) have argued for an embodied theory with emotional information being the main contributor to the representation of abstract concepts (Etkin et al., 2006; Vigliocco et al., 2013).

The overlap of our topographies in areas of high affective association/social cognition suggest that while abstract concepts likely rely more on affective association/social cognition for their representation, concrete concepts can also be greatly influenced by affective association/social cognition. There is also the indication that high affective association/social cognition can lead to abstract concepts becoming more tangible, that is, more concrete, as indicated by the positive association between affective association/social cognition and imageability. This overlap may lead to a strengthening of the networks for these concepts leading to collective processing advantages that Kousta et al. (2009) found for words high in affective association. It should be noted that these areas of overlap are even more surprising as we only chose concepts that were at the extreme ends of the concreteness spectrum.

The ACF approach allowed us to create a single multidimensional semantic space. This approach obviates the need for multiple semantic systems (e.g., language for abstract words, percepts for concrete words). By treating this topography as a continuous space, word meaning can be distributed in a flexible way that is untethered to any particular artificial dichotomy (e.g., abstract-concrete, imageable-non-imageable; for another unitary semantics account see Vigliocco et al., 2004). In this approach words were rated individually, therefore words collocated in this semantic space represent similar underlying properties and not merely linguistic properties. It should be noted that early work on dimensionality in semantics by Osgood et al. (1954) also found three dimensions that held importance in the evaluation of

concepts: evaluation, potency, and activity. This work, however, has mostly focused on determining the connotation of a concept, object, or event.

It still remains an open question, however, whether this semantic space is neurologically real or just a product of our data. We attempted to test this question through the use of a behavioral task with a patient with aphasia (Crutch et al., 2013). The patient, a 65 years old male, had a history of global aphasia which resolved into a mixed non-fluent aphasia. This patient, SKO, displayed deficits in verbal comprehension and phonological-orthographic transcoding. The patient was given a spoken word to written word matching paradigm. This consisted of SKO being shown two words and then being asked to point to the word just spoken by the examiner. The pairs of words were varied by distance. Some of the words were close in distance in the semantic space created in the current study while others were far. As we had predicted, pairs of words closer in semantic distance lead to greater interference than those further. We also determined that ACF ratings were better at predicting deficits than another common and well researched method of determining the strength of word association, latent semantic analysis (Landauer and Dumais, 1997). We argue that these findings suggest that this semantic space is somewhat representative of the underlying representation of concepts.

While the findings here are promising more can be done to improve the current semantic space. The 12 predictors chosen do not constitute an exhaustive list of potentially relevant dimensions. The sensation dimension, for instance, could be broken up into several dimensions (Visual, Auditory, etc.), which might lead to greater differentiation across more concrete concepts. The inclusion of greater dimensionality would also help decrease the amount of unexplained variance in the model, however, this will happen to a smaller and smaller degree as more dimensions are added. Also now that we have shown proof of concept, future work would benefit from expanding the concepts across grammatical class and concreteness (e.g., more middling concreteness concepts) as this will likely create a semantic space which is more ecologically valid.

Overall, this topographic approach also readily lends itself to computational investigations whereby particular dimensions (e.g., magnitude) or individual clusters (e.g., high emotion, low magnitude) might be selectively lesioned as functions of regional brain damage. Much of the utility of this approach will depend on specifying the nature and fluidity of the topography.

## ACKNOWLEDGMENTS

We are grateful to Alison O'Donoghue for her assistance with numerous aspects of this project. This work was supported by US Public Health Service grants DC010197 (Jamie Reilly) and DC013063 (Jamie Reilly), Alzheimer Research UK Senior Research Fellowship (Sebastian Crutch), and by the NIHR Queen Square Dementia Biomedical Research Unit (Sebastian Crutch).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00360/abstract>

## REFERENCES

- Aldenderfer, M. S., and Blashfield, R. K. (1984). *Cluster Analysis*. London: Sage Publications.
- Allen, R., and Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *J. Mem. Lang.* 55, 64–88. doi: 10.1016/j.jml.2006.02.002
- Allman, M. J., and Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain* 135, 656–677. doi: 10.1093/brain/awr210
- Altarriba, J., Bauer, L., and Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behav. Res. Methods* 31, 578–602. doi: 10.3758/BF03200738
- Anderson, R. C., and Nagy, W. E. (1991). “Word meanings,” in *Handbook of Reading Research*, Vol. 2, eds R. Barr, M. L. Kamil, P. B. Mosenthal, and P. D. Pearson (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc), 690–724.
- Anderson, T. W., and Rubin, H. (1956). “Statistical inference in factor analysis,” in *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 5 (London: Cambridge University Press), 111–150.
- Andrews, M., Vigliocco, G., and Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychol. Rev.* 116, 463–498. doi: 10.1037/a0016261
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behav. Brain Sci.* 22, 637–660. doi: 10.1017/S0140525X99532147
- Barsalou, L. W., and Wiemer-Hastings, K. (2005). “Situating abstract concepts,” in *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*, D. Pecher and R. Zwaan (New York, NY: Cambridge University Press), 129–163. doi: 10.1017/CBO9780511499968.007
- Bleasdale, F. A. (1987). Concreteness-dependent associative priming: separate lexical organization for concrete and abstract words. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 582–594. doi: 10.1037/0278-7393.13.4.582
- Bloom, L. (1998). “Language acquisition in its developmental context,” in *Handbook of Child Psychology: Volume 2: Cognition, Perception, and Language* ed W. Damon (Hoboken, NJ: John Wiley & Sons Inc), 309–370.
- Bonner, M. F., Vesely, L., Price, C., Anderson, C., Richmond, L., Farag, C., et al. (2009). Reversal of the concreteness effect in semantic dementia. *Cogn. Neuropsychol.* 26, 568–579. doi: 10.1080/02643290903512305
- Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., Scorolli, C., Borghi, A., et al. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Front. Psychol.* 2:15. doi: 10.3389/fpsyg.2011.00015
- Breedin, S. D., Saffran, E. M., and Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cogn. Neuropsychol.* 11, 617–660. doi: 10.1080/02643299408251987
- Buhrmester, M., Kwang, T., and Gosling, S. D. (2011). Amazon's mechanical turk. *Perspect. Psychol. Sci.* 6, 3–5. doi: 10.1177/1745691610393980
- Coltheart, M. (1981). The MRC psycholinguistic database. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 33A, 497–505. doi: 10.1080/14640748108400805
- Coltheart, V., Laxon, V. J., and Keating, C. (1988). Effects of word imageability and age of acquisition on children's reading. *Br. J. Psychol.* 79, 1–12. doi: 10.1111/j.2044-8295.1988.tb02270.x
- Connell, L., and Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition* 125, 452–465. doi: 10.1016/j.cognition.2012.07.010
- Cree, G. S., McNorgan, C., and McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: implications for theories of semantic memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 643–658. doi: 10.1037/0278-7393.32.4.643
- Cree, G. S., and McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J. Exp. Psychol. Gen.* 132, 163–201. doi: 10.1037/0096-3445.132.2.163
- Crutch, S. J., and Jackson, E. C. (2011). Contrasting graded effects of semantic similarity and association across the concreteness spectrum. *Q. J. Exp. Psychol.* 64, 1388–1408. doi: 10.1080/17470218.2010.543285
- Crutch, S. J., Lehmann, M., Schott, J. M., Rabinovici, G. D., Rossor, M. N., and Fox, N. C. (2012a). Posterior cortical atrophy. *Lancet Neurol.* 11, 170. doi: 10.1016/S1474-4422(11)70289-7
- Crutch, S. J., Troche, J., Reilly, J., and Ridgway, G. R. (2013). Abstract conceptual feature ratings: the role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Front. Hum. Neurosci.* 7:186. doi: 10.3389/fnhum.2013.00186

- Crutch, S. J., and Warrington, E. K. (2003). Spatial coding of semantic information: knowledge of country and city names depends on their geographical proximity. *Brain* 126, 1821–1829. doi: 10.1093/brain/awg187
- Crutch, S. J., and Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain* 128, 615–627. doi: 10.1093/brain/awh349
- Crutch, S. J., Williams, P., Ridgway, G. R., and Borgenicht, L. (2012b). The role of polarity in antonym and synonym conceptual knowledge: evidence from stroke aphasia and multidimensional ratings of abstract words. *Neuropsychologia* 50, 2636–2644. doi: 10.1016/j.neuropsychologia.2012.07.015
- Danielsson, P.-E. (1980). Euclidean distance mapping. *Comput. Graph. Image Process.* 14, 227–248. doi: 10.1016/0146-664X(80)90054-4
- Dilkina, K., and Lambon Ralph, M. A. (2012). Conceptual structure within and between modalities. *Front. Hum. Neurosci.* 6:333. doi: 10.3389/fnhum.2012.00333
- Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., and Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: evidence from the visual-world paradigm. *Cognition* 110, 284–292. doi: 10.1016/j.cognition.2008.11.012
- Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., and Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51, 871–882. doi: 10.1016/j.neuron.2006.07.029
- Franklin, S., Howard, D., and Patterson, K. (1994). Abstract word meaning deafness. *Cogn. Neuropsychol.* 11, 1–34. doi: 10.1080/02643299408251964
- Franklin, S., Howard, D., and Patterson, K. (1995). Abstract word anomia. *Cogn. Neuropsychol.* 12, 549–566. doi: 10.1080/02643299508252007
- Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310
- Garrard, P., Lambon Ralph, M. A., Hodges, J. R., and Patterson, K. (2001). Prototypicality, distinctiveness, and intercorrelation: analyses of the semantic attributes of living and nonliving concepts. *Cogn. Neuropsychol.* 18, 125–174. doi: 10.1080/02643290125857
- Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., and Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. *Brain Lang.* 93, 79–94. doi: 10.1016/j.bandl.2004.08.003
- Gathercole, V. C. (1985). "He has too much hard questions:" the acquisition of the linguistic mass-count distinction in much and many. *J. Child Lang.* 12, 395–415. doi: 10.1017/S0305000900006504
- Gibb, W., Luther, P., and Marsden, C. (1989). Corticobasal degeneration. *Brain* 112, 1171–1192. doi: 10.1093/brain/112.5.1171
- Gilhooly, K., and Logie, R. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behav. Res. Methods* 12, 395–427. doi: 10.3758/BF03201693
- Jefferies, E., Patterson, K., Jones, R. W., and Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology* 23, 492–499. doi: 10.1037/a0015452
- Jones, G. V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain Lang.* 24, 1–19. doi: 10.1016/0093-934X(85)90094-X
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34. doi: 10.1037/a0021446
- Kousta, S.-T., Vinson, D. P., and Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112, 473–481. doi: 10.1016/j.cognition.2009.06.007
- Kroll, J. F., and Merves, J. S. (1986). Lexical access for concrete and abstract words. *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 92–107. doi: 10.1037/0278-7393.12.1.92
- Lakoff, G. (1990). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press.
- Lakoff, G. (2012). Explaining embodied cognition results. *Top. Cogn. Sci.* 4, 773–785. doi: 10.1111/j.1756-8765.2012.01222.x
- Lakoff, G., and Johnson, M. (2008). *Metaphors We Live By*. Chicago, IL: University of Chicago press.
- Landauer, T. K., and Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* 104, 211. doi: 10.1037/0033-295X.104.2.211
- Locke, J. (1685). "An essay on human understanding," in *Of Abstract and Concrete Terms*, Chapter 7, ed M. Ruse (New York, NY; Amherst, MA: Prometheus books), 384–385.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., and Grafman, J. (2005). The neural basis of human moral cognition. *Nat. Rev. Neurosci.* 6, 799–809. doi: 10.1038/nrn1768
- Newcombe, P. I., Campbell, C., Siakaluk, P. D., and Pexman, P. M. (2012). Effects of emotional and sensorimotor knowledge in semantic processing of concrete and abstract nouns. *Front. Hum. Neurosci.* 6:275. doi: 10.3389/fnhum.2012.00275
- Osgood, C. E., Sebeok, T. A., Gardner, J. W., Carroll, J. B., Newmark, L. D., Ervin, S. M., et al. (1954). Psycholinguistics: a survey of theory and research problems. *J. Abnorm. Soc. Psychol.* 49(4 pt 2), i. doi: 10.1037/h0063655
- Paivio, A. (1991). Dual coding theory: retrospect and current status. *Can. J. Psychol.* 45, 255–287. doi: 10.1037/h0084295
- Paivio, A. (2013). Dual coding theory, word abstractness, and emotion: a critical review of Kousta et al. (2011). *J. Exp. Psychol. Gen.* 142, 282–287. doi: 10.1037/a0027004
- Paivio, A., Yuille, J. C., and Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. Exp. Psychol.* 76(1 pt 2), 1–25. doi: 10.1037/h0025327
- Plaut, D. C., and Shallice, T. (1993). Deep dyslexia: a case study of connectionist neuropsychology. *Cogn. Neuropsychol.* 10, 377–500. doi: 10.1080/02643299308253469
- Reilly, J. (2005). *A tale of two imageabilities: An interaction of sound and meaning in natural language perception*. (66), Dissertation Abstracts International: Section B: The Sciences and Engineering, US. Available online at: <https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2005-99024-143&site=ehost-live>
- Reilly, J., Chrysikou, E. G., and Ramey, C. H. (2007). Support for hybrid models of the age of acquisition of English nouns. *Psychon. Bull. Rev.* 14, 1164–1170. doi: 10.3758/BF03193107
- Reilly, J., and Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: analyses and theoretical implications. *Cogn. Sci.* 31, 157–168. doi: 10.1080/03640210709336988
- Reilly, J., Westbury, C., Kean, J., and Peelle, J. E. (2012). Arbitrary symbolism in natural language revisited: when word forms carry meaning. *PLoS ONE* 7:e42286. doi: 10.1371/journal.pone.0042286
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychol. Rev.* 111, 205–235. doi: 10.1037/0033-295X.111.1.205
- Rosch, E. H. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350. doi: 10.1016/0010-0285(73)90017-0
- Sadoski, M. (2005). A dual coding view of vocabulary learning. *Read. Writing Q.* 21, 221–238. doi: 10.1080/10573560590949359
- Sadoski, M., and Paivio, A. (2004). A dual coding theoretical model of reading. *Theor. Models Processes Read.* 5, 1329–1362. doi: 10.1598/0872075028.47
- Schwanenflugel, P. J., Akin, C., and Luh, W.-M. (1992). Context availability and the recall of abstract and concrete words. *Mem. Cognit.* 20, 96–104. doi: 10.3758/BF03208259
- Schwanenflugel, P. J., Harnishfeger, K. K., and Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *J. Mem. Lang.* 27, 499–520. doi: 10.1016/0749-596X(88)90022-8
- Schwanenflugel, P. J., and Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 82–102. doi: 10.1037/0278-7393.9.1.82
- Strain, E., Patterson, K., and Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 207–214. discussion: 215–220. doi: 10.1037/0278-7393.28.1.207
- Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2013). The neural representation of abstract words: the role of emotion. *Cereb. Cortex*. doi: 10.1093/cercor/bht025. [Epub ahead of print].
- Vigliocco, G., Vinson, D. P., Lewis, W., and Garrett, M. F. (2004). Representing the meanings of object and action words: the featural and unitary semantic space hypothesis. *Cogn. Psychol.* 48, 422–488. doi: 10.1016/j.cogpsych.2003.09.001
- Walker, I., and Hulme, C. (1999). Concrete words are easier to recall than abstract words: evidence for a semantic contribution to short-term serial recall.



- J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1256–1271. doi: 10.1037/0278-7393.25.5.1256
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244. doi: 10.1080/01621459.1963.10500845
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Q. J. Exp. Psychol.* 27, 635–657. doi: 10.1080/14640747508400525
- Warrington, E. K. (1981). Concrete word dyslexia. *Br. J. Psychol.* 72, 175–196. doi: 10.1111/j.2044-8295.1981.tb02175.x
- Westbury, C., and Moroschan, G. (2009). Imageability x phonology interactions during lexical access: effects of modality, phonological neighbourhood, and phonological processing efficiency. *Ment. Lex.* 4, 115–145. doi: 10.1075/ml.4.1.05wes
- Wiemer-Hastings, K., and Xu, X. (2005). Content differences for abstract and concrete concepts. *Cogn. Sci.* 29, 719–736. doi: 10.1207/s15516709cog0000\_33
- Zwaan, R. A., and Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *J. Exp. Psychol. Gen.* 135, 1–11. doi: 10.1037/0096-3445.135.1.1
- Zwaan, R. A., and Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychon. Bull. Rev.* 10, 954–958. doi: 10.3758/BF03196557
- 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.

Received: 10 January 2014; accepted: 06 April 2014; published online: 28 April 2014.

Citation: Troche J, Crutch S and Reilly J (2014) Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Front. Psychol.* 5:360. doi: 10.3389/fpsyg.2014.00360

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Troche, Crutch and Reilly. 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.



# Toward a more embedded/extended perspective on the cognitive function of gestures

Wim T. J. L. Pouw<sup>1\*</sup>, Jacqueline A. de Nooijer<sup>1</sup>, Tamara van Gog<sup>1</sup>, Rolf A. Zwaan<sup>1</sup> and Fred Paas<sup>1,2</sup>

<sup>1</sup> Department of Social Sciences, Institute of Psychology, Erasmus University Rotterdam, Rotterdam, South Holland, Netherlands

<sup>2</sup> Early Start Research Institute, University of Wollongong, Wollongong, NSW, Australia

## Edited by:

Guy Dove, University of Louisville,  
USA

## Reviewed by:

Hecke Schrobsdorff,  
Max-Planck-Institute for Dynamics and  
Self-Organization, Germany  
Olivier Le Guen, Centro de  
Investigaciones y Estudios Superiores  
en Antropología Social, Mexico

## \*Correspondence:

Wim T. J. L. Pouw, Department of  
Social Sciences, Institute of  
Psychology, Erasmus University  
Rotterdam, Burgemeester Oudlaan  
50, TBuilding, Room T13-36, 3062 PA  
Rotterdam, South Holland,  
Netherlands  
e-mail: pouw@fsw.eur.nl

Gestures are often considered to be demonstrative of the embodied nature of the mind (Hostetter and Alibali, 2008). In this article, we review current theories and research targeted at the intra-cognitive role of gestures. We ask the question how can gestures support internal cognitive processes of the gesturer? We suggest that extant theories are in a sense *disembodied*, because they focus solely on embodiment in terms of the sensorimotor neural precursors of gestures. As a result, current theories on the intra-cognitive role of gestures are lacking in explanatory scope to address how gestures-as-bodily-acts fulfill a cognitive function. On the basis of recent theoretical appeals that focus on the possibly embedded/extended cognitive role of gestures (Clark, 2013), we suggest that gestures are external physical tools of the cognitive system that *replace and support* otherwise solely internal cognitive processes. That is gestures provide the cognitive system with a stable external physical and visual presence that can provide means to think with. We show that there is a considerable amount of overlap between the way the human cognitive system has been found to use its environment, and how gestures are used during cognitive processes. Lastly, we provide several suggestions of how to investigate the embedded/extended perspective of the cognitive function of gestures.

**Keywords:** gestures, embodied cognition, embedded cognition, extended cognition

## INTRODUCTION

Gestures reflect internal cognitive processes. This is arguably the most fundamental, uncontroversial, and straightforward assumption in the current literature concerning gesticulation. Gestures provide a “window on the mind” (Goldin-Meadow, 2003), which provides a peek into the “embodied nature of the mind” (Hostetter and Alibali, 2008). In less metaphorical terms, it is argued that gestures are direct outcomes of multimodal, sensorimotor or embodied representations that constitute thought processes and speech production. Although not all theoretical perspectives on the function and underpinnings of gestures suggest a purely sensorimotor based approach to mental representations (see Krauss, 1998; Kita, 2000 for alternative views), it is commonly held that activation of the motor-system supports speech production and thought, at least when the conceptual content is visuospatial in nature (Alibali, 2005). Several perspectives on gesticulation (e.g., McNeill, 1992; Kita, 2000; Wesp et al., 2001) have abandoned the view that gestures are merely communicative tools that are elicited *after* central cognitive processes (e.g., lexical retrieval, conceptualization) have taken place (Graham and Argyle, 1975; Kendon, 1994). Instead, in these perspectives the motor-system has been upgraded from a mere output system to a constitutive system for (some of the) central processes underlying thought and speech production. This resonates well with a wider movement in embodied cognitive science (Wilson, 2002; Shapiro, 2010) in which mental representations are thought to be multimodal (Barsalou, 1999, 2008; Svensson, 2007) and coupled to the body’s current state (Glenberg and Kaschak, 2002).

In this article, we focus on the possible intra-cognitive function of gestures, as opposed to their inter-cognitive or communicative function, which we will touch upon only briefly. That is, gestures seem to support internal cognitive processes of the gesturer (e.g., Rauscher et al., 1996; Goldin-Meadow et al., 2001; Morsella and Krauss, 2004; Marstaller and Burianová, 2013). We argue that the current theoretical “embodied” movement in gesture research has fueled the upsurge of inquiry into the beneficial role of gestures in cognitive processes such as speech and visuospatial cognition, but that this line of thought is underspecified with regard to explaining how gestures *as bodily movements* aid cognitive processing. In a sense, current perspectives on gestures are still *disembodied* and too *internalistic* because they seem to implicitly reduce gestures to *cognitively trivial* bodily outputs of (sensorimotor) neural precursors.

We seek to provide a more embodied account of gesticulation on the basis of recent philosophical and theoretical appeals within embodied cognitive science (e.g., Wilson, 2002) that focus on the possibly embedded/extended role of gestures (Kirsh, 1995; Clark, 2008, 2013; Wheeler, 2013), and a review of related empirical literature (e.g., Gray and Fu, 2004; Kirsh, 2009). This account is “more embodied” because embedded/extended perspectives traditionally seek to provide an anti-internalist perspective on cognition (e.g., Hutchins, 1995a), in which cognition is understood as being on-line, that is, being tightly coupled with, embedded in, if not extended over, the body and the environment (Shapiro, 2010). This stands in stark contrast with more internalist notions of embodiment that are currently dominating the gesture literature

and that focus on decoupled, or “off-line” cognition and the sensorimotor nature of mental representations (Wilson, 2002). We suggest that the embedded/extended account of the cognitive function of gestures could be successful in explaining how gestures fulfill a cognitive function if it makes clear how gestures as self-generated bodily acts *generate and support* rather than execute thought processes (Clark, 2013). Therefore, we focus on the idea that gestures may at times serve as external tools of the cognitive system that *replace and support* otherwise solely internal cognitive processes. By reviewing research on the beneficial role of gesture production in (visuo-spatial) cognition (e.g., Chu and Kita, 2008; Delgado et al., 2011) and connecting the resulting insights with research on embedded cognition (e.g., Kirsh and Maglio, 1994; Hutchins, 1995a; Gray and Fu, 2004) we aim to contribute to a more embedded/extended account of gestures.

Before we will elaborate on the main goals of this paper, we need to point out what this article is not about. First, we do not suggest that current perspectives in the gesture literature are incorrect. In fact, our embedded/extended perspective is largely complementary to, and in some instances builds on, contemporary accounts of the function of gestures we review here. Second, although we argue in favor of a more embodied account of gestures and their cognitive function, this does not require us to make any additional, more radical, claims about the supposed sensorimotor nature of conceptual representations that are currently under discussion in the literature (e.g., Dove, 2010; Arbib et al., 2014; Zwaan, in press). Third, we will not provide philosophical claims about whether gestures should be considered as an extended as opposed to an embedded cognitive phenomenon (e.g., Adams and Aizawa, 2001; Clark, 2008, 2013; Wheeler, 2013). That is, we do not make explicit claims about whether gestures as extra-neural events are part of the cognitive process (extended claim) or whether gestures merely support internal cognitive processes but strictly speaking should not be considered as part of the cognitive process (embedded claim). Rather, we aim to provide an empirical view through the embedded/extended perspective, on the basis of the shared anti-internalist goal of these perspectives, by focusing on extra-neural factors that support, shape, and replace internal cognitive processes. We suggest that our embedded/extended account of the cognitive function of gestures can fill an explanatory gap in the current literature concerning the possible intra-cognitive role of gestures and is supported by extant findings.

This article is structured into four main sections. The next section reviews findings that show that co-speech and -thought gestures have a (beneficial) cognitive function (primarily in visuo-spatial cognition). Section three provides an overview of some important theoretical perspectives on the role of gestures in cognition. We suggest that the current theoretical perspectives on the function and underpinnings of gestures leave an explanatory gap concerning how gestures as external bodily acts might be conducive to internal cognitive processes. Having exposed the explanatory gap, we introduce an embedded/extended account of gestures (Clark, 2008, 2013) and provide a new interpretation of the research reviewed in the previous section in light of recent research in the field of embedded cognition (Kirsh and Maglio, 1994; Ballard et al., 1995; Gray and Fu, 2004; Kirsh, 2009; Risko et al., 2013). Finally, we summarize and discuss our main points.

## THE FUNCTION OF GESTURE: EMPIRICAL EVIDENCE

### THE INTER-COGNITIVE ROLE OF GESTURES

Before we consider evidence for the beneficial or supportive role of gestures for cognitive processes, it is important to acknowledge the evidence for the common assertion that gestures fulfill a communicative function. When speakers produce gestures, this seems to be intended to increase listeners' understanding of their message. Indeed, when speaker and listener are face-to-face, more gestures with semantic content are produced than when there is no visual contact (Alibali et al., 2001). Also, when speakers are aware of listeners' knowledge gaps, they tend to convey the information unknown to listeners in both speech and gesture, while they tend to only use verbal information when relevant knowledge is already shared between the interlocutors (Holler and Stevens, 2007). These results suggest that speakers adjust their gestures for their listeners' benefit. And indeed, listeners' comprehension has been shown to improve by speakers' use of gestures from an early age on. For example, 3- to 5-year-olds understand indirect requests (Kelly, 2001) and new abstract concepts (Valenzano et al., 2003) better when the request is accompanied by deictic (i.e., pointing) gestures. In addition, preschoolers understand complex spoken messages better when these are accompanied by representational gestures (McNeil et al., 2000). Moreover, co-speech gestures do not only contribute to *what* is understood, but also to *how* something is understood. When deictic gestures are used, listeners are more likely to correctly interpret utterances compared to when the utterance was not combined with a gesture, suggesting that co-speech gestures play a role in pragmatic understanding. For example, when hearing the utterance “it's getting hot in here,” people were sooner inclined to interpret this as an indirect request (i.e., could you please open the window) when the speaker pointed to the window, than when the speaker did not point, in which case the listener might interpret the utterance as a mere statement (Kelly et al., 1999). All in all, there is a great deal of evidence for the contention that gestures fulfill inter-cognitive (i.e., communicative) functions (Goldin-Meadow and Alibali, 2012).

### THE INTRA-COGNITIVE ROLE OF GESTURES

There is mounting evidence that gestures fulfill intra-cognitive functions in addition to inter-cognitive ones. This is relevant to our present purposes. For example, co-speech gestures affect speakers' own cognitive processes. Several studies have suggested that lexical access is disrupted or promoted when gesticulation is prohibited vs. allowed to naturally emerge. When speakers are prohibited from gesturing during speech with spatial content, they are less fluent than when gesticulation is allowed, suggesting that lexical access is disrupted (Rauscher et al., 1996; Morsella and Krauss, 2004; see, however, Hoetjes et al., 2014). Moreover, speech is more fluent when co-speech gestures are produced and gesture rates are higher when lexical access is difficult (e.g., during the tip of the tongue phenomenon; Chawla and Krauss, 1994). Furthermore, when gesticulation is prohibited, the content of speech is less likely to be spatial in nature, suggesting that gestures support speech that is spatial in content (Rimé et al., 1984). Not only can online speech be influenced by co-speech gestures, these gestures can also have an influence off-line. For example, making gestures during the recollection of

a previous event, can improve retrieval of details of that event compared to when gesticulation is not allowed (Stevanoni and Salmon, 2005). In addition, gesticulation prior to recalling previously learned words aids recall performance (De Nooijer et al., 2013).

Gestures primarily arise during the processing of visuospatial information (e.g., Alibali et al., 2001; Seyfeddinipur and Kita, 2001; Allen, 2003; Kita and Özyürek, 2003). For example, people are more likely to gesture when describing visual objects from memory as opposed to when the object is visually present (Wesp et al., 2001; Morsella and Krauss, 2004; see also Ping and Goldin-Meadow, 2010), although gesticulation also occurs when the object is present (Morsella and Krauss, 2004). Moreover, gestures occur more often when objects are difficult to describe in speech, such as complex, not easily describable drawings (Morsella and Krauss, 2004). Indeed, the emergence of gesticulation appears to be related to the cognitive demands of the task (Goldin-Meadow et al., 2001; Wagner et al., 2004; Ping and Goldin-Meadow, 2010; Cook et al., 2012; Marstaller and Burianová, 2013; Smithson and Nicoladis, 2014). For example, participants who were given the dual task of remembering letters while explaining a difficult math problem, remembered more letters when they were allowed to gesture while explaining the problem than when they were not allowed to gesture (Goldin-Meadow et al., 2001). This suggests that gesticulation reduced the working memory load imposed by explaining the math problem, leaving more capacity available for performing the secondary task of remembering letters. Gesticulation when describing a mental rotation problem emerges primarily when describing the task-relevant rotation itself as opposed to describing the task-relevant static end-point of the rotation (Hostetter et al., 2011). This finding suggests that it is the high spatial cognitive demand, which is arguably higher during dynamic spatio-temporal rotation as opposed to describing static spatial information, that invokes the use of gestures (see also Smithson and Nicoladis, 2014). Furthermore, it has been found that encouraging participants to gesture during a mental rotation task enhances their performance (Chu and Kita, 2011).

The findings described here primarily involved iconic gestures. However, even deictic (pointing) gestures occur more often when cognitive demand is higher. Infants and young children (between 1 and 2 years of age) sometimes point for non-communicative reasons (Bates et al., 1975; Delgado et al., 2009). Furthermore, pointing gestures can aid the regulation of the speaker's attention in non-communicative and challenging problem-solving situations (Delgado et al., 2011). In two studies, children ranging in age from 2 to 4 years old saw a toy being hidden in one of three containers on a rotation table. This was followed by a delay of 45–60 s during which the children either had to remember where the toy was hidden by the experimenter (cognitive demand group) or had to wait for the experimenter to retrieve the toy for them. During the delay the experimenter left the room. Additionally, the difficulty of the memory task was varied for half of the trials such that the table was turned for 540°. Analysis of the video-taped sessions showed not only that solitary pointing gestures occurred, but also that they occurred significantly more often in the cognitive demand condition than in the waiting condition (although no effects were found for task difficulty). A second experiment with children ranging

from 4 to 6 years old who performed a picture-matching task showed that constraining gestures resulted in poorer performance on the task than non-constraining gestures, but only for children who habitually pointed in the constrained condition, suggesting a cognitively beneficial role of solitary pointing gestures. This finding is surprising because deictic gestures have primarily been considered as serving communicative functions (Tomasello et al., 2007). Additional research on pointing gestures was conducted in the context of keeping track of counting. Children, adults, and even primates effectively use the hands in counting objects by pointing and touching gestures as to mark counted objects, and synchronize with counting expressed in speech (Boysen et al., 1995; Kirsh, 1995; Alibali and DiRusso, 1999). For example, participants who were allowed to use their hands for pointing during the counting of coins were faster and made fewer mistakes than those who were not allowed to use their hands (Kirsh, 1995). Thus, pointing gestures sometimes regulate visuo-spatial attentional processes, being especially helpful under high cognitive task demands.

These results converge with a recent correlational study that examined whether individual differences in spatial working memory capacity, spatial transformation ability, and conceptualization ability (amongst others) were associated with frequency of use of several types of gestures (Chu et al., 2013). Lower scores on all of these variables predicted higher frequency of spontaneously produced representational and conduit<sup>1</sup> gestures in a natural setting. Other evidence is consistent with this pattern. Particularly people with low working memory capacity are negatively impacted on a working memory task when they are not allowed to gesture as opposed to people with high working memory capacity (Marstaller and Burianová, 2013). Thus, in addition to the findings that gestures emerge during spatial information processing, gestures are also more likely to be produced by, and more likely to affect cognitive processes of, people with low spatial working memory and information processing ability (see also Chu and Kita, 2011).

Further evidence for gesturing as a compensatory mechanism comes from a study by Chu and Kita (2008). The type of spontaneous gestures that participants used during a mental rotation task followed a trajectory from external to more internalized solution strategies. That is, participants first gestured concretely as if manipulating the object to be rotated and subsequently changed their strategy and used their flat hand as stand-in for the object that needed to be rotated. Moreover, frequency of gesture use in aiding a spatial rotation task diminished over time, suggesting that cognitive operations became gradually internalized. A related phenomenon is that intermediate advanced abacus users use gestures during mental calculation. In the absence of the abacus, trained participants apply finger gestures as if manipulating an abacus ready to hand; but as abacus users become more advanced, they exhibit a reduced reliance on gestures during mental calculation (Hatano et al., 1977; Hatano and Osawa, 1983). In line with the findings of Chu and Kita (2008) this shows that the use of gestures becomes more infrequent as familiarity

<sup>1</sup>Defined as “iconic depictions of abstract concepts of meaning and language” (McNeill, 1985, p. 357).



with the task increases. Moreover, when describing the solution of a particular spatial problem, people's gesticulation aligns with the medium that the problem has been introduced in Cook and Tanenhaus (2009). For example, participants who described solutions of the Tower of Hanoi with physical disks as opposed to a computer simulation tended to spontaneously produce gestures that aligned with the physical actions performed with physical disks.

Thus, if we consider (a) that working memory capacity is limited, and (b) that new tasks often impose a higher working memory demand that diminishes as the learner becomes more experienced with a task (e.g., Chase and Ericsson, 1982; Kalyuga et al., 2003) then the findings we just reviewed suggest that gestures are likely to emerge in novel situations so as to provide the cognizer with some kind of external support. We will discuss the nature of this external support in our embedded/extended account of the cognitive function of gestures.

Finally, gestures can aid in acquiring a solution during problem solving (Alibali et al., 2004; Stephen et al., 2009; Boncoddio et al., 2010). For example, participants were presented with two glasses with differing widths and equal heights and were asked to imagine the glasses being filled with water to the same level. Participants judged whether the water would spill when glasses were rotated at equal angles (Schwartz and Black, 1999). Participants were able to predict the answer correctly much more often when rotating the empty glasses with their eyes closed, compared to when they were only allowed to think about the solution (i.e., mentally rotate). Although the previous study was in a sense a form of direct action (by allowing the objects to be manipulated), there is evidence that suggests that gestures, as non-direct manipulations, equally support the use of particular problem-solving strategies. For example, a study in which participants were presented with an interlocking gear problem (Alibali et al., 2004) found that they judged the direction of movement of a gear through different strategies, depending on whether or not gesticulation was allowed. When they were allowed to gesture, participants were more likely to simulate the rotations of each gear by finger gestures in order to provide the solution of the end-gear's rotational direction (depictive strategy), whereas participants who were prohibited from gesticulation were more likely to achieve the solution through the parity rule (direction gear  $x$  has the same direction as gear  $x + 2$ ). Note that the participants who used the depictive strategy were not better at the task than those using the parity rule (Alibali et al., 2004; also see Hegarty et al., 2005). Indeed, the parity rule strategy is generally considered to be the most effective strategy (Boncoddio et al., 2010). It is interesting in this regard to note that preschoolers are more likely to achieve understanding of the parity rule through gesticulation (Boncoddio et al., 2010). That is, preschoolers who used more gestures supporting a depictive strategy, more efficiently acquired a strategy based on the parity principle, in comparison to preschoolers who gestured less. Thus in this particular instance, the repeated use of gestures by participants is more likely to lead to discovery of new strategies during problem-solving although the use of gestures does not necessarily invite learners to adopt the most efficient strategy (see also Stephen et al., 2009).

The research reviewed here provides evidence that gestures have an intra-cognitive cognitive function for the gesturer. Furthermore, it produces two intriguing and related questions that we think need to be answered in a theoretical account of the cognitive function of gesticulation. First, why do gestures occur more often when cognitive demand is high? Second, why are spatial cognitive ability and working-memory capacity negatively related to the use of gestures?

## CURRENT THEORY ABOUT THE ORIGIN AND FUNCTION OF GESTURE

In this section, we will discuss several prominent accounts that aim to elucidate the underlying mechanisms and function of gestures, most prominently the Gesture-as-Simulated-Action account (GSA; Hostetter and Alibali, 2008) and subsequently the Lexical Gesture Process (LGP) model (Krauss et al., 2000), the Information Packaging Hypothesis (IPH; Kita, 2000), and the Image Maintenance Theory (IMT; Wesp et al., 2001). We evaluate these models directly after summarizing their main points, by assessing their explanatory power regarding the question: how do gestures-as-bodily-acts support cognitive processes?

We have chosen to address this collection of accounts for several reasons. The GSA account is a prominent contemporary account that attempts to integrate the literature of embodied cognition and the literature on gesture into a single perspective. Yet, as mentioned in the introduction, it seems that this attempt has resulted in a "dis-embodied" perspective on gesticulation. The other accounts have been very influential in elucidating the cognitive function of gestures. Moreover, they differ significantly from the GSA account but also from each other. The result is a representative (but not exhaustive) overview of theories about the possible cognitive function of gestures.

### GESTURE-AS-SIMULATED-ACTION ACCOUNT

The GSA account (Hostetter and Alibali, 2008) relies heavily on the insights from embodied cognition that representations are based on the sensorimotor system (Barsalou, 1999, 2008; Glenberg and Kaschak, 2002). This embodied view is supported by mounting evidence that perceptuo-motor faculties of the brain are activated during concrete but also supposedly symbolic and abstract conceptual processes (e.g., Barsalou, 2008; Pulvermüller et al., 2014). For example, merely reading words that have olfactory, gustatory, or motor connotations (e.g., garlic, jasmine, salt, sour, kick, pick) as opposed to reading neutral words, activates brain regions that are involved in smelling, tasting, and moving (Hauk et al., 2004; Gonzalez et al., 2006; Barrós-Loscertales et al., 2012).

The GSA approach predicts that cognitive processes, such as conceptual processing, co-occur with sensorimotor reactivations. More importantly it is contended that meaningful cognitive processing is *dependent* on these reactivations or simulations of sensorimotor states (Barsalou, 2008; Hostetter and Alibali, 2008). Indeed, conceptual processing is hampered when participants are primed with inconsistent perceptual or motor information (e.g., Glenberg et al., 2005; Kaschak et al., 2006). For example, participants are quicker in verifying the sensibility of sentences (such as "Andy delivered the pizza to you

vs. you delivered the pizza to Andy”) when their response actions were consistent with the implied motion of the sentences (moving the hand forward or backward), whereas they were slower when the movement contrasted with the implied motion (Glenberg and Kaschak, 2002). As such, it is suggested that induced sensorimotor states impinge on conceptual representational states since both systems are tightly coupled (Barsalou, 2008).

Hostetter and Alibali (2008) have suggested that the phenomenon of co-speech and co-thought gestures fits nicely with the idea that cognitive processing depends on activations in the sensorimotor system. In fact, according to the GSA account gestures *are* the bodily realizations (or as they call it, “visible embodiments”) of otherwise covert sensorimotor activations. The main question that the GSA account aims to address, therefore, is how sensorimotor activations come to be reflected in gestures. Hostetter and Alibali (2008, p. 503) first provide a simple answer: “Simulation involves premotor action states; this activation has the potential to spread to motor areas and to be realized as overt action. When this spreading activation occurs, a gesture is born.” More specifically, the GSA account suggests that gestures emerge through sensorimotor re-activations underlying thought and speech processing that “leak into” the motor-executive system:

“As an analogy, we might imagine activation spreading from premotor areas to motor areas through a gate. Once the gate is opened to allow more activation for one task (speaking), it may be difficult to inhibit other premotor activation (that which supports gestures) from also spreading through the gate to motor areas, the activation for the simulations ‘rides along’ and may be manifested as a gesture” (Hostetter and Alibali, 2008, p. 505).

Hostetter and Alibali (2008) further propose three underlying factors that determine when gestures are likely to occur. First, the strength of the particular perceptuo-motor activation must surpass a certain *gesture threshold* for actual physical embodiment (i.e., gesticulation) to arise. This activation strength is dependent on the degree to which speakers evoke visuospatial imagery during conceptual processing. For instance, they argue that the same conceptual content can be processed verbal-propositionally or with visuo-spatial imagery (e.g., in the case of route-descriptions), the latter type of encoding being more likely to evoke gesticulation (e.g., Alibali et al., 2001; Seyfeddinipur and Kita, 2001; Allen, 2003; Kita and Özyürek, 2003). Second, visuo-motor simulations are likely to evoke gesticulation when the conceptual content that is being processed involves an action. For example, talking about action is likely to evoke gestures because it is dependent on motor-information (Hostetter and Alibali, 2008). Third, it is speculated that the height of speakers’ gesture-threshold can vary across individuals and situations. To illustrate, a higher degree of neural interconnectivity between pre-motor and motor areas may lower the gesture threshold of a particular individual. Furthermore, inhibiting gesticulation requires cognitive effort and as such the threshold might be lowered when cognitive load is high (e.g., Goldin-Meadow et al., 2001).

### **Explanatory power of the GSA account**

So how does the GSA account answer our question of how gestures-as-bodily-acts support cognitive processes? First, it is

held that speech production and thought processes are dependent on the conceptual system recruiting sensorimotor representations. Furthermore, according to Hostetter and Alibali (2008), gestures arise from and are dependent on the strength of sensorimotor activations. However, the model does not allow the conclusion that gestures-as-bodily-acts aid cognition, because gestures only *execute* sensorimotor information, they do not *produce* it. The sensorimotor information that is produced (e.g., proprioceptive and visual consequences of movement) does not fulfill a cognitive function in the GSA account. This is indicated by the motor-leakage metaphor, as gestures simply “ride along” with sensorimotor activations (Hostetter and Alibali, 2008, p. 505) and can be understood as a mere “outgrowth” (Risko et al., 2013) or “visible embodiments” (Hostetter and Alibali, 2008) of internal embodied simulations. Thus, the GSA account leaves us with the question why do cognitive processes sometimes recruit the body (gestures), as opposed to relying on purely internal mechanisms? Furthermore, what is the explanatory power of the GSA account in terms of the empirical literature on the cognitive function of gestures provided above? Most notably, why is high cognitive demand result in more use of gestures. This is explained by the GSA account in “that inhibiting activation from spreading to a gesture requires more cognitive resources than does producing the gesture” (Hostetter and Alibali, 2008, p. 505). From this point of view, gesticulation is the default and is simply hard-wired with cognitive processes. By accepting this, we would simply deflate the idea of there being any function of gestures as bodily acts, endow the cognitive system with functionally unnecessary expenditure of energy (hand-movements), and allow only a negative cognitive effect of not gesturing. Although this idea of costly active inhibition may very well be a correct explanation for some instances of gesticulation, we think its possible scope for explaining the function of gesture is somewhat reduced by the realization that possessing a superfluous and energy-demanding gesture system does not seem very adaptive or flexible. Moreover, we think that a non-deflationary account of the function of gesture is possible and in fact more promising for understanding the empirical findings on the cognitive function of gestures reviewed in this paper.

### **LEXICAL GESTURE PROCESS MODEL**

The LGP model proposed by Krauss et al. (2000) tries to explain why speech might be facilitated by gesticulation. According to this theory, gestures do not only fulfill a communicative role, but may serve to facilitate lexical retrieval on the part of the gesturer as well. Gestures that share features with the lexical semantic content of the word will facilitate lexical access. Krauss et al. (2000) hypothesize that this is the case because gesturing results in “cross-modal priming” in which features of the concept represented by the gesture can facilitate lexical retrieval. According to this LGP account, gesture production draws upon the activated representations in working memory that are expressed in speech. The assumption is that the content of conceptual memory is encoded in multiple ways, and that activation of one representational format can spread to activation in another representational format. In this account, gestures derive from non-propositional representational formats (mostly visuo-spatial), as

opposed to speech, which draws on propositional symbolic formats. LGP further suggests that non-propositional information becomes expressed in speech through a spatial/dynamic feature selector that transforms spatially and dynamically formatted information into a set of “abstract properties of movement.” The abstract specifications are then translated into a motor program by a motor planner. Motor systems output the set of instructions from the motor planner and the gestural movement is monitored kinesthetically. The motoric features that are picked up by the kinesthetic monitor promote retrieval of the concept for speech through *cross-modal priming*. Krauss and Hadar (1999, p. 21) specify:

“The spatio-dynamic information the gesture encodes is fed via the kinesthetic monitor to the formulator, where it facilitates lexical retrieval. Facilitation is achieved through cross-modal priming, in which gesturally represented features of the concept in memory participate in lexical retrieval. Of course, it is possible to locate the site of gestural input more precisely (e.g., the grammatical encoder or the phonological encoder).”

### **Explanatory power Lexical Gesture Process model**

Does LGP allow for a cognitive role of gestures-as-bodily-acts? That is, does it answer the question why gestures are produced, and how they are cognitively relevant? An affirmative response is appropriate, although the mechanism seems underspecified and unparsimonious. Indeed, when a gesture is outputted by the motor-system, the “kinesthetic” feedback that is produced acts as input to the formulator (i.e., the grammatical or phonological encoder or both) and can then facilitate lexical selection by way of additional cues or “cross-modal priming.” Thus, in this model, motor-information is externalized and is fed back into the system to promote lexical retrieval through supporting the processes of the “grammatical encoder” and the “phonological encoder.” Yet the question remains why this motor-information needs to loop out of the brain and then be retrieved again by the kinesthetic monitor. According to LGP, gesture will only facilitate lexical access when the gesture features match the lexical semantic content of the concept. Therefore, gestures will only facilitate lexical access when the kinesthetic information that was already present in a verbal form is fed back into the formulator. Thus it seems that the brain is “primed” with information that is already present in the internal system, given that gestures are outputs of an already constructed motor program. Thus, it is unclear with what kind of information the cognitive system is primed. Of course, gestures might indeed fulfill this function, but the model currently presented is not very illuminating why and how gestures-as-bodily-acts fulfill a cognitive function. So, although LGP also suggests an intra-cognitive role for gestures, it is still difficult to appreciate the added value of the kinesthetic information that is fed back into the system with regard to cognitive processing.

### **INFORMATION PACKAGING HYPOTHESIS**

A third prominent theory in the gesture literature is the IPH (Kita, 2000). This theory proposes that gestures aid speech production by breaking images into smaller bits to enhance the verbalize-ability of communicative content. A key idea is that

there are two modes of thinking that tend to converge during the linguistic act. There is analytical thinking as opposed to spatio-motoric thinking from which gestures follow, which involves the organization of information through hierarchical structuring and involves decontextualized conceptual templates. According to Kita, these templates can be non-linguistic (in the case of scripts), or linguistic, such as in the case of a lexical item’s semantic and pragmatic specifications. The templates are not multimodal as in the case of the GSA account, thus they do not involve “activation of ‘peripheral’ modules” (Kita, 2000, p. 164), yet can be translated into the other mode of thinking, which is spatio-motoric thinking. The spatio-motoric mode of thinking constitutes gestures and involves information organized in action schemas. Gestures should be considered as actions in a virtual environment, and are derived from practical actions.

A core idea behind IPH is that the two modes of thinking collaboratively organize information during speaking. Kita (2000, p. 163) suggests that (a) “The production of the representational gesture helps speakers organize rich spatiotemporal information”, (b) “Spatio-motoric thinking, which underlies representational gestures helps speaking by providing an alternative informational organization that is not readily accessible to analytic thinking” and (c) “Spatio-motoric thinking and analytic thinking have ready access to different sets of informational organizations. However, in the course of speech production, the representations in the two modes of thinking are coordinated and tend to converge.”

### **Explanatory power Information Packaging Hypothesis**

Does IPH have explanatory power of how gestures-as-bodily-acts support cognitive processes? The IPH does not provide a clear account of how gestures aid the “packaging of information” given that gestures are considered as the result of spatio-motoric thinking that is already internally realized. That is, just like the GSA, the IPH seems to regard gestures as mere output of spatio-motoric thinking, with the latter having the actual cognitive function (information packaging). Even if we allow for a possible different reading of the IPH, in which gesticulation actually supports spatio-motoric thinking, the IPH account does not go into any detail about how gestures-as-bodily-acts feedback to or support internal cognitive processes to perform the function of spatio-motoric information packaging.

### **IMAGE MAINTENANCE THEORY**

The final theory under review here is the IMT by Wesp et al. (2001). Although this theory is only briefly presented in an empirical paper it has become an influential view on the cognitive role of gestures (Alibali, 2005). Arguably, the main thesis of the IMT, which is often contrasted with the LGP, is “that gestures are not directly involved in the search for words; rather, they keep the non-lexical concept in memory during the lexical search, a process of data maintenance not unlike that needed in other problem-solving activities” (Wesp et al., 2001, p. 592). This is further explained; “a prelinguistic representation of spatial information is established through spatial imagery and maintenance of these spatial images is facilitated by gestures” (Wesp et al., 2001, p. 595). Wesp et al. (2001) base this idea on the idea that spatial information is held in the



visuospatial scratchpad of working memory (Baddeley, 2003). The items (visuospatial information) in the scratchpad decay rapidly and must be rehearsed to be maintained in working memory. Just like articulatory loops, gestures serve the function of “refreshing” the visual scratchpad to sustain activation of the image in working memory. Importantly, gestures are therefore not necessary for lexical retrieval but may indirectly facilitate it through, “motoric refreshing” of the image (p. 597).

### **Explanatory power Image Maintenance Theory**

Does the IMT have explanatory power of how gestures-as-bodily-acts, support cognitive processes? The answer is yes, although much is still needed to understand its function. “Yes” because the IMT suggests that the production of a *physical* gesture supports the maintenance of an internal spatial image (a cognitive process); without the physical gesture the internal spatial image becomes unstable and its activation is likely to decay. Yet, Wesp et al.’s (2001) account does not provide sufficient detail beyond this notion. How do gestures refresh motoric spatial images? What is the mechanism by which gestures-as-bodily-acts refresh motor spatial images? Furthermore, are not gestures redundant given that they provide the gesturer with information that is already present in the system that outputs the gestures (e.g., visual information)? Although these questions remain unanswered, of all the accounts presented here, the IMT is most compatible with an embedded/extended account that assumes gestures are cognitively relevant because they are bodily.

### **SUMMARY OF FINDINGS FROM THE THEORETICAL OVERVIEW**

In the previous subsections, we have discussed four models that have been put forth to explain the underlying mechanisms of gestures. We sought an answer to our question: how do gestures-as-bodily-acts support cognitive processes? Our review of the literature suggests that the cognitive function of gestures-as-bodily-acts cannot be adequately explained, or remains under-specified, in several different theories about the underpinnings and functions of gestures. In the GSA account gestures are seen as by-products of sensorimotor activation but cease to be supporting cognition the moment they are outputted by the motor-system. The IPH suggests that gestures help package the spatio-motoric thinking during speech, yet this account also assumes that gestures are the result of these processes as they are the realizations of spatio-motoric internal processes; they are pre-packaged the moment they are externalized as gestures and do no packaging of their own. In the LGP account, the gestures that are produced are fed back into the cognitive system to provide it with cross-modal primes. As such, gestures, as physical acts, attain a function. Yet, the LGP account is unclear about what exactly is primed, or what novel information gestures provide to the system, that was not already activated or present. Interestingly, the IMT does seem to ascribe a definite cognitive function to gestures by positing that they support the maintenance of mental images.

It is important to stress that our review is aimed at answering a specific question that may be different from the questions that the theories we discussed were designed to address. We have only considered these theories’ explanations (*explanantia*) of a particular aspect of gesticulation that we think needs to be explained

(*explanandum*), namely how gestures-as-bodily-actions have a cognitive function. This means that we do not suggest that the theories under discussion are wrong, nor do we suggest that they are incompatible with the upcoming perspective; rather the *explanantia* they offer are not (yet) suitable to cover the *explanandum* that is the focus of the current paper. In the next section, we aim to fill this explanatory gap through a more embedded/extended perspective on the cognitive function on gestures.

## **TOWARD A MORE EMBEDDED/EXTENDED PERSPECTIVE TO THE COGNITIVE FUNCTION OF GESTURES**

In this section, we attempt to answer the main question of how gestures can fulfill cognitive functions. In the following subsection, we will briefly introduce the embedded/extended cognition perspective (inspired by Clark, 2013), which is followed by a representative overview of research in this domain. Subsequently we apply the relevant theoretical and empirical findings to the cognitive function of gestures, which yields challenges and hypotheses for future research.

### **AN EMBEDDED/EXTENDED PERSPECTIVE: THEORY AND RESEARCH**

Embedded/extended cognition is considered part of the broader development of embodied cognitive science (Wilson, 2002; Shapiro, 2010) and has its roots (amongst others; Gallagher, 2009) in situated cognition (Bredo, 1994), robotics (Brooks, 1991) and the dynamical systems approach to cognition (Chemero, 2009). According to a loose description of “the” embedded/extended perspective on cognition (cf. Wilson, 2002), the main thesis is that the cognitive system is a coupled brain–body–world system (Wheeler, 2007; Clark, 2008). As such, cognition involves an ongoing transaction between current states of the brain, body, and the environment (Clark, 2008). Within this view, the classic internalist picture of cognition is disputed; thinking is something we do, rather than something that simply happens within us. Understanding cognition, therefore, requires a broader level of analysis that allows the study of how we use our body and the world during the unfolding of cognitive processes. For example, Hutchins (1995b) analyzed the goings-on of commercial airlines and suggested that a purely internalist perspective was ill-suited to understand its workings; flying a plane involves task-relevant information that is neither fully instantiated in the cockpit, the pilot, or co-pilots, it is rather distributed among them and all parts work together (see also Hutchins, 1995a). Everyday examples of embedded/extended cognitive phenomena would be, for instance, asking another person to remind you of something, using a tall building for navigating your way home, or reducing working memory load by taking notes during a conversation. Or in the case of drawing: “One draws, responds to what one has drawn, draws more, and so on. The goals for the drawing change as the drawing evolves and different effects become possible, making the whole development a mutual affair rather than a matter of one-way determinism” (Bredo, 1994, p. 28).

In philosophy, there is a debate on whether states of the body and the environment can be considered extra-neural contributors to cognition (Wilson, 2002), or in a more radical reading, external vehicles of cognition (Clark and Chalmers, 1998; Clark, 2008). According to the radical extended perspective, the internalist view



is provoked by the classic thesis that “If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process” (Clark and Chalmers, 1998, p. 8). The less radical thesis, the notion of embeddedness, also stresses a tight coupling between the agent and the world and suggests that the body and environment can, often in unexpected ways, causally impact cognition, yet suggest that the body and the environment are not part of cognition (Adams and Aizawa, 2001; Rupert, 2009). Thus the difference between embedded and extended cognition is whether extra-neural conditions causally impact cognition (embedded thesis) or are constitutive of it (extended thesis). As mentioned in the introduction, we will side-step this technical debate; for our present purposes it suffices to say that we follow the joint anti-internalist approach of embedded and extended cognition, which suggests that the cognitive system works in concert with the body and the environment.

The embedded/extended perspective has given rise to a large amount of empirical research on the way the cognitive system uses the body and the environment (e.g., Kirsh and Maglio, 1994; Ballard et al., 1995; Haselen et al., 2000; Martin and Schwartz, 2005; Fu, 2011; Risko et al., 2013; see also Pouw et al., 2014). A seminal study by Kirsh and Maglio (1994; see also Stull et al., 2012) found that expert Tetris players make more use of *epistemic actions*; actions that uncover (hidden) information that is cognitively demanding to compute. These types of actions are different from actions that bring one closer to one’s goal (pragmatic actions). For example, advanced players, instead of rotating “zoids” (i.e., falling block arrangements in Tetris) through mental simulation to judge whether it will fit the zoids in the bottom deck, they preferred rotating them physically as this allowed a direct matching of orientation and fit. The cognitive operation of rotation to determine a possible fit was thus off-loaded onto the environment.

Another classic study (Ballard et al., 1995, 1997; Haselen et al., 2000) showed that the cognitive system opts for retrieving information just-in-time, thereby minimizing constraints on working-memory. Participants were asked to recreate a configuration of colored blocks from a *model* by picking up colored blocks from a *resource space* and putting them in a *work-space*. The model, resource-, and work-space were all displayed in front of the participants. Eye-movement data were collected during this task. Participants made many switches of eye fixations between the model, work and -resource space. This indicated that participants adopt a “minimal memory strategy” in which information is gathered incrementally as opposed to memorized in one fell swoop. Instead of memorizing the position and color all at once, participants first memorized the color to be searched from the model, then after finding a color match in the resource space, looked up the position of the block of the model. Thus, information is gathered just in time to minimize working memory constraints (see also Cary and Carlson, 1999, who obtained similar results in an income calculation task).

Yet, findings indicate that the cognitive system does not seem to have an *a priori* preference for using the environment rather than internal cognitive resources in solving a cognitive problem; which

strategy is adopted depends on the context. For example, when Ballard et al. (1995) increased the distance between the workplace and the model, participants were more likely to adopt a memory-intensive strategy. This finding resonates with the study by Gray and Fu (2004; see also Fu, 2011) in which participants were confronted with the task of programing a simulated VCR. In this task, retrieval costs of attaining task-relevant information were subtly manipulated. That is, the ease of retrieval was manipulated in such a way that participants could either acquire the information through a simple glimpse or through performing an additional mouse-click to make the information available. The cognitive strategy that the subjects chose changed as a function of the ease of retrievability. When external information was directly accessible, participants primarily relied on retrieving information externally. Attaining this “perfect-knowledge-in-the-world” was shown to be a reliable strategy, as it reduces the number of mistakes made during the task. Moreover, when the information was only indirectly available, participants were more likely to rely on internal memory, which produced a larger number of mistakes. The reason why participants in this condition relied on “imperfect-knowledge-in-the-head” was that the internally stored information was more quickly available compared to externally available information, as was predicted by a computational model that expressed the amount of time it takes to retrieve or recall information. Thus people seem to opt for the quickest problem-solving strategy in which the cognitive system “tends to recruit, on the spot, whatever mix of problem-solving resources will yield an acceptable result with a minimum of effort” (Clark, 2008, p. 13).

Situational constraints bring about a trade-off decision whether the cognitive system relies on computation performed “on-line” (with the environment) or “off-line” (internally; Wilson, 2002). Relevant in this regard is a recent set of experiments conducted by Risko et al. (2013) in which participants were presented with a varying number of letters that were either presented upright or tilted at 45° or 90°. Participants spontaneously rotated their head, which indeed seemed to promote readability of tilted presentation of letters. Furthermore, participants were more likely to rotate their head when more letters were presented and tilt of the letters was more extreme, indicating that head-tilting (which they call external normalization) occurs when the cognitive demand of not tilting the head by means of “internal normalization” increases (more cognitive effort to read more letters in tilted position, and more extreme tilt of the letters). Thus, when internal computational demand increases, an externally mediated cognitive strategy becomes more attractive. This was also found in a study by Kirsh (2009), in which participants played a mental tic-tac-toe game with the experimenter. During the mental tic-tac-toe game participants have to keep their own “moves” and those of the opponent, in mind. In the critical conditions, participants were given a sheet of paper with a tic-tac-toe matrix depicted on it or a blank sheet. External support of a tic-tac-toe matrix aided participants’ efficiency of playing the game in comparison to having no support or a white sheet. Apparently, participants are able to project the progression of the moves on the matrix through visual simulation. This is very similar to chess-players who think through moves on a chess-board without manipulating the board (Kirsh, 2009). Interestingly,

however, the external support was only beneficial when the tic-tac-toe game was complex ( $4 \times 4$  matrix as opposed to a  $3 \times 3$  matrix), and especially for participants who scored low on spatial ability. Thus, this study suggests that projection on external support is especially helpful when cognitive demand is high, and relatedly, primarily for those who are low in spatial cognitive ability.

As a final example, the study conducted by Martin and Schwartz (2005) shows how active manipulation of the environment may foster learning through exploration of the solution space. In two studies, children (9–10 years old) were learning how to solve fraction operator problems (e.g., one-fourth of eight candies), using physical tiles and pie-wedges that were movable *and* in another set of trials, using line drawings of pies or tiles which they could highlight and circle with a pen. The difficulty that children often experience in this task is that they focus on the numerator, leading them to understand “one-fourth of eight candies” to be “one candy.” Martin and Schwartz (2005) predicted that physical interaction with manipulable objects would increase the chance that children come to interpret that one-fourth of eight means four groups of two because rearranging the tiles results in new groupings. Thus they reasoned that the agent and the environment mutually adapt each other (as in the case of drawing), where one acts without a preconceived goal on the environment which in turn feeds back information that might align with the correct solution. Indeed, children performed better with manipulable objects than without them (Experiments 1 and 2). Interestingly, presenting the children with the correct organization of tiles did not aid understanding; rather the physical open-ended interaction with the environment drove understanding and performance on the task (see also Manches et al., 2010).

Let us summarize. First, the cognitive system makes use of the environment to distribute computational load but also to enable exploration of a problem-space that is difficult to achieve off-line (i.e., to achieve through purely internal computations). Moreover, the cognitive system is not *a priori* driven to reduce internal computational load by off-loading onto the environment, rather the environment is exploited if it offers a cheaper resource than internal means of computation to achieve an acceptable performance on a task (Gray and Fu, 2004). Although not conclusive, it further seems that when cognitive demand is high, either due to external constraints (higher cognitive load of the task) or internal constraints (e.g., low visuospatial cognitive ability) the cognitive system is more likely to opt for and benefit from external computational strategies. However, these findings do not allow us to draw definitive conclusions about when and how the cognitive system trades external with internal computational resources. Thus one of the major challenges for research in embedded/extended cognition is to determine which external (e.g., availability of external information) and internal (e.g., working memory ability) constraints affect whether and how problem-solving strategies become externally or internally mediated (Risko et al., 2013). Furthermore, is it possible to identify a trajectory of problem-solving strategies as expertise develops? Specifically, does the cognitive system first rely on external support – given that it is still ill-equipped to perform stand-alone internal computations – and are computations increasingly performed off-line

when the cognitive system becomes more equipped (e.g., because of acquired strategy knowledge or chunking mechanisms) to hold task-relevant information internally?

Even though such questions cannot yet be answered by the embedded/extended cognition frameworks, it is not difficult to see the relevance of this framework for gesture research; there is a clear analogy between these findings and the findings from some of the gesture studies reviewed in the section on “the intra-cognitive role of gestures.”

## AN EMBEDDED/EXTENDED PERSPECTIVE ON THE COGNITIVE FUNCTION OF GESTURES

Recently, Clark (2008, 2013; see also Wheeler, 2013) provided a purely extended perspective on gesticulation. Clark (2013) provides a detailed discussion of why gestures should be seen as constitutive to – as opposed to merely causally impinging on – cognitive processes (cf. Wheeler, 2013). Here we only briefly address his account to further develop an embedded/extended perspective that is able to provide an explanation of the empirical data on the cognitive function of gestures as well as produce hypotheses and identify challenges for further research.

According to Clark (2013) we should *not* understand the cognitive role of gestures purely in terms of its neural pre- and post-cursors:

“The wrong image here is that of a central reasoning engine that merely uses gesture to clothe or materialize performed ideas. Instead, gesture and overt or covert speech emerge as interacting parts of a distributed cognitive engine, participating in cognitively potent self-stimulating loops whose activity is as much an *aspect* of our thinking as its *result*.” (p. 263)

Furthermore, he states that:

“The physical act of gesturing is part and parcel of a coupled neural-bodily unfolding that is itself usefully seen as an extended process of thought.” (p. 257)

Clark further argues that by producing a gesture, something concrete is brought into being (arm posture) that subsequently affects ongoing thinking and reasoning. Much like using a notepad, gestures provide a *stable* physical presence that embodies a particular aspect of a cognitive task. We can appreciate Clark’s point if we consider that speech dissolves in midair and working memory allows only for a certain amount of thoughts to be consciously entertained. We can argue that gestures are not only a way to externalize speech and thought content, but also allow for temporal cognitive stability that might be more reliable than internal means of temporal cognitive extension (e.g., consciously attending to a thought to keep in mind).

Thus the key to an embedded/extended perspective on gestures is the view that gestures fulfill a cognitive function *because* they are bodily. That is, in contrast to what the GSA and the IPH propose, gesticulation produces an external physical presence that somehow supports internal cognitive processes. According to Clark’s (2013) purely extended account, this physical presence instantiated in gesture is actually part of thinking itself. Indeed, he thinks that a more moderate account of gestures’ function in which they merely *affect* inner neural cognitive processes is misconstrued. His argument for an extended cognitive understanding of gestures relies on the

appreciation that some crucial forms of neural activity arise in coordination with gestures, wherein gesture and neural activity are interdependent in achieving a particular cognitive state. Thus although, in some instances “neural goings-on’ may be sufficient for the presence of some cognitive state or the other” in other instances gestures, at times, should be given a genuine cognitive status (p. 261) because “gesture and speech emerge as interacting parts of a cognitive system” (p. 263) whereby no meaningful categorization can be made of what should be considered cognitive or non-cognitive on the basis of the distinction between inner (neural activity) and outer (gestures).

How and when do these specific physical conditions fulfill a supporting role for a particular cognitive function? It is instructive to compare the research from the embedded/extended cognition tradition with research on the cognitive function of gesture. We need to reconsider the research by Kirsh and Maglio (1994), which showed that expert Tetris players operate on the environment to alleviate internal computational load (epistemic actions). Determining where a zoid fits is not dependent on internally computed rotations of the zoid, but is achieved by actual rotation of the zoid. In mental rotation tasks in which participants have to judge whether a 3-d zoid matches one out of several 3-d zoids depicted in different rotational angles (classic S-M cube task; Shepard and Metzler, 1971), participants use gestures to aid in their judgments (Chu and Kita, 2008, 2011). We would submit, that gestures in this case *are* epistemic actions that reveal information that is hidden (since the 3-d zoids do not rotate by themselves) and difficult or more costly to compute internally. Chu and Kita (2008) also found that when participants first approach the mental rotation task they are more likely to use hand-movements *as-if* actively rotating the block. We would speculate that in this case gestures fulfill the function of providing a physical platform that supports the internal representational stability (a term earlier used by Hutchins, 2005) of a rotating 3-d zoid (see also Pouw et al., 2014). In this case the zoid is visually “projected” into the hands (Kirsh, 2009) and is manipulated *as if* it were actually in the hand. In this case the hands offer a reliable external support for performing the cognitive function of rotating the projected 3-d zoid through gestures. Furthermore, using pointing gestures to keep track of something in the environment similarly produces a reliable physical attentional marker that alleviates internal attentional tracking processes (e.g., Kirsh, 1995; Delgado et al., 2011). This might also be the case with abacus users doing mental calculations that perform gestures on, what seems to be, a mentally projected abacus (Hatano et al., 1977; Hatano and Osawa, 1983). In this case, physical gesticulation seems to be preferred by these users as opposed to internally simulating changes on the abacus. We would argue that because gestures allow a stable external physical presence, they support internal representational stability of the dynamically changing abacus during calculation. In line with Kirsh (2009), we argue that in these cases the cognitive system seems to be neither purely off-line nor on-line; rather, it uses partly environmental resources (e.g., gestures) and internal cognitive resources (e.g., visual simulation) to perform a task. Gestures are essentially a way to put on-line extra-neural resources into the mix of problem-solving resources.

Another possible embedded/extended function of gesture is exploration of a problem space. Martin and Schwartz (2005) found

that manipulation of objects promoted the understanding of fraction-operating principles. Relevantly, gesturing might sometimes allow the gesturer to become aware of structural correlations that would be difficult to generate through internal computation. For instance, this seemed to be the case in the rotating-gear problem, in which the number gestures used that simulated each rotation of a gear predicted the discovery of a more efficient problem-solving strategy that involved pick-up of the regularity that each gear  $N + 2$  rotates in the same direction (Delgado et al., 2011).

With regard to when gestures emerge to fulfill an embedded/extended function, the research that we have discussed in the domain of embedded/extended cognition has another interesting alignment with the gesture literature. We can summarize both streams of findings in one converging main principle: *When the costs of internal computation are high, either induced by external constraints (higher cognitive demand of the task; more cost of retrieving information from the environment) or internal constraints (e.g., lower working memory ability) the cognitive system is more likely to adopt, if cheaply available, an externally supported problem-solving strategy; be it the environment or gestures* (Goldin-Meadow et al., 2001; Gray and Fu, 2004; Wagner et al., 2004; Kirsh, 2009; Ping and Goldin-Meadow, 2010; Marstaller and Burianová, 2013; Risko et al., 2013; Smithson and Nicoladis, 2014). In other words, “cognitive processes flow to wherever it is cheaper to perform them” (Kirsh, 2010, p. 442). Understood in this manner, it is not surprising that people who are describing a physical object tend to gesture less when the object is present as opposed to absent (Morsella and Krauss, 2004), since the task-relevant information is cheaply available in the environment. Or that gestures are more likely to be used to lighten the cognitive load when pressure is put on internal computational system (cognitive demand of the task; e.g., Goldin-Meadow et al., 2001; Smithson and Nicoladis, 2014).

This embedded/extended perspective on the cognitive function of gestures, leads to several testable questions and further challenges for future research.

First, an interesting avenue for further research is to determine how changes in the external constraints – such as the cognitive demands of a task – and in the ease of availability of external resources, changes the likelihood of gesturing. For example, one could devise a mental rotation task in which participants can rotate a 3-d zoid either through a mouse, by using gestures, or solely by internal strategies. According to the present perspective, if we manipulate the speed in which the 3-d zoid can be manipulated by a mouse, we would predict that participants are more likely to use gestures when the manipulation takes more time (as relative cost decreases). Another, more unorthodox manipulation would be to put varying weights on the wrists of participants, which may induce costs in terms of energy expense, leading participants to an earlier adoption of an internal solution strategy. Many more constraints could be considered to assess the trade-off decision between internal and external resources that the cognitive system seems to make.

Second, gesture use evolves (Chu and Kita, 2008). When the task is more familiar, hand-gestures evolve from “*as-if* manipulations” to a stand-in-for relation of the 3-d zoid by means of a rotating flat hand, eventually eliminating the use of gestures



altogether. In a similar vein, when abacus users become more advanced they tend to use less and less gestures during mental calculations. Indeed, it seems that gestures itself are costly to perform, and contrary to the GSA account, may under certain circumstances hinder performance (De Nooijer et al., in press), or learning (Post et al., 2013) relative to other strategies. Interesting in this regard, is research that suggests that different types of body-movements have their own cognitive load (or come with particular cognitive costs) and may at times be traded for less costly bodily movements. That is dancers who rehearsed a dance-routine performed better when they rehearsed through “marking” (minimal movements and use of gestures to stand in for full-out movements) as opposed to rehearsing the routine full out (Warburton et al., 2013). Thus, it seems that under certain conditions, gestures, once cheap resources to think with, become relatively costly in comparison to, and are therefore traded in for, purely internal strategies. This raises several questions. For example, do gestures help in the internalization process? Thus, are embedded/extended solution strategies shaping the way internal computations are performed?

Relatedly, when the cognitive system has a lower ability to produce internal object rotations (i.e., low spatial cognitive ability) it will rely more on external resources such as gestures (e.g., Chu et al., 2013; Marstaller and Burianová, 2013). An important research question that relates to this idea is whether people who score “low” on spatial cognitive ability test are actually only scoring low on *mental* spatial cognitive ability, and may not underperform when gestures are allowed. Indeed, when gesture is prohibited people who are low in working memory perform only more poorly on a mental rotation task with no performance deficits in the gesture condition, suggesting that they can fully compensate with external problem-solving strategies (Marstaller and Burianová, 2013). Furthermore, consider findings that prohibiting gesturing has a negative effect on performance. Seen in this light, this negative effect of not gesturing may not arise because it imposes cognitive load, and thereby imposes constraints on cognition (as proposed by the GSA account), but precisely because the prohibition to gesture withholds the cognitive system from the use of external resources in the performance of a task. Thus, whereas the GSA account suggests that not-gesturing imposes a cognitive load since the agent has to prevent automatic activations of gestures, we propose that the prohibition of gesturing takes external bodily resources away from the agent and drives the agent to rely exclusively on internal computational processes. This is an important empirical question that future research should address, as it is both related to how we should define and measure cognitive abilities, as well as to the particular cognitive function of gestures.

A more fundamental question that currently remains unanswered in the embedded/extended perspective on gesturing is what type of information is being made available through gesturing. Is it the proprioceptive, kinesthetic, haptic, and/or visual consequences of movement that allow gestures to support cognitive processes? Or both, as these systems are tightly coupled (e.g., Radman, 2013)? For example, it is well-known that the visually impaired people use gestures (Iverson, 1998). Do they still benefit from gestures through proprioception or other consequences of movement? Clark (2013) raised a similar question

in relation to patients with a rare disease that leads to loss of proprioception; yet these patients are still able to gesture quite naturally (see Gallagher, 2005). Would gestures still fulfill an embedded/extended cognitive function for such patients through visual feedback? This question is somewhat harder to address since the disease is, luckily, quite rare. An interesting avenue for research therefore would be to interfere with the information that gestures might provide as to identify factors that might underlie the embedded/extended cognitive function of gestures. For example, obstructing visibility of one’s own gestures, by putting a screen at the level of the shoulders (Gallagher, 2005). Thus the current challenge for the present account is to provide an account of what information gestures produce that might be supportive for cognitive processes.

## CONCLUSION

By means of our review of the empirical literature we have tried to assess explanatory power of current theories with regard to the question of how gestures might fulfill cognitive functions. Although all the accounts we have addressed here claim that gestures indeed fulfill a cognitive function, we have shown that in these accounts, this claim often does not refer to gestures, but rather to their neural precursors. Importantly, there are accounts that suggest that gestures fulfill the cognitive role of priming or activating internal action representations (e.g., Krauss et al., 2000; Goldin-Meadow and Beilock, 2010), yet we think the reason why bodily movements fulfill this function is not clearly stated and seems to differ from the embedded/extended cognitive function we have identified here. We have tried to analyze the cognitive functions of gestures, by integrating the literature of embedded/extended cognition with the gesture literature. There is a considerable amount of overlap between the ways cognizers have been found to use their environment as well as how gestures support cognitive processes. Although further research into the exact mechanisms of embedded/extended functions of gestures is necessary, we put forth the notion that gestures provide the cognitive system with a stable external, physical, and visual presence that can provide a means to think with.

Importantly, we should stress two related concerns that apply to the current proposal. First, it is evident that the embedded/extended view on gestures, as presented here, does not address the full gamut of gesticulation. We have primarily focused on co-thought gestures in problem-solving contexts instead of, for example, beat gestures, or gestures that primarily emerge in communicative contexts. Therefore, at this point we remain agnostic to whether all gestures fulfill an embedded/extended cognitive function (for the gesturer). Indeed, extant “alternative” theories that we have addressed here may very well be complementary to our proposal. These theories are complementary to our proposal in that they might address cognitive functions and underpinnings of gestures that we have not addressed here. For example, it is possible that gestures emerge from action-related motor simulations that are activated during visuo-spatial cognition (Hostetter and Alibali, 2008) with the added proposal that the bodily externalizations of these motor simulations have a cognitive function themselves of the kind we have



proposed here. Thus although we maintain that current theories in the gesture literature are not very suitable to address why gestures-as-bodily-acts might fulfill a cognitive function, our proposal does not deny any explanatory power of these theories regarding other aspects of the nature and cognitive function of gestures.

Secondly, it is clear that gestures have a developmental trajectory and primarily emerge in intersubjective contexts (e.g., McNeill, 1992; Iverson and Thelen, 1999; Tomasello, 2008; Liszkowski et al., 2012). As such, the current embedded/extended account of the cognitive function of gestures is still presented in an “ontogenetic vacuum” and is still rather individualistic. Although this is a concern that needs to be addressed in future work, there is much room for exploring how the embedded/extended function of gestures might be related to developmental and social dimensions. For example, Iverson and Thelen (1999) have provided a detailed account of how the hands, mouth, and the brain should be regarded as one dynamical system; more specifically of how these components become entrained throughout development. Although they focus primarily on the way language and gesture become constitutively interdependent, the kind of gestures that have been the focus of this paper (gestures in problem-solving contexts) can be scaffolded onto their developmental account as another way of how “perception, action, and cognition can be mutually and flexibly coupled” (Iverson and Thelen, 1999, p. 37). On the other hand, how does our account relate to the intersubjective context in which gestures most often emerge? It would fare well with appeals coming from embodied cognitive science which suggest that an important way humans achieve interpersonal understanding is not from a spectatorial third-person stance, but rather from an interactive and second-person stance (e.g., De Jaegher and Di Paolo, 2007; De Jaegher et al., 2010; Anderson et al., 2012; Schilbach et al., 2013; Pouw et al., under review). In these approaches interpersonal understanding involves “know-how that allows us to sustain interactions, form relations, understand each other, and act together” (De Jaegher et al., 2010, p. 442), instead of two brains trying to predict each other’s mental contents through observation alone. In such a portrayal of intersubjectivity, gestures are always already considered as having an embedded function for both the gesturer and the interlocutor since gestures are co-constitutive of the social coordination itself. To put it another way, in social interaction gestures are a non-neural component that is part of an organism–organism–environment coordinative structure (Anderson et al., 2012). The challenge for further work is to show how non-social embedded/extended gestures that we have focused on here might develop from these social contexts.

In closing, our aim with this article to point out the necessity of understanding the role of the body in thinking. We tried to accomplish this by developing an embedded/extended perspective on the cognitive role of gestures. In this perspective, the body is not a trivial output-appendage of the cognitive system but an important component thereof. The body is a resource with particular qualities that is recruited in the coordination of cognitive processes. This perspective intended to promote research that tries to further address when, why, and how gestures are recruited during cognitive processes.

## AUTHOR CONTRIBUTIONS

Wim T. J. L. Pouw drafted, Jacqueline A. de Nooijer co-drafted, Tamara van Gog, Rolf A. Zwaan, and Fred Paas provided critical revision of the manuscript. All authors approved the final manuscript.

## ACKNOWLEDGMENTS

This research was funded by the Netherlands Organisation for Scientific Research (NWO-PROO, project number: 411-10-908) and the National Initiative Brain and Cognition (project number: 056-33-016).

## REFERENCES

- Adams, F., and Aizawa, K. (2001). The bounds of cognition. *Philos. Psychol.* 14, 43–64. doi: 10.1080/09515080120033571
- Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spat. Cogn. Comput.* 5, 307–331. doi: 10.1207/s15427633scc0504\_2
- Alibali, M. W., and DiRusso, A. A. (1999). The function of gesture in learning to count: more than keeping track. *Cogn. Dev.* 14, 37–56. doi: 10.1016/S0885-2014(99)80017-3
- Alibali, M. W., Heath, D. C., and Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *J. Mem. Lang.* 44, 169–188. doi: 10.1006/jmla.2000.2752
- Alibali, M. W., Spencer, R. C., and Kita, S. (2004). “Spontaneous gestures influence strategy choice in problem solving,” in *Paper Presented at the Biennial Meeting of the Society for Research in Child Development*, Atlanta, GA.
- Allen, G. L. (2003). Gestures accompanying verbal route directions: do they point to a new avenue for examining spatial representations? *Spat. Cogn. Comput.* 3, 259–268. doi: 10.1207/s15427633scc0304\_1
- Anderson, M. L., Richardson, M. J., and Chemero, A. (2012). Eroding the boundaries of cognition: implications of embodiment. *Top. Cogn. Sci.* 4, 1–14. doi: 10.1111/j.1756-8765.2012.01211.x
- Arbib, M. A., Gasser, B., and Barrès, V. (2014). Language is handy but is it embodied? *Neuropsychologia* 55, 57–75. doi: 10.1016/j.neuropsychologia.2013.11.004
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201
- Ballard, D. H., Hayhoe, M. M., and Pelz, J. B. (1995). Memory representations in natural tasks. *J. Cogn. Neurosci.* 7, 66–80. doi: 10.1162/jocn.1995.7.1.66
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., and Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behav. Brain Sci.* 20, 723–742.
- Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J. C., Costumero, V., et al. (2012). Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cereb. Cortex* 22, 2554–2563. doi: 10.1093/cercor/bhr324
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Bates, E., Camaioni, L., and Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill Palmer Q.* 21, 205–226.
- Boncoddo, R., Dixon, J. A., and Kelley, E. (2010). The emergence of a novel representation from action: evidence from preschoolers. *Dev. Sci.* 13, 370–377. doi: 10.1111/j.1467-7687.2009.00905.x
- Boysen, S. T., Berntson, G. G., Shreyer, T. A., and Hannan, M. B. (1995). Indicating acts during counting by a chimpanzee (*Pan troglodytes*). *J. Comp. Psychol.* 109, 47. doi: 10.1037/0735-7036.109.1.47
- Bredo, E. (1994). Reconstructing educational psychology: situated cognition and Deweyan pragmatism. *Educ. Psychol.* 29, 23–35. doi: 10.1207/s15326985ep2901\_3
- Brooks, R. (1991). Intelligence without representation. *Artif. Intell. J.* 47, 139–160. doi: 10.1016/0004-3702(91)90053-M
- Cary, M., and Carlson, R. A. (1999). External support and the development of problem-solving routines. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1053–1070. doi: 10.1037/0278-7393.25.4.1053
- Chase, W. G., and Ericsson, K. A. (1982). Skill and working memory. *Psychol. Learn. Motiv.* 16, 1–58. doi: 10.1016/S0079-7421(08)60546-0

- Chawla, P., and Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *J. Exp. Soc. Psychol.* 30, 580–601. doi: 10.1006/jesp.1994.1027
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: The MIT Press.
- Chu, M., and Kita, S. (2008). Spontaneous gestures during mental rotation tasks: insights into the microdevelopment of the motor strategy. *J. Exp. Psychol. Gen.* 137, 706. doi: 10.1037/a0013157
- Chu, M., and Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *J. Exp. Psychol. Gen.* 140, 102–115. doi: 10.1037/a0021790
- Chu, M., Meyer, A., Foulkes, L., and Kita, S. (2013). Individual differences in frequency and salience of speech-accompanying gestures: the role of cognitive abilities and empathy. *J. Exp. Psychol. Gen.* 143, 694–709. doi: 10.1037/a0033861
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/978019533213.001.0001
- Clark, A. (2013). "Gesture as thought," in *The Hand, an Organ of the Mind: What the Manual Tells the Mental*, ed. Z. Radman (Cambridge, MA: MIT Press), 255–268.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19. doi: 10.1093/analys/58.1.7
- Cook, S. W., and Tanenhaus, M. K. (2009). Embodied communication: speakers' gestures affect listeners' actions. *Cognition* 113, 98–104. doi: 10.1016/j.cognition.2009.06.006
- Cook, S. W., Yip, T. K., and Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Lang. Cogn. Process.* 27, 594–610. doi: 10.1080/01690965.2011.56707
- De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making. *Phenomenol. Cogn. Sci.* 6, 485–507. doi: 10.1007/s11097-007-9076-9
- De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci.* 14, 441–447. doi: 10.1016/j.tics.2010.06.009
- Delgado, B., Gomez, J. C., and Sarria, E. (2009). "Private pointing and private speech: developing parallelisms," in *Private Speech, Executive Function and the Development of Verbal Self-Regulation*, eds A. Winsler, C. Fernyhough, and I. Montero (Cambridge: Cambridge University Press), 153–162.
- Delgado, B., Gómez, J. C., and Sarriá, E. (2011). Pointing gestures as a cognitive tool in young children: experimental evidence. *J. Exp. Child Psychol.* 110, 299–312. doi: 10.1016/j.jecp.2011.04.010
- De Nooijer, J. A., van Gog, T., Paas, F., and Zwaan, R. A. (2013). Effects of imitating gestures during encoding or during retrieval of novel verbs on children's test performance. *Acta Psychol.* 144, 173–179. doi: 10.1016/j.actpsy.2013.05.013
- De Nooijer, J. A., van Gog, T., Paas, F., and Zwaan, R. A. (in press). Words in action: using gestures to improve verb learning in primary school children. *Gesture*.
- Dove, G. (2010). On the need for embodied and dis-embodied cognition. *Front. Psychol.* 1:242. doi: 10.3389/fpsyg.2010.00242
- Fu, W. T. (2011). A dynamic context model of interactive behavior. *Cogn. Sci.* 35, 874–904. doi: 10.1111/j.1551-6709.2011.01173.x
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Clarendon Press, 173–178.
- Gallagher, S. (2009). "Philosophical antecedents to situated cognition," in *Cambridge Handbook of Situated Cognition*, eds P. Robbins and M. Aydede (Cambridge: Cambridge University Press), 35–53.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help us Think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S., and Alibali, M. W. (2012). "Gesture's role in learning and development," *Oxford Handbook of Developmental Psychology*, ed. P. Zelazo (Oxford: Oxford University Press), 953–973.
- Goldin-Meadow, S., and Beilock, S. L. (2010). Action's influence on thought: the case of gesture. *Perspect. Psychol. Sci.* 5, 664–674. doi: 10.1177/1745691610388764
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., and Wagner, S. (2001). Explaining math: gesticulation lightens the load. *Psychol. Sci.* 12, 516–522. doi: 10.1111/1467-9280.00395
- Glenberg, A. M., Havas, D., Becker, R., and Rinck, M. (2005). "Grounding language in bodily states," in *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, eds D. Pecher and R. A. Zwaan (Cambridge: Cambridge University Press), 115–128. doi: 10.1017/CBO9780511499968.006
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565. doi: 10.3758/BF03196313
- Gonzalez, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., and Sanjuán, A. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage* 32, 906–912. doi: 10.1016/j.neuroimage.2006.03.037
- Graham, J. A., and Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gesture. *Int. J. Psychol.* 10, 57–67. doi: 10.1080/00207597508247319
- Gray, W. D., and Fu, W. T. (2004). Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cogn. Sci.* 28, 359–382. doi: 10.1016/j.cogsci.2003.12.001
- Haselen, G. L. V., van der Steen, J., and Frens, M. A. (2000). Copying strategies for patterns by children and adults. *Percept. Mot. Skills* 91, 603–615. doi: 10.2466/pms.2000.91.2.603
- Hatano, G., Miyake, Y., and Binks, M. G. (1977). Performance of expert abacus operators. *Cognition* 5, 47–55. doi: 10.1016/0010-0277(77)90016-6
- Hatano, G., and Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition* 15, 95–110. doi: 10.1016/0010-0277(83)90035-5
- Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9
- Hegarty, M., Mayer, S., Kriz, S., and Keehner, M. (2005). The role of gestures in mental animation. *Spat. Cogn. Comput.* 5, 333–356. doi: 10.1207/s15427633scc0504\_3
- Hoetjes, M., Krahmer, E., and Swerts, M. (2014). Does our speech change when we cannot gesture? *Speech Commun.* 57, 257–267. doi: 10.1016/j.specom.2013.06.007
- Holler, J., and Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *J. Lang. Soc. Psychol.* 26, 4–27. doi: 10.1177/0261927X06296428
- Hostetter, A. B., and Alibali, M. W. (2008). Visible embodiment: gestures as simulated action. *Psychon. Bull. Rev.* 15, 495–514. doi: 10.1016/j.jml.2010.04.003
- Hostetter, A. B., Alibali, M. W., and Bartholomew, A. E. (2011). "Gesture during mental rotation," in *Proceedings of the 33rd Annual Conference of Cognitive Science Society* (Austin, TX: Cognitive Science Society), 1448–1453.
- Hutchins, E. (1995a). *Cognition in the Wild*. Cambridge: MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cogn. Sci.* 19, 265–288. doi: 10.1016/0364-0213(95)90020-9
- Hutchins, E. (2005). Material anchors for conceptual blends. *J. Pragmat.* 37, 1555–1577. doi: 10.1016/j.pragma.2004.06.008
- Iverson, J. M. (1998). Gesture when there is no visual model. *New Dir. Child Dev.* 79, 89–100. doi: 10.1002/cd.23219987907
- Iverson, J. M., and Thelen, E. (1999). Hand, mouth, and brain: the dynamic emergence of speech and gesture. *J. Conscious. Stud.* 6, 19–40.
- Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (2003). The expertise reversal effect. *Educ. Psychol.* 38, 23–31. doi: 10.1207/S15326985EP3801\_4
- Kaschak, M. P., Zwaan, R. A., Aveyard, M., and Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cogn. Sci.* 30, 733–744. doi: 10.1207/s15516709cog0000\_54
- Kelly, S. D. (2001). Broadening the units of analysis in communication: speech and nonverbal behaviours in pragmatic comprehension. *J. Child Lang.* 28, 325–349. doi: 10.1017/S0305000901004664
- Kelly, S. D., Barr, D. J., Church, R. B., and Lynch, K. (1999). Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory. *J. Mem. Lang.* 40, 577–592. doi: 10.1006/jmla.1999.2634
- Kendon, A. (1994). Do gestures communicate? A review. *Res. Lang. Soc. Interact.* 27, 175–200. doi: 10.1207/s15327973rlsi2703\_2
- Kirsh, D. (1995). The intelligent use of space. *Artif. Intell.* 73, 31–68. doi: 10.1016/0004-3702(94)00017-U
- Kirsh, D. (2009). "Interaction, external representation and sense making," in *Proceedings of the Thirty First Annual Conference of the Cognitive Science Society*, eds N. A. Taatgen and H. V. Rijn (Austin, TX: Cognitive Science Society), 1103–1108.
- Kirsh, D. (2010). *Thinking with External Representations. AI and Society*. Springer: London, 25, 441–454.
- Kirsh, D., and Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cogn. Sci.* 18, 513–549. doi: 10.1016/0364-0213(94)90007-8
- Kita, S. (2000). "How representational gestures help speaking," in *Language and Gesture*, ed. D. McNeill (Cambridge: Cambridge University Press), 162–185.
- Kita, S., and Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? evidence for an interface representation of spatial thinking and speaking. *J. Mem. Lang.* 48, 16–32. doi: 10.1016/S0749-596X(02)00505-3

- Krauss, R. M. (1998). Why do we gesture when we speak? *Curr. Dir. Psychol. Sci.* 7, 54–60. doi: 10.1111/1467-8721.ep13175642
- Krauss, R. M., Chen, Y., and Gottesmann, R. F. (2000). “Lexical gestures and lexical access: a process model,” in *Language and Gesture*, ed. D. McNeill (Cambridge: Cambridge University Press), 261–283.
- Krauss, R. M., and Hadar, U. (1999). “The role of speech-related arm/hand gestures in word retrieval,” in *Gesture, Speech, and Sign*, eds R. Campbell and L. Messing (Oxford: Oxford University Press), 93–116.
- Liszkowski, U., Brown, P., Callaghan, T., Takada, A., and de Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cogn. Sci.* 36, 698–713. doi: 10.1111/j.1551-6709.2011.01228.x
- Manches, A., O'Malley, C., and Benford, S. (2010). The role of physical representations in solving number problems: a comparison of young children's use of physical and virtual materials. *Comput. Educ.* 54, 622–640. doi: 10.1016/j.compedu.2009.09.023
- Marstaller, L., and Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychon. Bull. Rev.* 20, 496–500. doi: 10.3758/s13423-012-0365-0
- Martin, T., and Schwartz, D. L. (2005). Physically distributed learning: adapting and reinterpreting physical environments in the development of fraction concepts. *Cogn. Sci.* 29, 587–625. doi: 10.1207/s15516709cog0000\_15
- McNeil, N. M., Alibali, M. W., and Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: now they need it, now they don't. *J. Nonverbal Behav.* 24, 131–150. doi: 10.1023/A:1006657929803
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychol. Rev.* 92, 350–371. doi: 10.1037/0033-295X.92.3.350
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL: University of Chicago Press.
- Morsella, E., and Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *Am. J. Psychol.* 117, 411–424. doi: 10.1007/s10936-005-6141-9
- Ping, R., and Goldin-Meadow, S. (2010). Gesticulation saves cognitive resources when talking about nonpresent objects. *Cogn. Sci.* 34, 602–619. doi: 10.1111/j.1551-6709.2010.01102.x
- Post, L. S., van Gog, T., Paas, F., and Zwaan, R. A. (2013). Effects of simultaneously observing and making gestures while studying grammar animations on cognitive load and learning. *Comput. Hum. Behav.* 29, 1450–1455. doi: 10.1016/j.chb.2013.01.005
- Pouw, W. T. J. L., van Gog, T., and Paas, F. (2014). An embedded and embodied cognition review of instructional manipulatives. *Educ. Psychol. Rev.* 26, 51–72. doi: 10.1007/s10648-014-9255-5
- Pulvermüller, F., Moseley, R., Egorova, N., Shebani, Z., and Boulenger, V. (2014). Motor cognition—motor semantics: action perception theory of cognition and communication. *Neuropsychologia* 34, 602–619. doi: 10.1016/j.neuropsychologia.2013.12.002
- Radman, Z. (2013). *The Hand, an Organ of the Mind: What the Manual Tells the Mental*. Cambridge, MA: MIT Press.
- Rauscher, F. H., Krauss, R. M., and Chen, Y. (1996). Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychol. Sci.* 7, 226–231. doi: 10.1111/j.1467-9280.1996.tb00364.x
- Rimé, B., Schiaratura, L., Hupet, M., and Ghyselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motiv. Emot.* 8, 311–325. doi: 10.1007/BF00991870
- Risko, E. F., Medimorec, S., Chisholm, J. D., and Kingstone, A. (2013). Rotating with rotated text: a natural behavior approach to investigating cognitive offloading. *Cogn. Sci.* doi: 10.1111/cogs.12087 [Epub ahead of print].
- Rupert, R. D. (2009). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195379457.001.0001
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660
- Schwartz, D. L., and Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 116. doi: 10.1037/0278-7393.25.1.116
- Seyfeddinipur, M., and Kita, S. (2001). “Gesture and dysfluency in speech,” in *Oralite et gesturalite: Interactions et comportements multimodaux dans la communication*, eds C. Cave, I. Guaitella, and S. Santi (Paris: L'Harmattan), 266–270.
- Shapiro, L. (2010). *Embodied Cognition*. London: Taylor & Francis.
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703. doi: 10.1126/science.171.3972.701
- Smithson, L., and Nicoladis, E. (2014). Lending a hand to imagery? The impact of visuospatial working memory interference upon iconic gesture production in a narrative task. *J. Nonverbal Behav.* 38, 247–258. doi: 10.1007/s10919-014-0176-2
- Stephen, D. G., Dixon, J. A., and Isenhower, R. W. (2009). Dynamics of representational change: entropy, action, and cognition. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1811–1832. doi: 10.1037/a0014510
- Stevanoni, E., and Salmon, K. (2005). Giving memory a hand: instructing children to gesture enhances their event recall. *J. Nonverbal Behav.* 29, 217–233. doi: 10.1007/s10919-005-7721-y
- Stull, A. T., Hegarty, M., Dixon, B., and Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cogn. Instr.* 30, 404–434. doi: 10.1080/07370008.2012.719956
- Svensson, H. (2007). *Embodied Simulation as Off-line Representation*. Doctoral dissertation, Department of Computer and Information Science, Linköping University, Linköping.
- Tomasello, M. (2008). *Origins of Human Communication*. Boston: MIT Press.
- Tomasello, M., Carpenter, M., and Liszkowski, U. (2007). A new look at infant pointing. *Child Dev.* 78, 705–722. doi: 10.1111/j.1467-8624.2007.01025.x
- Valenzano, L., Alibali, M. W., and Klatzky, R. (2003). Teachers' gestures facilitate students' learning: a lesson in symmetry. *Contemp. Educ. Psychol.* 28, 187–204. doi: 10.1016/S0361-476X(02)00007-3
- Wagner, S. M., Nusbaum, H., and Goldin-Meadow, S. (2004). Probing the mental representation of gesture: is handwaving spatial? *J. Mem. Lang.* 50, 395–407. doi: 10.1016/j.jml.2004.01.002
- Warburton, E. C., Wilson, M., Lynch, L., and Cuykendall, S. (2013). The cognitive benefits of movement reduction: evidence from dance marking. *Psychol. Sci.* 24, 1732–1739. doi: 10.1177/0956797613478824
- Wesp, R., Hesse, J., Keutmann, D., and Wheaton, K. (2001). Gestures maintain spatial imagery. *Am. J. Psychol.* 114, 591–600. doi: 10.2307/1423612
- Wheeler, M. (2007). *Reconstructing the Cognitive World: The Next Step*. Cambridge: MIT Press.
- Wheeler, M. (2013). “Is cognition embedded or extended? The case of gestures,” in *The Hand, an Organ of the Mind: What the Manual Tells the Mental*, ed. Z. Radman (Cambridge, MA: MIT Press), 269–301.
- Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322
- Zwaan, R. A. (in press). Embodiment and language comprehension: reframing the discussion. *Trends Cogn. Sci.* doi: 10.1016/j.tics.2014.02.008

**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.

Received: 30 January 2014; accepted: 06 April 2014; published online: 24 April 2014.  
Citation: Pouw WTJL, de Nooijer JA, van Gog T, Zwaan RA and Paas F (2014) Toward a more embedded/extended perspective on the cognitive function of gestures. *Front. Psychol.* 5:359. doi: 10.3389/fpsyg.2014.00359  
This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.  
Copyright © 2014 Pouw, de Nooijer, van Gog, Zwaan and Paas. 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.



# The body and the fading away of abstract concepts and words: a sign language analysis

Anna M. Borghi<sup>1\*</sup>, Olga Capirci<sup>2</sup>, Gabriele Gianfreda<sup>3</sup> and Virginia Volterra<sup>2\*</sup>

<sup>1</sup> Department of Psychology, University of Bologna and Institute of Cognitive Sciences and Technologies, Italian National Research Council, Rome, Italy

<sup>2</sup> Institute of Cognitive Sciences and Technologies, Italian National Research Council, Rome, Italy

<sup>3</sup> National Institute for the Deaf, Rome, Italy

## Edited by:

Guy Dove, University of Louisville,  
USA

## Reviewed by:

Arthur M. Glenberg, Arizona State  
University, USA

Katja Wiemer, Northern Illinois  
University, USA

## \*Correspondence:

Anna M. Borghi, Department of  
Psychology, University of Bologna,  
Viale Berti Pichat, 5, 40127 Bologna,  
Italy

e-mail: anna.borghi@gmail.com;

Virginia Volterra, Institute of  
Cognitive Sciences and  
Technologies, Via Nomentana 56,  
00171 Rome, Italy

e-mail: vvolterra@teletu.it

One of the most important challenges for embodied and grounded theories of cognition concerns the representation of abstract concepts, such as “freedom.” Many embodied theories of abstract concepts have been proposed. Some proposals stress the similarities between concrete and abstract concepts showing that they are both grounded in perception and action system while other emphasize their difference favoring a multiple representation view. An influential view proposes that abstract concepts are mapped to concrete ones through metaphors. Furthermore, some theories underline the fact that abstract concepts are grounded in specific contents, as situations, introspective states, emotions. These approaches are not necessarily mutually exclusive, since it is possible that they can account for different subsets of abstract concepts and words. One novel and fruitful way to understand the way in which abstract concepts are represented is to analyze how sign languages encode concepts into signs. In the present paper we will discuss these theoretical issues mostly relying on examples taken from Italian Sign Language (LIS, *Lingua dei Segni Italiana*), the visual-gestural language used within the Italian Deaf community. We will verify whether and to what extent LIS signs provide evidence favoring the different theories of abstract concepts. In analyzing signs we will distinguish between direct forms of involvement of the body and forms in which concepts are grounded differently, for example relying on linguistic experience. In dealing with the LIS evidence, we will consider the possibility that different abstract concepts are represented using different levels of embodiment. The collected evidence will help us to discuss whether a unitary embodied theory of abstract concepts is possible or whether the different theoretical proposals can account for different aspects of their representation.

**Keywords:** abstract concepts, abstract words, Italian Sign Language (LIS), sign languages, embodied cognition, metaphor, signs, iconicity

## INTRODUCTION

To what extent are cognitive capacities learnt through action? According to embodied and grounded views, acting and interacting with the objects and the physical and social entities present in the environment represent the basis of our cognitive abilities (e.g., Wilson, 2002). Research on embodied and grounded cognition has rapidly grown in the last 10–15 years, as widely acknowledged by different scholars (e.g., Chatterjee, 2010; Gentner, 2010; for a review see Borghi and Caruana, in press).

In the last years, much behavioral and neuroscience evidence has been provided, showing that concepts and language are grounded on perception and action systems (for reviews, see Gallese and Lakoff, 2005; Barsalou, 2008; Fischer and Zwaan, 2008; Gallese, 2008; Jirak et al., 2010; Meteyard et al., 2012; for special issues, see Borghi and Pecher, 2011). However, the perspective of embodied and grounded cognition is confronted with some unsolved issues and open challenges. One of the major challenges is represented by the possibility to account for the representation of abstract concepts and words meanings (see

the recent special issue by Tomasino and Rumiati, 2013). With “abstract words meanings” we intend the meaning of words such as “philosophy” and “truth,” that apparently do not have a single, easily identifiable, imaginable and concrete referent. Their referents are instead situations, events, mental states, conditions. Specifically, whether the embodied account holds only for concrete concepts and words or whether it can be extended to abstract concepts and words as well is still a matter of debate. A number of scholars have argued that, while embodied theories are able to account for words referring to concrete objects (e.g., bottle), supported by convincing evidence, the story is completely different if we consider the domain of abstract words, due both to theoretical limits and to the lack of compelling empirical evidence (e.g., Dove, 2009, 2011).

Our paper deals with abstract concepts representation. First, we will consider the possibility that different degrees of embodiment are involved in the representation of concrete and abstract concepts. Second, we will verify whether different abstract concepts are represented using different levels of embodiment. We



will distinguish between direct forms of involvement of the body and forms in which concepts are grounded differently, for example relying on linguistic experience. To handle these theoretical issues in the present paper we will first provide a brief outline of the major recent accounts of abstract concepts within embodied and grounded theories (for recent reviews see Pecher et al., 2011; Borghi and Binkofski, 2014). The embodied cognition perspective has indeed developed different proposals that attempt to explain abstract concepts representation.

The novelty of the present contribution is that we will verify the solidity of these theories in light of examples taken from one of the many Sign Languages (from now on SL): the Italian Sign Language (from now on LIS, *Lingua dei Segni Italiana*), the language used within the Italian Deaf community, described and analyzed since about 30 years.

### DEFINITION

Defining abstract concepts and words is not an easy task. It is noteworthy that the term “abstract” is represented in LIS by a sign located near the head and referring to something that cannot be touched and grasped, to something that is not material and concrete but that rather fades away.

Here we will adopt a rather broad operational definition of abstract terms. We define abstract the words and the signs that, differently from concrete ones, do not refer to single, concrete and manipulable items, but are rather grounded in situations, events, mental states, etc. Abstract words are typically rated as less imaginable as concrete ones, they are more complex than concrete words since they often refer to relations between elements rather than to single objects/entities, and they are characterized by higher intersubjective and intra-subjective variability (see Borghi and Binkofski, 2014, for clarifications on this definition). Notice however that the opposition between concrete and abstract concepts might not be a dichotomy but rather a continuum. Ratings asking people to judge the concreteness of large sets of words showed that concrete and abstract concepts are distributed in a bimodal way, falling into two big clusters (according to features, such as tangibility or visibility); within each cluster, however, the entities had different concreteness degrees (Nelson and Schreiber, 1992; Wiemer-Hastings et al., 2001).

Despite the difficulty in finding a shared definition, embodied theories of abstract concepts are numerous; below we will briefly illustrate the most important ones.

### MAIN EMBODIED THEORIES OF ABSTRACT WORDS

According to classical Embodied Cognition (EC) theories of abstract words there would not be a substantial difference between concrete and abstract words, since both are grounded in perception, action and emotional systems. For example, the abstract concepts of number would be grounded in action due to finger counting experience (for a review, see Fischer and Brugger, 2011). Further evidence in support of this view is obtained by studies that link words to action, for example by evidence on the Action-sentence Compatibility Effect (ACE). Results showed that judging the sensibility of sentences which describe the transfer of both concrete objects and abstract information (e.g., “giving the pizza” vs. “giving the information”) requires less time when the action implied by the sentence matches the action required to make the response (Glenberg and Kaschak, 2002; Glenberg et al., 2008a,b). This finding suggests that the mechanisms underlying transfer of abstract concepts (e.g., “the information”) are the same as those underlying transfer of concrete ones (e.g., “the pizza”) (see also Guan et al., 2013).

The other EC theories we will illustrate posit that abstract and concrete concepts and words are represented differently. The most influential one is probably the Conceptual Metaphor Theory, which states that abstract concepts are represented by image schemas derived from concrete domains. Evidence supporting this theory has shown for example that similarity is represented as closeness, categories as containers, and that the abstract notion of time is mapped onto the concrete domain of space (e.g., Lakoff and Johnson, 1980; Gibbs and Steen, 1999; Boroditsky and Ramscar, 2002; Casasanto and Boroditsky, 2008; Boot and Pecher, 2010, 2011; Casasanto et al., 2010; Flusberg et al., 2010; Lai and Boroditsky, 2013).

Further theories identify differences in content between concrete and abstract concepts. According to Barsalou and Wiemer-Hastings (Barsalou, 1999; Barsalou and Wiemer-Hastings, 2005), abstract concepts differ from concrete concepts as the first activate situations and introspective relationships more frequently. Evidence in favor of this approach is based mainly on results of feature generation tasks, showing that, whereas with concrete concepts, such as “bottle,” people tend to produce mostly properties referring to perceptual characteristics such as color, size, shape, matter, parts (e.g., “green,” “plastic,” “neck”), abstract concepts such as “freedom” evoke more frequently situations, events, introspective states (e.g., “running on the grass,” “exiting from prison,” etc.).

A novel proposal advanced by Vigliocco and colleagues (Kousta et al., 2011; Vigliocco et al., 2014) states that abstract concepts differ from concrete ones in content, since they rely more on emotional experience. Analyzing a large database Kousta et al. (2011) demonstrated that, when imageability was kept constant, emotional valence was a significant predictor of concreteness ratings. Recent brain imaging evidence (Vigliocco et al., 2014) further supports this view.

Other recent approaches, such as the Language and Situated Simulation Theory (LASS) (Barsalou et al., 2008; Simmons et al., 2008), the Symbol Interdependence Theory (Louwerse and Connell, 2011), the proposal by Dove (2011, 2014) and the Words As social Tools (WAT) proposals (Borghi and Cimatti, 2009; Borghi, 2014; Borghi and Binkofski, 2014; evidence in Borghi et al., 2011; Scorolli et al., 2011, 2012; Sakreida et al., 2013), argue that both linguistic and sensorimotor information are crucial for conceptual representation. LASS does not specifically focus on abstract concepts, but on conceptual representation more generally. According to LASS, both the linguistic and the simulation system are activated during conceptual processing; the linguistic system is faster and more superficial, while the simulation system is engaged for understanding of meaning. In some situations using the linguistic system represents a shortcut as it allows to respond immediately to a task (particularly to linguistic tasks) without necessarily accessing to

conceptual meaning (Pecher and Boot, 2011). In a similar vein, Louwerse's Symbol Interdependency Theory states that shallow linguistic representations precede deeper perceptual representations (Louwerse, 2011; Louwerse and Connell, 2011; Connell and Lynott, 2012).

Compared to the other multiple representation theories, WAT (Borghi and Cimatti, 2009; Borghi and Binkofski, 2014) and Dove's view (Dove, 2014) focus specifically on the difference between concrete and abstract concepts and words. According to both views abstract concepts representation relies more on language than representation of concrete words. In his proposal on abstract concepts Dove (2011, 2014) stresses the important scaffolding role language can play and the fact that the abilities acquired thanks to language allow its use not only as a means of communication but of thought as well. The main tenets of WAT are the following: a. both concrete and abstract concepts are embodied and grounded in perception and action systems, b. for abstract concepts linguistic information plays a more crucial role than for concrete ones, c. this is due to the different acquisition modality of concrete and abstract words; d. this distributional difference is reflected in the representation in the brain of concrete and abstract concepts, e. given that representation of abstract concepts is more influenced by language, linguistic diversity has a major impact on abstract concepts representation. An important principle of the WAT proposal concerns the acquisition mechanism of the two kinds of words: with concrete words, the concrete entities (e.g., book) can be perceived together with their linguistic labels. In the case of abstract words, the linguistic experience might be more important, because typically abstract words do not have a single concrete referent and also because they usually refer to exemplars differing to a great extent. Verbal labels are hence used to assemble a set of quite sparse and diverse sensorimotor experiences (e.g., we probably put together different experiences of freedom once we have learned the word "freedom"). Evidence in support of this proposal is multifaceted (for review see Borghi and Binkofski, 2014). Brain imaging studies demonstrated greater engagement of the verbal system for processing of abstract concepts and greater engagement of the perceptual and motor system for concrete concepts (e.g., Binder et al., 2005; Sabsevitz et al., 2005; Rüschemeyer et al., 2007; Desai et al., 2010; Sakreida et al., 2013), and behavioral research has shown a high cross-linguistic variability with abstract words (e.g., Boroditsky, 2011). Notably, acquisition evidence has shown that the process of acquisition of the two kinds of words might differ (e.g., Wauters et al., 2003; Borghi et al., 2011). In particular, studies on Mode of Acquisition (MOA) (e.g., Wauters et al., 2003) have shown that children acquire the meaning of concrete words, such as "bottle," associating the word with its referent, the bottle, or with an action typically performed with or on the bottle by themselves or by another individual (Capirci et al., 2005). The meaning of abstract words like "grammar" or "philosophy," instead, has to be explained by means of language. Finally, the meaning of a word like "tundra" can be acquired in both ways, depending on the environment where it is learned. MOA ratings, which correlate but are not totally explained by age of acquisition, concreteness and imageability, gradually change with age: initially acquisition is mainly perceptual, later it is mainly linguistic.

## THE CHALLENGE

The question theorists adopting an EC approach have to ask is the following: is it possible to account for abstract words with a unified framework? Isn't it possible, instead, that the domain of abstract words is not homogeneous, and that the different subsets of abstract words have to be explained relying on different mechanisms? Recent studies showing fine-grained differences between subsets of abstract words (e.g., Ghio et al., 2013; Roversi et al., 2013) suggest that this might be the case. For example, abstract words as diverse as "category," "truth," and "risk" could rely on different mechanisms: the first could metaphorically evoke a container (Boot and Pecher, 2010), the second could evoke linguistic information and the third might activate situations. If this is true, this would lead us to abandon the overall notion of abstractness and to partition the domain into sub-domains of abstract words.

One intriguing way to understand the way in which abstract words are represented and to deal with the challenge abstract words pose to the EC perspective is to analyze how they are dealt with in sign languages. In our opinion, the way in which sign languages encode concepts into signs can help us understand how abstract linguistic items are represented, and which theory among those on abstract concepts can better account for their meaning.

Linguistic research undertaken since Stokoe's (1960) seminal work on American Sign Language (ASL) has led to the discovery and description of a very large number of national sign languages, now widely recognized by the scientific community as full-fledged, natural languages, which include Italian Sign Language or LIS (Volterra, 1987; Pizzuto and Corazza, 1996). In the last edition of the Ethnologue database 137 Sign Languages (SL) are listed. It has been shown that, even though these languages are perceived and produced in the visual-gestural (rather than in the vocal-auditory) modality, they satisfy the communicative and expressive needs of a community and possess all the basic linguistic components including phonological, lexical, syntactic and grammatical systems. Just as words of a spoken language are formed on the basis of phonemes in various combinations, all signs of a signed language are formed by combining a defined number of formational parameters (called also as cheremes). More precisely, a sign can be broken down into four basic parameters: the form or configuration taken on by the hand; the orientation the hand takes on while making the sign; the location in which the sign is performed; the movement the hand describes.

As Penny Boyes Braem pointed out already in 1981, signed lexical units are often made up of formal features visually motivated and thereby iconic. Their visual motivation is not idiosyncratic, it derives from regularities at the level of formational parameters. Handshapes, for example, are often linked to features of a sign's meaning via reference to some peculiar visual forms (Pizzuto et al., 1995; Pietrandrea and Russo, 2007). The same holds true for location and often for movement (for a comprehensive analysis of the iconicity of the LIS parameters, see Pietrandrea, 2002). In spite and beyond important structural resemblances between Sign Languages and Vocal Languages, equally relevant structural differences need to be taken in due account (Sutton-Spence, 2005; Cuxac and Sallandre, 2007; Pizzuto et al., 2007; Perniss et al., 2010; West and Sutton-Spence, 2010; Boyes Braem et al., 2012; Meurant et al., 2013; Perniss and Vigliocco, 2014). The grammar

and the syntax of a sign language are expressed in various ways, including use of space, modulation of movement, facial expression and position of the trunk and shoulders. A great deal of research has been carried out on the signs used by the Deaf Italian community (a complete bibliography on LIS is available at biblioLIS [http://www.istc.cnr.it/sites/default/files/u182/bibliolis\\_arg\\_2011.pdf](http://www.istc.cnr.it/sites/default/files/u182/bibliolis_arg_2011.pdf)).

To our knowledge the relationship between sign languages and abstract concepts has been investigated in a few studies so far (e.g., West and Sutton-Spence, 2010). In 2005 the Journal “Sign Language Studies” devoted a Special Issue to a crosslinguistic analysis of SL in the metaphorical domains of thought and communication. Linguists studying different sign languages (British, American, Catalan, and Italian) examined the mappings involved in SL metaphors, showing the process of embodiment active in metaphorical structures. Some structures share similarities across sign languages but there are also some interesting differences. Russo (2005) suggests that signed language metaphors are intrinsically related to aspects of the linguistic and cultural dimensions of a specific deaf community. More recently Roush (2011) has addressed the issue of the cognitive representation of abstract terms in sign languages. The author analyzed how a number of abstract words are represented in American Sign Language (ASL). Roush (2011) applied a specific linguistic-cognitive framework, the Conceptual Metaphor Theory, to investigate how the area of (im)politeness is conceptualized through metaphors and reflected and iconically represented in ASL. Our approach shares with Roush the view that using sign languages is an important perspective helping understand the way in which concepts are represented, however the ultimate aim why we use sign languages for investigating cognitive issues is slightly different. While Roush focuses on a specific theory we move from a variety of embodied theories struggling to account for abstract concepts representation. Specifically, our investigation is aimed at analyzing how abstract concepts belonging to different domains are represented in LIS, assuming that this analysis will allow us to understand whether the category of abstract terms is homogeneous or whether it needs to be re-organized into different sub-sets.

## HYPOTHESES

We advance the following hypotheses. First, in line with all embodied theories we predict that all the considered abstract concepts are at least in part grounded in the sensorimotor system. This guarantees the fact that the problem of symbol grounding (Harnad, 1990) is not present, since symbols used to represent abstract concepts are not arbitrarily linked to their referents.

At the same time, however, we predict that theories taking into account only sensorimotor nonlinguistic information will not be able to explain all examples we provide. In our view a unified framework, either based only on sensorimotor (for a review, see Pecher et al., 2011) or only on linguistic information (e.g., Paivio, 1986) will not be able to account for the differences between kinds of abstract concepts. In line with multiple representation theories we predict, instead, that to account for some abstract concepts a combination of sensorimotor, emotional, and linguistic information will be necessary. With “linguistic information” we

intend any kind of exploitation of forms derived from any kind of language, be the same sign language or a different sign or spoken language. An example is the concept of “causation”: it is grounded in sensorimotor information since it might activate a variety of situations in which, for example, one element determines an effect on another one (e.g., a ball hurting another ball and provoking its movement, a handle being pressed to open a door etc.); at the same time, however, to acquire the concept children might rely on explanations of what causation is provided by others, such as parents or teachers, or by authoritative written sources, such as dictionaries, encyclopedias, etc. Another example highlighting how the formation of abstract concepts can rely on linguistic sources is the concept of “linguistics,” which originates from and refers to the more concrete concept of “language.” Specific examples pertaining SLs, such as LINGUISTICS, LANGUAGE, TRUTH, etc., are discussed later in the paper. To highlight the role of linguistic information we have selected on purpose concepts where the role of linguistic elements is particularly evident, even if sensorimotor information still plays a role. This combination of sensorimotor and linguistic information is what we mean when we speak of “different levels of embodiment.”

## LIS EVIDENCE

In the present section we will provide novel evidence on LIS signs supporting the most important theories we have presented. The examples we are going to illustrate and discuss are mainly taken from a corpus collected by Gianfreda (2011; Gianfreda et al., 2014). The corpus was originally collected to explore the linguistic forms through which Italian Sign Language (LIS) signers realize communicative functions related to the expression of certainty and uncertainty, focusing on dimensions already explored for spoken Languages and for which theoretical constructs such as epistemic modality and evidentiality have been proposed. Conversations in LIS between deaf people communicating through a video-chat software have been collected and analyzed. In this type of interaction, the technological instrument itself permits to record the conversations in a less intrusive manner. Both participants are obliged to maintain themselves in front of the webcam and to optimize video quality in order to understand their sign language productions. The software automatically creates, in real time, two video windows for each interlocutor; through split-screen it is possible to analyze efficiently the synchronization between signs, facial expressions and body actions produced by both participants. Focusing on low-structured interactions we have been able to observe linguistic units typical of LIS as they spontaneously emerge in effective situations of language use.

The corpus consisted of six exchanges: four completely free and two on a suggested topic. The time duration range of conversations was from 23 to 51 min. Conversational exchanges in which signers were expressing certainty and/or uncertainty have been identified and transcribed through Sign Writing (SW: Sutton, 1999). SW is a system based on a set of “glyphs,” which, combined together in graphic units, permit to write or transcribe signs, allowing an external reader to reconstruct sign language forms. A textual qualitative analysis has been conducted to better identify and describe the linguistic forms used by the LIS signers.



All examples of signs provided and words reported in the present paper to support different theories of abstract concepts are selected from the corpus above described except for the last three LIS signs mentioned in the present paper: LANGUAGE/LINGUAGGIO, LINGUISTICS, and COMMUNICATION. Our analysis has obviously no pretense to be exhaustive. However, we believe that providing examples supporting or disconfirming a given theory is a useful strategy. Consider for example studies providing support to the Conceptual Metaphor Theory: in one study it is shown that similarity is conceived as spatial contiguity (Boot and Pecher, 2010), in another that category is intended in terms of container (Boot and Pecher, 2011), in many studies it is shown that the abstract notion of time is conceived in terms of the more concrete notion of space (e.g., Boroditsky and Ramscar, 2002; Casasanto, 2008). These examples provide support to the theory, even though they do not tell us that the theory is necessarily always true. At the same time, providing even one single example disconfirming a theory can widely limit its application range, or its generality. This is exactly the strategy we will follow in the present paper. In the present text signs are reported by English glosses and often by figures.<sup>1</sup> A complete list of all the figures can be found in the supplementary materials.

Different signs can provide support for the Conceptual Metaphor Theory. Specifically, we will refer to examples that highlight the use of body parts in an iconic way to refer to underlying metaphors. These manual signs are executed in different iconically motivated body parts (e.g., eyes, head, chest).

Concrete examples are represented by the LIS signs glossed as SEE and HEAR. Both verbs refer to the acquisition of characteristics of external reality through the appropriate sensorial organs. The movement of the first sign starts from the eye toward the external space while the second sign is executed near the ear with a movement toward the body. Two further signs are executed in these face locations, i.e., PERCEIVE-THROUGH-SIGHT and PERCEIVE-THROUGH-HEARING.

These two signs share the same configuration and the same movement, but their different locations indicate the different sensorial modalities (sight and hearing) through which the perceptions occur. Notice that deaf people tend to exclude audition when they refer to perceptual activity in general since this modality is not very useful in their representation of the world. The verbs HEAR and PERCEIVE-THROUGH-HEARING are strictly associated to experiences of hearing individuals. This aspect helps us understand why in LIS the notion KNOWING IS SEEING is more meaningful and therefore more used. Several metaphors rely on this concept and explain many LIS lexical units. For example in the sign CLEAR (Figure 1) both hands are initially located in front

of the eyes with hand configurations suggesting an initial partial obscurity. The two hands move laterally, away from the body, expressing broad, unimpeded perception. The same hand configuration is used for the sign SEEM, which is typically used to express something acquired through perception. The association between the perceived entity and its interpretation is uncertain (for the corresponding ASL sign, see Wilcox and Wilcox, 1995; Wilcox and Shaffer, 2006).

The location in which the sign SEEM is produced, i.e., the space between the forehead and the eyes, reflects perceptual and cognitive processes. The signer indicates that his/her epistemic belief concerning the content he/she is expressing is grounded on some kind of evidence, which should be further verified. The sign can be linked not only with inferences based on acquired evidence but also on memory retrieval. When the sign SEEM is produced with half-closed eyes, and sometimes also with tensed cheeks, it expresses a focusing process concerning perception or memory.

Many verbs are produced around the forehead. For example, TO LEARN, TO KNOW, TO UNDERSTAND, TO FORGET, TO REMEMBER, and ACKNOWLEDGED all seem to link to the underlying metaphor of the head as the location of cognitive and memory activities. For the sign ACKNOWLEDGED, the signer first locates his/her index finger in the direction of the head; after this first movement a quick rotation of the wrist with the open hand follows, representing the sign translatable as FINISH, which allows indicating the completion of the action expressed from the main verb. The mental process is signaled in a slightly different way from the sign TO KNOW (Figure 2) in which the fingers thumb, index and medium, extended, quickly touch each other.



FIGURE 1 | LIS sign CLEAR.



FIGURE 2 | LIS sign TO KNOW.

<sup>1</sup>Glosses, better known as interlinear glosses, are used in different areas of linguistics in order to give an account of the meaning/description of the morphemes of a given language. The use of glosses in sign language research is a useful practice, but should not be considered a self sufficient representation system neglecting the general requirement of being associated with a transcription of the form of the morpheme. Otherwise it is not possible to verify (discuss or contradict) any morphological analysis conducted, since no formal property of the sign can be used in order to check the consistency of data and analysis provided (Pizzuto and Pietrandrea, 2001; Petitta et al., 2013).



In the sign TO REMEMBER, instead, the index and medium finger, extended and joined, are placed on the forehead, suggesting that the remembered object is stably located within the head. Some of these verbs, also located around the forehead, i.e., TO LEARN, TO UNDERSTAND, and TO FORGET, rely on the underlying metaphor of the MIND AS CONTAINER: perceptual traces, recalls, linguistic information, conceptual nets are formed and stored in the head. Clearly present in the conceptual metaphor here is a movement toward or away from the head. One of the clearest examples is the sign TO LEARN in which all the extended digits quickly touch each other and move toward the signer's forehead as if bringing in something from the external space (see Supplementary Materials). The same digit configuration, but with the palm of the hand orientated laterally to the head and combined with a repeated circular movement is found in the sign TO THINK (see supplementary materials). The forehead location, symbolizing the place where the "objects" of perceptual, mnemonic and cognitive processes can be seen and manipulated, explains the formation of many lexical units in a variety of sign languages (see Brennan, 2005; Jarque, 2005; Russo, 2005; Wilcox, 2005, 2007).

Another interesting example is the sign TO UNDERSTAND (Figure 3), which uses the same movement found in LIS to indicate grasping of physical objects. The main difference between the signs TO UNDERSTAND and TO GRASP is in their location: to grasp is located in the neutral space in front of the signer's chest, whereas TO UNDERSTAND is produced near the signer's head; this clearly represents a form of metaphorical extension, as it suggests that understanding is grasping and putting something in the head-container (Russo, 2004). This metaphor reflects the Latin etymology of the word *com-prehendere*, which is maintained also in other European sign languages. In ASL, a different underlying metaphor is present: the concept TO UNDERSTAND is conveyed through a fist-like handshape placed near the forehead from which the index finger is then extended, indicating the emergence of a thought-object from mental processes (Wilcox, 2005).

The metaphor of the head as container underlies also the LIS sign TO FORGET (Figure 4), in which the closed hand moves to the other side of the head, symbolizing the sliding away of a mental object which had been previously "grasped" by the signer, and opens: the close hand indeed moves away from the head toward the lateral space.



FIGURE 3 | LIS sign TO UNDERSTAND.

The examples discussed so far support the idea that abstract terms are represented through conceptual metaphors. But some signs, such as TO LEARN, TO UNDERSTAND, TO FORGET, also support the ACE view, as actions executed with physical objects are relevant for the representation of the concept expressed through the metaphor.

Other LIS signs expressing uncertainty are linked to a concrete physical object such as a balance.

In the LIS sign TO DOUBT (Figure 5) the oscillating movement of the two hands with downward orientated palms expresses uncertainty. The ASL sign MAYBE looks very similar but the hand configuration differs, as the hand palms are oriented upwards, referring more explicitly to a balance with two similar weights, metaphorically extended to cognitive activity (Wilcox and Wilcox, 1995; Wilcox, 1996).

The LIS signs PERHAPS/MAYBE and ABOUT both have hand-shapes and locations which are very similar to that of TO DOUBT, but differ in their movement of an oscillating wrist. These two signs occur, however, in different contexts, in which they are accompanied by different mouth<sup>2</sup> patterns. PERHAPS tends to reinforce hypothetical statements, or to reduce the impact of the speaker's statements. ABOUT, instead, can be mostly found in

<sup>2</sup>In LIS, as in all sign languages analyzed so far, signs are often accompanied by mouth patterns. Two main categories are distinguished: (i) mouthings which are derived and represent words or parts of words from a spoken language, and (ii) mouth gestures which are idiomatic gestures produced by the mouth not related to a spoken language (Boyes Braem and Sutton Spence, 2001; Fontana, 2008).



FIGURE 4 | LIS sign TO FORGET.



FIGURE 5 | LIS sign TO DOUBT.

expressions in which the signer defines numerical quantities or time periods, ascribing a character of approximation to the expressed values.

Other signs are executed on different body locations, which can also provide a motivation from an iconic point of view. For example many LIS signs executed on the chest are referring to feelings, such as LOVE, HATRED, RAGE. However, signs linked to mental activity can also be produced near the chest. For example, the sign TO BELIEVE (**Figure 6**) is made with the upper side of the two fists touching the heart; in LIS this sign can also mean TRUST.

A sign that specifically supports the ACE view is TO CONSTRAIN. In this “agreement verb,” the hand (thumb and index finger bent as if to grasp a small object) can move toward the signer’s neck or with reversed palm orientation move toward another point in space. This change in palm orientation and movement direction specifies the arguments of the verb (“x is constrained by y,” “x constrains y”). The underlying metaphor is clearly linked to the expression “Grab somebody by the throat.”

A more abstract version of this sign is made in neutral space, with a sharp downward wrist flexion. This version of the sign is glossed as BY FORCE (**Figure 7**). In this sign the constraining agent is less salient or completely absent and the sign refers to actions where a norm should be applied. It is often used with an epistemic value: to ascertain that the described facts are as they should be, or that given qualities or actions are necessary to realize or accomplish a given state of affairs. Another LIS sign directed toward the speaker’s neck expresses the signer’s obligation but with a different hand configuration (bent V). This sign (TO BE CONSTRAINED) expresses an obligation not determined by an agent but by the external events.

Evidence favoring the theory that emotions characterize abstract concept representation (Kousta et al., 2011) can be found not only in the LIS sign TO BELIEVE discussed previously, but also in the sign TO EXPRESS ONESELF (see Supplementary Materials). In this sign, the two hands move up and outward in an arc from the chest toward external space, opening to a spread “5 handshake,” an action resembling the way in which we throw objects out of a container. It might not be necessarily obvious how these two concepts imply emotional components; however, as clarified in the introduction, according to the view proposed by Kousta et al. (2011) and Vigliocco et al. (2014) view all abstract concepts have emotional components, even if in different degrees. Compared to the head, the chest activates more general

metaphors, linked not only to cognitive aspects but to emotional elements as well.

The specific metaphors underlying the signs often reflect cultural differences. For example, in Japanese Sign Language, signs related to thinking are executed in the area surrounding the chest (Wilcox, 2005). In Catalan Sign Language (LSC) ideas can be conceived as having liquid form and the results of learning process can be shown as a liquid contained in the learners’ lower torso (Jarque, 2005).

A variety of signs provide support for the theory according to which abstract terms refer more frequently to situations compared to concrete terms, which refer instead more often to objects and their properties. The three LIS signs in **Figures 8, 9** highlight the importance of situations for concepts etymology and representation: they show that signs used in specific situations develop from signs used in similar situations and could all be glossed with the same English word IMPOSSIBLE. These three signs, however, all have different forms, different origins, and are used in different sentences to express a slightly different meaning.

These three signs are examples of the phenomena of semantic change: signs that are initially grounded can become progressively more abstract and less transparent<sup>3</sup> from an iconic perspective.

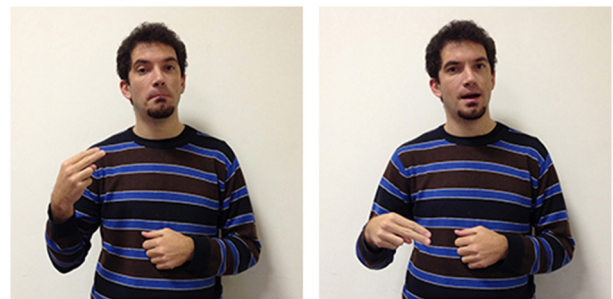
<sup>3</sup>Research on iconicity has traditionally distinguishes between transparent (the meaning can be guessed by everyone), translucent (a non-signer can choose among alternative the right ones, once the meaning is known) and opaque (no iconically motivated link can be found) signs (Bellugi and Klima, 1976; Klima and Bellugi, 1979; Pizzuto and Volterra, 2000; Perniss and Vigliocco, 2014).



**FIGURE 7 | LIS sign BY FORCE.**



**FIGURE 6 | LIS sign TO BELIEVE.**



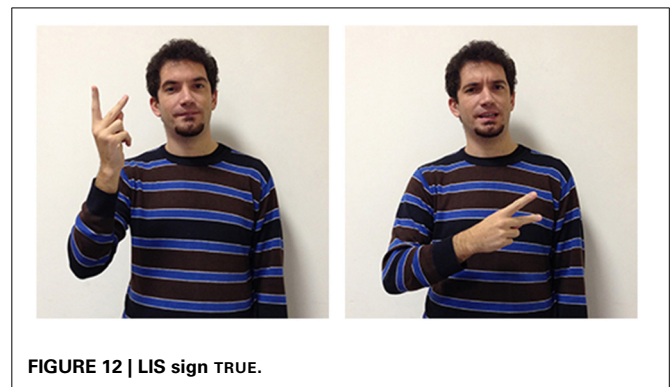
**FIGURE 8 | LIS sign IMPOSSIBLE<sub>H-pa-pa</sub>.**





The sign glossed as IMPOSSIBLE<sub>H-pa-pa</sub><sup>4</sup>, is probably derived from another sign, FORBID, with which it shares the same handshape (extended index and middle fingers) and downward movement. In IMPOSSIBLE<sub>H-pa-pa</sub>, however, the movement is repeated and more rapid. This form has assumed a more general meaning, allowing the signer to express the impossibility of an event or action, due to a decision taken from an authority, to the presence of unfavorable circumstances or to the absence of the necessary conditions for its implementation. The signer would use another sign glossed as IMPOSSIBLE<sub>H-fff</sub> in which the extended fingers move upward in a circular movement to categorically exclude the possibility that the conditions for an event to take place could exist. Wilcox et al. (2010) have proposed an interesting hypothesis on the origin of this LIS sign, which is relevant for us as it supports the idea that abstract words refer to events and situations. The sign IMPOSSIBLE<sub>H-fff</sub> seems to originate with the blessing gesture typical of Christian religion, and is similar to the gesture that has been historically reported to be used by speakers from the South of Italy to refer to a dead or dying person. It is worth noticing that this last variant has been incorporated into LIS as an autonomous lexical unit, i.e., the sign DEAD, produced without the mouth gesture “fff” which is co-produced in IMPOSSIBLE<sub>H-fff</sub>. The conceptual link between the blessing gesture and the sign expressing death is motivated by a metonymic contiguity, since priests are commonly required to bless dead people or people who are going to die. Given that death is associated to the preclusion of the possibility to live, it would have led metaphorically to the emergence of the extreme notion of impossibility expressed through the sign IMPOSSIBLE<sub>H-fff</sub>.

The third LIS sign, IMPOSSIBLE<sub>AA</sub> (Figure 10), has a semantic value that is less specific than the other two signs, as it expresses the notion that the conditions allowing a given action or event are absent, or that something cannot have given characteristics. This two-handed sign derives from the sign POSSIBLE<sub>AA</sub> (Figure 11), in which the signer expresses an evaluation on the existence of actual or potential conditions allowing an action or event. Both IMPOSSIBLE<sub>AA</sub> and POSSIBLE<sub>AA</sub> have the same hand configuration (two fists) but are performed with different movements. In POSSIBLE<sub>AA</sub> the two hands execute simultaneous repeated downward



movements, while in IMPOSSIBLE<sub>AA</sub> the negation of a possibility is expressed through the alternate rotation of the forearms; this negation can be reinforced through a shaking head “no” movement. The close similarity between these two signs, POSSIBLE<sub>AA</sub> and IMPOSSIBLE<sub>AA</sub>, illustrates how similarities and differences in the forms of signs are linked to semantic relations and/or oppositions (see Wilcox et al., 2010; Gianfreda et al., 2014).

A different kind of situational conditioning is found in signs whose forms are influenced by the spoken or written language. For example, the LIS sign TRUE (Figure 12) has a handshape which is also used for the letter V in the manual alphabet (extended index and middle fingers) and adds movement down and to the left of the face. This sign is typically used by signers, either to convey the idea that the described state of affair is true, or in order to clarify that the expressed position is valid.

<sup>4</sup>The letter “H” reported in subscript is conventionally used because this handshape represents the letter H in the manual alphabet. The symbol “pa-pa” refers to the mouth gesture obligatorily requested in the sign execution.

The abstract meaning of “true” and “truth” is thus conveyed in LIS using a strategy known as “initialization.” In sign languages some signs are linked to the corresponding words through the use of a hand configuration which in the manual alphabet (used also in fingerspelling) represents the initial letter of the word having a corresponding meaning. In spoken/written Italian the corresponding words to the English words “true” and “truth” are “vero” and “verità,” both starting with the letter V. Other parameters of the lexical unit, such as movement and location, are not linked to the spoken/written language but are motivated by other factors. While LIS does not distinguish between “true” and “truth,” in ASL the two notions are represented differently. TRUE is represented by using a sign grounded on the straight-path image schema (Roush, 2011), placing the dominant index finger against the signer’s lips and then moving the finger forward several inches using a quick motion. So, the meaning of “true” is represented through the image of an object sent from the mouth along a straight line. In the nominalization form, TRUTH, the sign is slightly varied in that the dominant hand with extended index and middle fingers move in a straight line to make contact with the open palm of the nondominant hand.

These examples help us understand how, in keeping with the WAT theory, the formation of abstract concepts can be influenced by multiple factors, some of which have linguistic origin.

These analyses show that the parameters of the sign’s form can be motivated both by factors internal to the sign language as well as by the signers’ relationship with another language having other characteristics, such as the spoken/written language.

A further example of how forms are influenced by other languages are seen in two other LIS signs. In Italian two different terms are used to distinguish the faculty for language (*linguaggio*) from a specific language used by a community of users (*lingua*) while in English the two concepts are labeled with the same term: “language.”

These concepts are also differentiated by two different signs in LIS: in LANGUAGE/LINGUAGGIO (Figure 13) the hand moves up from the chest toward the external space and opens to a spread 5 handshape (very similar to the sign TO EXPRESS ONESELF); in LANGUAGE/LINGUA (Figure 14) both hands have an handshape associated with the letter “L” in the manual alphabet (extended index finger, thumb extended laterally). The hands, which are initially located in proximity of the mouth, move symmetrically forward with a wrist rotation. The sign LINGUISTICS (Figure 15)

is very similar to the sign LANGUAGE/LINGUA, with the only exception that at the end of the movement the hands close into fists.

A final example is the LIS sign COMMUNICATION. This sign is similar to the ASL sign for the same concept: both hands have a handshape like the letter “C” in the manual alphabet and move forward and backward with a reciprocal alternate movement, possibly reflecting the underlying metaphor that “interaction is exchanging objects” (Roush, 2011). In LIS this sign has undergone interesting changes. In the past the sign was made in front of the mouth; now the sign is executed in the neutral space in front of the signer, perhaps related to a more recent cultural change in the concept resulting in communication not being conceived as being limited to spoken communication, but as also including manual and more general body communication.

All of the examples discussed above are interesting because they combine a strategy based on initialization with a process in which specific body parts (mouth, hand) and movements are involved to constrain and delimit the meaning.

## CONCLUSION

Our analyses and the examples provided are consistent with embodied and grounded theories of cognition, according to which abstract concepts are grounded in perception, action and emotional systems. What we find most important, however, is that sign languages can clarify the different kinds of grounding and thus contribute to the debate about how embodied theories can account for abstractness. We considered and found examples supporting different kinds of embodied theories. The examples



FIGURE 14 | LIS sign LANGUAGE/LINGUA.



FIGURE 13 | LIS sign LANGUAGE/LINGUAGGIO.



FIGURE 15 | LIS sign LINGUISTICS.



we made do not allow us to claim that a given theory is more valid compared to other theories. More systematic analyses would be necessary to advance such a claim. However, we think we are entitled to argue a. that an example can support or not a theory, or more than one theory; b. that, if the theory A is not able to explain a given sign which is rather explained by the theory B, the theory A cannot be considered as exhaustive.

We will discuss below what we consider the most important theoretical implications of the present work.

### DIFFERENT LEVELS OF BODILY INVOLVEMENT

First, our analysis indicates that, even if in sign languages the body is always involved to convey meanings, this involvement occurs at different levels. Skeptics of an embodied cognition perspective might object that it is not completely surprising that sign languages would provide evidence of grounding, given their visual nature and in particular the large amount of iconicity utilized by the language. In sign languages the coupling between language processing and sensori-motor processing becomes indeed more evident than in spoken languages. The body is always involved in spoken languages, for example through vocal articulators but in Sign languages the body, the hands and facial expressions become the main articulators. For example, the hands used for everyday activities such as pointing, enumerating or manipulating objects are also used for representing the same activities.

At the same time, however, it is possible to detect different levels of embodiment through a sign language analysis. The continuity between praxis, gesture and sign is easily recognizable at different levels of SLs structure: formational parameters, lexicon, morphology and syntax (see below for a more detailed discussion of this point). Despite this special characteristic of SLs has been widely recognized (e.g., Sandler and Lillo-Martin, 2006), only a few studies have explored the relationship between sign language and embodied theories, stressing the role of iconicity in sign languages (e.g., Pizzuto and Volterra, 2000; Boyes Braem et al., 2002; Morgan et al., 2008; Perniss et al., 2010). Iconicity can provide an additional mechanism for the grounding of language in sensori-motor systems; in SLs the presence of iconicity is pervasive, as a consequence SLs can be considered a special open window to better understand how language can be grounded. For example, according to Taub's (2001) cognitive-linguistic view, iconicity "is not an objective relationship between image and referent; rather, it is a relationship between our mental models of image and referent." She claims that the creation of an iconic sign involves four successive stages: conceptualization, image selection, schematization, and sign encoding. The choice of the mental image is always mediated by cultural conventions, modality factors and language-specific conventions. This explains also why there is not an "Universal Sign Language" but rather many different Sign Languages. In a recent paper, Perniss and Vigliocco (2014) have highlighted the role of iconicity in both spoken and sign languages considering iconicity as a major vehicle for linking language and human sensory-motor experience. According to their perspective, iconicity represents the key to understand language evolution, development and processing providing a mechanism for displacement, referentiality and embodiment. They have also distinguished different types of iconic mapping, from a form of

iconicity based more on imitative resemblance between the sign and the referent to a form of iconicity requiring more abstract mapping of features.

The novelty of our work, that recognizes the special and more evident role played by iconicity in Sign Languages, consists in focusing not only on the different levels of abstraction of the sign-referent mapping, but in identifying and examining a special case of referents, those of abstract concepts. Analyzing how signs can express abstract concepts in different ways (or through different iconic and not iconic mechanisms) provides some contributions to the debate on how different theories may account for abstract words representation. LIS can indeed provide interesting insights on the different degrees in which the various parameters of the signs are linked to the expressed concepts. In many cases specific locations assume an iconic meaning (for example, the majority of signs for mental activity are performed on the forehead), in other cases also the configuration and/or the movement performed are salient (for example, the sign CLEAR is performed with an open hand configuration moving away from the eyes; a grasping movement characterizes the sign UNDERSTAND) (Pietrandrea, 2002).

### SUPPORT FOR THE DIFFERENT EMBODIED THEORIES OF ABSTRACT CONCEPTS

More crucially to the aim of the present paper, our work provides some insights and has a number of theoretical implications for the debate on how embodied and grounded theories might account for abstract concepts and words (see also Dove, 2009, 2011). The novelty of our work consists in investigating whether signs can provide support for the different embodied theories of abstract concepts.

In line with the previous literature on Conceptual Metaphor Theory, we found that many signs convey a metaphorical meaning and are based on underlying metaphors (e.g., the metaphors of knowing as seeing, of the head as container of mental activities, of the chest as container of feelings and emotions), in keeping with the view that abstract concepts are represented through a metaphorical mapping mechanism. However, in contrast with previous studies we have seen that this is not the whole story, for two main reasons.

The first is that our data support further embodied cognition theories according to which action, situations and emotions are important for abstract concepts representation. Some signs (e.g., the sign for IMPOSSIBLE<sub>H-fff</sub>) provide evidence in favor of the view according to which abstract concepts are grounded on situations; other signs (e.g., the sign TO CONSTRAIN) offer support to the ACE view and other signs (e.g., the sign TO EXPRESS ONESELF) provide evidence favoring the emotion theory of abstract concepts. At a theoretical level the complex framework we obtained cast doubts on the possibility that a single explanation, for example based on a metaphorical mapping mechanism, is valid for the entire domain of abstract concepts and terms (See Prinz, 2002, 2012, for a similar view, according to which different abstract concepts can be explained referring to situation, to metaphors, to action as well as to linguistic information). At the same time, it confirms the necessity to perform fine-grained analyses of the differences between kinds of abstract concepts, analyses

which some authors have started to conduct (e.g., Ghio et al., 2013; Roversi et al., 2013).

The second conclusion we can make is that, even if the analysis on LIS we performed provides support to all the aforementioned theories, at the same time it highlights their limitations. All these theories together are not able to fully account for the whole variety of signs we described. More importantly, they are not able to account for signs expressing some abstract concepts, such as truth.

We think that one of the main contributions of the present work consists in showing that, for some abstract concepts (e.g., the name of a discipline such as “linguistics,” a concept such as “truth,” etc.), LIS exploits linguistic information. This linguistic information could derive from different sources: from the same sign language (e.g., the LIS IMPOSSIBLE<sub>AA</sub> sign derives from the LIS sign POSSIBLE<sub>AA</sub>), from a foreign sign language as ASL (e.g., LANGUAGE/LINGUA and LINGUISTICS) or from spoken/written Italian (e.g., TRUE). This finding challenges many current embodied theories of abstract concepts and clearly supports the WAT view. More generally, it supports multiple representation views according to which not only sensorimotor but also emotional and especially linguistic information, differently distributed, characterize abstract concepts representation (beyond the WAT theory, see also Barsalou et al., 2008; Louwerse, 2011; see Kousta et al., 2011, for a multiple representation view stressing the role of emotions for abstract concepts and Dove, 2014, for a multiple representation view stressing the importance of language, similarly to WAT).

## A METHODOLOGICAL NOTE

Finally, a methodological note. LIS has proved to be an interesting and powerful mean to access how concepts are represented. We hope we have been able to suggest that the study of sign languages represents a fruitful and promising research line to investigate issues crucial for embodied and grounded cognition perspectives, in particular whether different degrees of embodiment exist (Taub, 2001) and whether they vary depending on the domain. Other studies have already demonstrated the importance of the study of sign languages for an embodied and grounded perspective. However, to our knowledge the present study is the first in which examples from a sign language are used to test and validate different theories on abstract concepts. Obviously a certain caution should be used, since, even though they are performed with the body, signs are, like words, arbitrary, so it is difficult to argue that they reflect directly the way concepts are represented. However, they are surely more grounded and to a certain extent more “visible” than words, thus they certainly represent an important cue to help understand conceptual representation. The present paper, being a theoretical paper rather than an experimental one, intends to indicate a possible direction of work. In order to perform a more systematic and thorough analysis, one would need to ask LIS signers to rate different kinds of signs in terms of abstractness, and then select a subset of signs evaluated as abstract and analyze them. Future work is planned to perform such an analysis.

Overall, we think our work provides some hints for how to address issues related to the future of embodied cognition and to the notion of body. Our LIS analyses suggest that, even if the

signs we described always involve the body, different degrees of embodiment might be present. Furthermore, our results suggest that to account for abstract concepts not only sensorimotor and emotional experience should be called into play, but that also linguistic information plays a major role. This might appear in conflict with an embodied approach. We believe it is not, since language is not a disembodied activity but an important part of our total human experience. A challenge for future research is to identify sub-sets of abstract concepts, and to determine whether linguistic information becomes progressively more relevant, the higher the degree of concepts abstractness is.

## ACKNOWLEDGMENTS

We would like to thank Ferdinand Binkofski, Felice Cimatti, Carmen Granito, Claudia Scorolli, and Luca Tummolini for the frequent discussions on abstract concepts, and Paolo Rossini and Alessio Di Renzo for discussion on the LIS examples. Thanks to Giulia Petitta for help with the references and to Luca Lamano and Stefano Marta for help with the pictures. A special thanks to Penny Boyes Braem, to whom we are deeply indebted for a careful revision and for insightful comments on the manuscript. Thanks also to the people of the EMCO lab ([www.emco.unibo.it](http://www.emco.unibo.it)) and of the LaCAM lab ([www.istc.cnr.it/group/lacam](http://www.istc.cnr.it/group/lacam)).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00811/abstract>

## REFERENCES

- Barsalou, L. W. (1999). Perceptual symbol systems. *Brain Behav. Sci.* 22, 577–609.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Barsalou, L. W., Santos, A., Simmons, K. W., and Wilson, C. D. (2008). “Language and simulations in conceptual processing,” in *Symbols, Embodiment and Meaning*, eds M. De Vega, A. M. Glenberg, and A. C. Graesser (Oxford: Oxford University Press), 245–283. doi: 10.1093/acprof:oso/9780199217274.003.0013
- Barsalou, L. W., and Wiemer-Hastings, K. (2005). “Situating abstract concepts,” in *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*, eds D. Pecher and R. Zwaan (New York, NY: Cambridge University Press), 129–163. doi: 10.1017/CBO9780511499968.007
- Bellugi, U., and Klima, E. S. (1976). “Two faces of sign: iconic and abstract,” in *Origins and Evolution of Language and Speech*, eds S. Harnad, D. Hoest, and I. Lancaster (New York, NY: New York Academy of Sciences), 514–538.
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., and Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract words. *J. Cogn. Neurosci.* 17, 905–917. doi: 10.1162/0898929054021102
- Boot, I., and Pecher, D. (2010). Similarity is closeness: metaphorical mapping in a perceptual task. *Q. J. Exp. Psychol.* 63, 942–954. doi: 10.1080/17470210903134351
- Boot, I., and Pecher, D. (2011). Representation of categories. *Exp. Psychol.* 58, 162–170. doi: 10.1027/1618-3169/a000082
- Borghi, A. M. (2014). “Embodied cognition and word acquisition: the challenge of abstract words,” in *Body-Language-Communication: An International Handbook on Multimodality in Human Interaction Handbooks of Linguistics and Communication Science (HSK) 38/2*, eds C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and J. Bressemer (Berlin; Boston, MA: Mouton De Gruyter), 1841–1848.
- Borghi, A. M., and Binkofski, F. (2014). *Words as Social Tools: An Embodied View on Abstract Concepts*. Berlin; New York, NY: Springer.
- Borghi, A. M., and Caruana, F. (in press). “Embodiment theories,” in *International Encyclopedia of the Social and Behavioral Sciences, Section of Cognitive Neuroscience, 2nd Edn.*, eds J. Wright and S. Cappa (Amsterdam; New York, NY: Elsevier).

- Borghi, A. M., and Cimatti, F. (2009). "Words as tools and the problem of abstract words meanings," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, eds N. Taatgen and H. van Rijn (Amsterdam: Cognitive Science Society), 2304–2309.
- Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., and Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words acquisition. *Front. Psychol.* 2:15. doi: 10.3389/fpsyg.2011.00015
- Borghi, A. M., and Pecher, D. (2011). Introduction to the special topic Embodied and Grounded Cognition. *Front. Psychol.* 2:187. doi: 10.3389/fpsyg.2011.00187
- Boroditsky, L. (2011). "How languages construct time," in *Space, Time and Number in the Brain: Searching for the Foundation of Mathematical Thought*, eds S. Dahene and E. M. Brannon (Amsterdam; New York, NY: Elsevier), 333–341.
- Boroditsky, L., and Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychol. Sci.* 13, 185–188. doi: 10.1111/1467-9280.00434
- Boyes Braem, P. (1981). *Feature of the Handshape in American Sign Language*. Doctoral dissertation, Berkeley: University of California.
- Boyes Braem, P., Groeber, S., Stocker, H., and Tissi, K. (2012). *Weblexikon für Fachbegriffe in Deutschschweizerischer Gebärdensprache (DSGS) und Deutsch*. Hamburg: eDITION Fachzeitschrift für Terminologie.
- Boyes Braem, P., Pizzuto, E., and Volterra, V. (2002). "The interpretation of signs by (Hearing and Deaf) members of different cultures," in *Progress in Sign Language Research. In Honor of Siegmund Prillwitz*, eds R. Schulmeister and H. Reinitzer (Hamburg: Signum-Verlag), 187–219.
- Boyes Braem, P., and Sutton Spence, R. (Eds.). (2001). *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Language*. Hamburg: Signum Verlag.
- Brennan, M. (2005). Conjoining word and image in British Sign Language (BSL): an exploration of metaphorical signs in BSL. *Sign. Lang. Stud.* 5, 360–382. doi: 10.1353/sls.2005.0007
- Capirci, O., Contaldo, A., Caselli, M. C., and Volterra, V. (2005). From action to language through gesture: a longitudinal perspective. *Gesture* 5, 155–177. doi: 10.1075/gest.5.1-2.12cap
- Casasanto, D. (2008). Similarity and proximity: when does close in space mean close in mind? *Mem. Cognit.* 36, 1047–1056. doi: 10.3758/MC.36.6.1047
- Casasanto, D., and Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition* 106, 579–593. doi: 10.1016/j.cognition.2007.03.004
- Casasanto, D., Fotakopoulou, O., and Boroditsky, L. (2010). Space and time in the child's mind: evidence for a cross-dimensional asymmetry. *Cogn. Sci.* 34, 387–405. doi: 10.1111/j.1551-6709.2010.01094.x
- Chatterjee, A. (2010). Disembodying cognition. *Lang. Cognit.* 2, 79–116. doi: 10.1515/langcog.2010.004
- Connell, L., and Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition* 125, 452–465. doi: 10.1016/j.cognition.2012.07.010
- Cuxac, C., and Sallandre, M. A. (2007). "Iconicity and arbitrariness in French Sign Language: highly iconic structures, degenerated iconicity and diagrammatic iconicity," in *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, eds E. Pizzuto, P. Pietrandrea, and R. Simone (Berlin; New York, NY: Mouton de Gruyter), 13–33.
- Desai, D. H., Binder, J. R., Conant, L. L., and Seidenberg, M. S. (2010). Activation of sensory-motor areas in sentence comprehension. *Cereb. Cortex* 20, 468–478. doi: 10.1093/cercor/bhp115
- Dove, G. (2009). Beyond perceptual symbols: a call for representational pluralism. *Cognition* 110, 412–431. doi: 10.1016/j.cognition.2008.11.016
- Dove, G. (2011). On the need for embodied and disembodied cognition. *Front. Psychol.* 1:242. doi: 10.3389/fpsyg.2010.00242
- Dove, G. (2014). Thinking in words: language as an embodied medium of thought. *Top. Cogn. Sci.* doi: 10.1111/tops.12102. [Epub ahead of print].
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260
- Fischer, M. H., and Zwaan, R. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 825–850. doi: 10.1080/17470210701623605
- Flusberg, S. J., Thibodeau, P. H., Sternberg, D. A., and Glick, J. J. (2010). A connectionist approach to embodied conceptual metaphor. *Front. Psychol.* 1:197. doi: 10.3389/fpsyg.2010.00197
- Fontana, S. (2008). Mouth actions as gesture in sign language. *Gesture* 8, 104–123. doi: 10.1075/gest.8.1.08fon
- Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Soc. Neurosci.* 3, 317–333. doi: 10.1080/17470910701563608
- Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310
- Gentner, D. (2010). Psychology in cognitive science: 1978–2038. *Top. Cogn. Sci.* 2, 328–344. doi: 10.1111/j.1756-8765.2010.01103.x
- Ghio, M., Vaghi, M. M. S., and Tettamanti, M. (2013). Fine-grained semantic categorization across the abstract and concrete domains. *PLoS ONE* 8:e67090. doi: 10.1371/journal.pone.0067090
- Gianfreda, G. (2011). *Analisi Conversazionale e Indicatori Linguistici Percettivi e Cognitivi nella Lingua dei Segni Italiana (LIS)*. Ph.D. dissertation, University of Macerata.
- Gianfreda, G., Volterra, V., and Zuczkowski, A. (2014). L'espressione dell'incertezza nella Lingua dei Segni Italiana (LIS). *Ricerche di Pedagogia e Didattica* 9, 199–234. doi: 10.6092/issn.1970-2221/4298
- Gibbs, R. W., and Steen, G. J. (1999). *Metaphor in Cognitive Linguistics: Selected Papers from the Fifth International Cognitive Linguistics Conference, Amsterdam, 1997*. London: Benjamins.
- Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565. doi: 10.3758/BF03196313
- Glenberg, A. M., Sato, M., and Cattaneo, L. (2008a). Use-induced motor plasticity affects the processing of abstract and concrete language. *Curr. Biol.* 18, R290–R291. doi: 10.1016/j.cub.2008.02.036
- Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., and Buccino, G. (2008b). Processing abstract language modulates motor system activity. *Q. J. Exp. Psychol.* 61, 905–919. doi: 10.1080/17470210701625550
- Guan, C. Q., Meng, W., Yao, R., and Glenberg, A. M. (2013). Motor system contribution to the comprehension of abstract language. *PLoS ONE* 8:e75183. doi: 10.1371/journal.pone.0075183
- Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335–346.
- Jarque, M. J. (2005). Double mapping in metaphorical expressions of thought and communication in Catalan Sign Language. *Sign. Lang. Stud.* 5, 292–316. doi: 10.1353/sls.2005.0008
- Jirak, D., Menz, M., Buccino, G., Borghi, A. M., and Binkofski, F. (2010). Grasping language. A short story on embodiment. *Conscious. Cogn.* 19, 711–720. doi: 10.1016/j.concog.2010.06.020
- Klima, E. S., and Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.
- Kousta, S. T., Vigliocco, G., Vinson, D., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34. doi: 10.1037/a0021446
- Lai, V. T., and Boroditsky, L. (2013). The immediate and chronic influence of spatio-temporal metaphors on the mental representations of time in English, Mandarin, and Mandarin-English speakers. *Front. Psychol.* 4:142. doi: 10.3389/fpsyg.2013.00142
- Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: Chicago University Press.
- Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Top. Cogn. Sci.* 3, 273–302. doi: 10.1111/j.1756-8765.2010.01106.x
- Louwerse, M. M., and Connell, L. (2011). A taste of words: linguistic context and perceptual simulation predict the modality of words. *Cogn. Sci.* 35, 381–398. doi: 10.1111/j.1551-6709.2010.01157.x
- Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. doi: 10.1016/j.cortex.2010.11.002
- Meurant, L., Sinte, A., Van Herreweghe, M., and Vermeerbergen, M. (2013). *Sign Language Research, Uses and Practices. Crossing Views on Theoretical and Applied Sign Language Linguistics*. Berlin: Mouton-De Gruyter and Ishara Press. doi: 10.1515/9781614511472
- Morgan, G., Herman, R., Barriere, I., and Woll, B. (2008). The onset and mastery of spatial language in children acquiring British Sign Language. *Cogn. Dev.* 23, 1–19. doi: 10.1016/j.cogdev.2007.09.003
- Nelson, D. L., and Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *J. Mem. Lang.* 31, 237–260. doi: 10.1016/0749-596X(92)90013-N
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York, NY: Oxford University Press.

- Pecher, D., and Boot, I. (2011). Numbers in space: differences between concrete and abstract situations. *Front. Psychol.* 2:121. doi: 10.3389/fpsyg.2011.00121
- Pecher, D., Boot, I., and van Dantzig, S. (2011). "Abstract concepts: sensory-motor grounding, metaphors, and beyond," in *The Psychology of Learning and Motivation*, Vol. 54, ed B. Ross (Burlington, VT: Academic Press), 217–248.
- Perniss, P., Thompson, T., and Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Front. Psychol.* 1:227. doi: 10.3389/fpsyg.2010.00227
- Perniss, P., and Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philos. Trans. R. Soc. B.*
- Petitta, G., Di Renzo, A., Chiari, I., and Rossini, P. (2013). "Sign language representation: new approaches to the study of face-to-face and written Italian Sign Language (LIS)," in *Sign Language Research, Uses and Practices*, ed Meurant et al. (Berlin: Mouton de Gruyter), 137–158.
- Pietrandrea, P. (2002). Iconicity and arbitrariness in Italian Sign Language. *Sign. Lang. Stud.* 2, 296–321. doi: 10.1353/sls.2002.0012
- Pietrandrea, P., and Russo, T. (2007). "Diagrammatic and imagic hypoicons in signed and verbal languages," in *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, eds E. Pizzuto, P. Pietrandrea, and R. Simone (Berlin; New York, NY: Mouton de Gruyter), 35–56.
- Pizzuto, E., Cameracanna, E., Corazza, S., and Volterra, V. (1995). "Terms for spatio-temporal relations in Italian Sign Language (LIS)," in *Iconicity in Language*, ed R. Simone (Amsterdam: John Benjamins), 237–256. doi: 10.1075/cilt.110.17piz
- Pizzuto, E., and Corazza, S. (1996). Noun morphology in Italian Sign language (LIS). *Lingua* 98, 169–196. doi: 10.1016/0024-3841(95)00037-2
- Pizzuto, E., and Pietrandrea, P. (2001). The notation of signed texts: open questions and indications for further research. *Sign Lang. Linguist.* 4, 29–43. doi: 10.1075/sll.4.12.05piz
- Pizzuto, E., Pietrandrea, P., and Simone, R. (2007). "Introduction," in *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, eds E. Pizzuto, P. Pietrandrea, and R. Simone (Berlin: Mouton de Gruyter), 1–10.
- Pizzuto, E., and Volterra, V. (2000). "Iconicity and transparency in sign languages: a cross-linguistic cross-cultural view," in *The Signs of Language Revisited: An Anthology in Honor of Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. Lane (Hillsdale, NJ: Erlbaum), 261–286.
- Prinz, J. J. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA: MIT Press.
- Prinz, J. J. (2012). *Beyond human nature. How culture and experience shape our lives*. London; New York, NY: Penguin; Norton.
- Roush, D. R. (2011). Language between bodies: a cognitive approach to understanding linguistic politeness in American Sign Language. *Sign. Lang. Stud.* 1, 329–374. doi: 10.1353/sls.2011.0000
- Roversi, C., Borghi, A. M., and Tummolini, L. (2013). A marriage is an artefact and not a walk that we take together: an experimental study on the categorization of artefacts. *Rev. Philos. Psychol.* 4, 527–542. doi: 10.1007/s13164-013-0150-7
- Rüschmeyer, S. A., Brass, M., and Friederici, A. D. (2007). Comprehending pre-hending: neural correlates of processing verbs with motor stems source. *J. Cogn. Neurosci.* 19, 855–865. doi: 10.1162/jocn.2007.19.5.855
- Russo, T. (2004). *La mappa Poggiata Sull'isola. Iconicità e metafora nelle lingue dei segni e nelle lingue vocali*. Rende: Centro Editoriale e Libreria Università della Calabria.
- Russo, T. (2005). A crosslinguistic, cross-cultural analysis of metaphors in two Italian Sign Language (LIS) registers. *Sign. Lang. Stud.* 5, 333–359. doi: 10.1353/sls.2005.0009
- Sabsevitz, D. S., Medler, D. A., Seidenberg, M., and Binder, J. R. (2005). Modulation of the semantic system by word imageability. *Neuroimage* 27, 188–200. doi: 10.1016/j.neuroimage.2005.04.012
- Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., and Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Front. Hum. Neurosci.* 7:125. doi: 10.3389/fnhum.2013.00125
- Sandler, W., and Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139163910
- Scorolli, C., Binkofski, F., Buccino, G., Nicoletti, R., Riggio, L., and Borghi, A. M. (2011). Abstract and concrete sentences, embodiment, and languages. *Front. Psychol.* 2:227. doi: 10.3389/fpsyg.2011.00227
- Scorolli, C., Jacquet, P., Binkofski, F., Nicoletti, R., Tessari, A., and Borghi, A. M. (2012). Abstract and concrete phrases processing differently modulates cortico-spinal excitability. *Brain Res.* 1488, 60–71. doi: 10.1016/j.brainres.2012.10.004
- Simmons, W. K., Hamann, S. B., Harenski, C. L., Hu, X. P., and Barsalou, L. W. (2008). fMRI evidence for word association and situated simulation in conceptual processing. *J. Physiol. Paris* 102, 106–119. doi: 10.1016/j.jphysparis.2008.03.014
- Stokoe, W. C. (1960). *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. Studies in Linguistics*. Occasional Paper 8, Department of Anthropology and Linguistics, University of Buffalo (Rev. ed Linstok press, Silver spring, MD, 1978).
- Sutton, V. (1999). *Lessons in SignWriting. Textbook and workbook, 2nd Edn., 1st Edn., 1995*. La Jolla, CA: Deaf Action Committee for Sign Writing.
- Sutton-Spence, R. (2005). *Analysing Sign Language Poetry*. Basingstoke: Palgrave.
- Taub, S. F. (2001). *Language from the Body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511509629
- Tomasino, B., and Rumiati, R. I. (2013). Special topic on: what does neuropsychology say about the role of sensorimotor processes in conceptual knowledge and abstract concepts. *Front. Hum. Neurosci.* 7:498. doi: 10.3389/fnhum.2013.00498
- Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025
- Volterra, V. (1987). *LIS. Lingua Italiana dei Segni*. Bologna: Il Mulino.
- Wauters, L. N., Tellings, A. E. J. M., Van Bon, W. H. J., and Van Haften, A. W. (2003). Mode of acquisition of word meanings: the viability of a theoretical construct. *Appl. Psychol.* 24, 385–406. doi: 10.1017/S0142716403000201
- West, D., and Sutton-Spence, R. (2010). Shared thinking processes with four deaf poets: a window on "the creative" in "Creative Sign Language." *Sign. Lang. Stud.* 12, 188–210. doi: 10.1353/sls.2011.0023
- Wiemer-Hastings, K., Krug, J., and Xu, X. (2001). "Imagery, context availability, contextual constraints and abstractness," in *Proceedings of 23rd Annual Meeting of the Cognitive Science Society* (Hillsdale, NJ: Erlbaum), 1106–1111.
- Wilcox, P. (1996). "Deontic and epistemic modals in ASL: a discourse analysis," in *Conceptual Structure, Discourse and Language*, ed A. Goldberg (Cambridge: Cambridge University Press), 481–492.
- Wilcox, P. P. (2005). What do you think? Metaphor in thought and communication domains in American Sign Language. *Sign. Lang. Stud.* 5, 267–291. doi: 10.1353/sls.2005.0011
- Wilcox, P. P. (2007). "Constructs of the mind: cross-linguistic contrast of metaphor in verbal and signed languages," in *Verbal and Signed Languages: Comparing Structures, Constructs and Methodologies*, eds E. Pizzuto, P. Pietrandrea, and R. Simone (Berlin; New York, NY: Mouton de Gruyter), 251–272.
- Wilcox, S., Rossini, P., and Antinoro Pizzuto, E. (2010). "Grammaticalization in sign languages," in *Sign Languages*, ed D. Brentari (Cambridge, UK: Cambridge University Press), 332–354. doi: 10.1017/CBO9780511712203.016
- Wilcox, S., and Shaffer, B. (2006). "Modality in american sign language," in *The Expression of Modality*, ed W. Frawley (Berlin: Mouton de Gruyter), 207–237.
- Wilcox, S., and Wilcox, P. P. (1995). "The gestural expression of modality in ASL," in *Modality in Grammar and Discourse*, eds J. Bybee and S. Fleischman (Amsterdam/Philadelphia: John Benjamins), 135–162. doi: 10.1075/tsl.32.07wil
- Wilson, M. (2002). Six views on embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322

**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.

Received: 31 January 2014; accepted: 08 July 2014; published online: 29 July 2014.

Citation: Borghi AM, Capirci O, Gianfreda G and Volterra V (2014) The body and the fading away of abstract concepts and words: a sign language analysis. *Front. Psychol.* 5:811. doi: 10.3389/fpsyg.2014.00811

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Borghi, Capirci, Gianfreda and Volterra. 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.





# Sensory motor mechanisms unify psychology: the embodiment of culture

Tamer Soliman, Alison Gibson and Arthur M. Glenberg\*

Department of Psychology, Arizona State University, Tempe, AZ, USA

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Andrew D. Wilson, Leeds Metropolitan University, UK  
Tomás A. Palma, University of Lisbon, Portugal  
Margarida Vaz Garrido, ISCTE-Instituto Universitário de Lisboa, Portugal

## \*Correspondence:

Arthur M. Glenberg, Department of Psychology, Arizona State University, 950 S. McAllister, Mail Code 1104, Tempe, AZ 85287, USA  
e-mail: glenberga@asu.edu

Sensorimotor mechanisms can unify explanations at cognitive, social, and cultural levels. As an example, we review how anticipated motor effort is used by individuals and groups to judge distance: the greater the anticipated effort the greater the perceived distance. Anticipated motor effort can also be used to understand cultural differences. People with interdependent self-construals interact almost exclusively with in-group members, and hence there is little opportunity to tune their sensorimotor systems for interaction with out-group members. The result is that interactions with out-group members are expected to be difficult and out-group members are perceived as literally more distant. In two experiments we show (a) interdependent Americans, compared to independent Americans, see American confederates (in-group) as closer; (b) interdependent Arabs, compared to independent Arabs, perceive Arab confederates (in-group) as closer, whereas interdependent Americans perceive Arab confederates (out-group) as farther. These results demonstrate how the same embodied mechanism can seamlessly contribute to explanations at the cognitive, social, and cultural levels.

**Keywords:** embodied cognition, distance perception, motor effort, culture, self-construal

## TOWARD A UNIFIED PSYCHOLOGY

Academic psychology compartmentalized the mind into cognitive, social and cultural partitions, and developed for each a self-delimited conceptual paradigm and explanatory tradition. Typically, the cognitive, social, and cultural psychologists believe that they target three different mental structures in the minds of the same people. For the first, the research participants are computer-like information processors (e.g., Newell, 1980), for the second, they are social agents driven by basic motivations to fulfill interpersonal goals (Forgas et al., 2005), and, for the third, they are normative populations immersed each in their local system of values, beliefs, and worldviews (e.g., Shweder, 1996). With these disparate levels of construct specification, cross-talk over the epistemological fence is limited (e.g., Messick and Mackie, 1989; Hong et al., 2000; Nisbett, 2003; Knoblich and Sebanz, 2006). In these accounts, the levels of the mind may, at best, interact, but remain conceptually intact, much like billiard-balls that maintain their self-contained identities through their collisions.

Our goal is to take steps toward a unified account of the human mind by finding theoretical units of analysis that apply equally to understanding the cognitive, social, and cultural aspects of behavior. Alongside others (e.g., Schubert and Semin, 2009; Glenberg, 2010), we believe that the body has this unification potential; its sensorimotor mechanisms can explain behavior that plays out in a physical, social, or cultural context. Our strategy is to use the bodily level of description to side-step the three different characterizations of the mind found in the three sub-disciplines, and thereby demonstrate the possibility of specifying level-neutral mechanisms that could uniformly explain cognitive, social, and cultural behavior.

Specifically, the empirical plan is to identify a sensorimotor mechanism with proven explanatory power at one of these levels, then to examine whether this same mechanism can predict behavioral patterns that are well established at the other two levels. Fortunately, one such mechanism has already been characterized and its cognitive (Proffitt, 2006) and social (Schnall et al., 2008) effects successfully demonstrated. After reviewing these, we present our own account to explain how the mechanism can generate plausible predictions in a cultural context, then we report on two studies that generally confirmed our predictions.

## MOTOR EFFORT IN BASIC COGNITIVE AND SOCIAL PROCESSES

In cognitive psychology, Proffitt and his associates forged a link between two characteristics of the motor system, and they used this link to propose a novel reformulation of the mechanism of distance perception (Proffitt, 2006). On Proffitt's account, the voluntary muscle system is sensitive to the bioenergetic status of the body (Davis et al., 1997; Achten et al., 2004; Coyle, 2004) while being simultaneously tightly coupled with the visual system (Hommel et al., 2001). On the basis of this link, it was proposed that distance perception could not only be conceptualized as an algorithmic process determined exclusively by visual cues (e.g., Cutting and Vishton, 1995), but that it is an ecological integrative process in which the motor system plays an important role. Specifically, Proffitt predicts that visual perception of distance to a target should be scaled by the motor effort required to interact with (e.g., walk up to) that target.

The hypothesized effect of motor effort was confirmed (see Proffitt, 2006, for a review): participants reported inflated visual distance to targets that required more motor effort to reach.

Participants who were wearing a backpack, exhausted, in poor fitness, elderly, or in ill-health reported hills to appear steeper when compared with their fit, healthy, younger, or rested counterparts.

Could the same sensorimotor mechanism extend to the realm of the social? Schnall and her colleagues (Schnall et al., 2008) argued that a supportive other is construed by the body as a potential resource, either providing a surplus of energy or easing the burden on internal resources. Thus, social support enhances the efficacy of the individual's motor and cognitive systems during task performance. To test this hypothesis, a group of solo participants and another group accompanied by their friends were asked to estimate the slope of hills. Others were asked to imagine the presence of a friend, a neutral individual, or a disliked person before offering their estimates. In both experiments, the real or imagined presence of a (potentially) supportive other led to smaller estimates of the hill's slope.

### **MOTOR EFFORT AND CULTURAL ORIENTATIONS OF INDEPENDENCE AND INTERDEPENDENCE**

To push the explanatory and predictive power of this mechanism into the domain of cultural behavior, we brought together findings from several lines of research. The first is that the motor system is involved in interpersonal interactions. Of course the motor system is needed to talk, to observe (e.g., move the eyes), to move toward, and to cooperate in physical tasks. But additionally, the motor system is used to help recognize the goals of others using an automatic resonance process based on their movements (Rizzolatti et al., 2001; Blakemore and Frith, 2005; Wilson and Knoblich, 2005). When person A observes person B act, A's motor repertoire (predominately in premotor cortex) is automatically activated, or resonates, and provides a model of what B is doing. When successful, this resonance generates A's goals when engaging in this action, and A uses these goals as an understanding of B's goals. Note that this simulation, or resonance, is not in anticipation of the B's actions, but close to simultaneous with those actions.

Conversely, when the relevant motor program does not exist in A's repertoire, or when it cannot be fluently implemented, then A's recognition of the B's behavior is not as fluent, perhaps because of the more energetic investment the motor system requires to simulate the perceived action (Calvo-Merino et al., 2005; Casile and Giese, 2006; Petroni et al., 2010). For example, perception of familiar actions that the participant can fluently reproduce is accompanied by a reduced BOLD signal at the motor cortex, which is an index of low energy demands for simulating the action (Tanaka et al., 2001; Muhlau et al., 2005).

Differences in communication style across cultures are one important source of this familiarity effect on automatic motor resonance as demonstrated using fMRI to investigate modulatory effects of the perceived cultural membership on the activity of the putative human homolog of the mirror neuron system (MNS). Liew et al. (2011), for example, documented a higher BOLD signal at the MNS sites of mainland Chinese participants when watching American communicative hand gestures that were unfamiliar (e.g., "quail") vs. familiar expressive American gestures (e.g., thumbs up). Caucasian Americans watching a Nicaraguan actor

modeling either native Nicaraguan or American communicative hand gestures showed signs of more effortful motor resonance for the former (whereas the effect disappeared when the model was American; Molnar-Szakacs et al., 2007). White and black female (but not male) participants showed signs of more effortful motor resonance with the same simple finger movements when modeled by the other race compared to the same race (Desy and Theoret, 2007).

Thus, we propose that interactions with cultural out-group members are expected to be more effortful than interactions with in-group members. Note that out-group members potentially differ from in-group members along many communicative dimensions (Archer, 1997). For example, Russians may point with the middle finger, not the index finger. Also, although facial expression of emotions is qualitatively universal, differences in the rules of display (e.g., of intensity) may be misleading in cross-cultural encounters (Ekman et al., 1987; Matsumoto et al., 2002). Relative to Westerners, for example, the Japanese tend to mask both negative (Ekman, 1972) and positive (Matsumoto and Kuppertsbusch, 2001) emotional expressions. Consequently, they rely (more so than Westerners) on vocally conveyed emotional tone when inferring underlying emotional states (Tanaka et al., 2010). And of course, accented pronunciation by those speaking a second language often differs significantly from native pronunciation in both segmental (place and method of articulation, e.g., Gatbonton, 1975) and supra-segmental (stress, rhythm, and intonation, e.g., Fokes and Bond, 1984) characteristics, and is actually perceived as less intelligible by native speakers (Flege, 1988). All these differences are taxing for the motor system as it attempts to resonate with observed actions [including resonance with articulatory actions, as demonstrated by Fadiga et al. (2002)].

Importantly, the more costly effort of cross-cultural encounters relative to within-cultural encounters is not only experienced during interaction, but also shapes the default expectation of interaction with out- vs. in-group members. Consistent with this assumption, a meta-analysis of the social-projection literature shows that projecting one's own state is stronger onto in-group than onto out-group others, specifically due to the perception of higher self-other similarity with the former group (Robbins and Krueger, 2005). That is, even when the interaction has not yet started, people have an expectation of less (sensorimotor, communicative) similarity with an out-group member than with an in-group member (which we will use later to justify the design of our experiments).

In addition to the proposed main effect of group-membership on the expected effort of interaction, a moderating effect needs to be added. Cross-cultural psychology suggests that people may develop an interdependent cultural orientation that stresses relatedness and harmony with their in-groups, or an independent one that emphasizes the uniqueness of their individual selves (Markus and Kitayama, 1991). People with interdependent self-construals tend to live in societies with fairly homogenous ethnic composition (e.g., East Asia), and exhibit lower levels of mobility within these settings, whereas independents typically live in ethnically diverse populations (e.g., North America) and are much more mobile relative to their interdependent counterparts (Triandis

et al., 1988; Oishi and Kisling, 2009; Oishi, 2010; Schug et al., 2010). It is important to note, however, that cross-cultural psychology is moving away from identifying these categorical and geographical cultural differences in social orientation (e.g., all East Asians or all North Americans) to acknowledging that both interdependent and independent self-construals can be found, to a greater or lesser extent, around the world.

Thus, we come to the major hypothesis that drives our empirical work, namely the cultural motor-effort hypothesis. First, we suppose that cultures, and the self-construals they engender, should be conceived more as a continuum than as categories. Thus, what we describe next for interdependent and independent self-construals should be considered the ends of the continuum. Second, people who live in a predominately collectivist culture (and develop interdependent self-construals) tend to interact with family, friends, and an in-group consisting of ethnically and culturally similar people. Consequently, the motor system is strongly tuned to resonate to the behaviors of the in-group, and interaction with the in-group is smooth and relatively effortless. However, for two reasons, these interdependents are at a disadvantage when it comes to interacting with members of the out-group. Because they have little experience with out-group members, they have had little opportunity to tune their motor systems to the behaviors of out-group members. Also, because of the strong tuning or specialization for the in-group, their motor systems will have even more difficulty resonating to the different accents, gestures, etc. of the out-group than a non-tuned system. [We see this as analogous to the development of speech perception. Before an infant is strongly tuned to its native language, it can perceive phonetic distinctions that are not incorporated into the native language (e.g., Kuhl et al., 1992; Aslin et al., 1998). However, once the infant has had considerable experience with the native language, the ability to perceive non-native distinctions is lost]. Thus, interdependents experience a costly demand for motor control and prediction during cross-cultural episodes of interaction.

Third, people who live in a predominately individualistic society are forced to interact with a diversity of others. Although not as strongly tuned as interdependents to interactions with the in-group, interactions with out-group members allow these people to develop moderate skill to process and respond to people with different accents, different communicative gestures and postures, and so on. Thus, in contrast with interdependents, for independents interactions with out-group members are literally less effortful.

This hypothesis predicts that (a) interdependents anticipate motor effort upon the prospect of interacting with out-group members. This, in turn, modulates their subjective visual experience of the distance to out-group members such that their estimates of distance are inflated relative to estimates of distance to in-group members. (b) People with independent orientations should show a smaller difference in estimated distance to in-group and out-group members; they anticipate much less differential effort to interact with out-group individuals owing to the diversity of their motor social repertoire acquired by immersion in ethnically diverse settings.

Experiment 1 provides an initial, cost-effective test of the cultural motor-effort hypothesis, albeit without sampling multiple cultures. The hypothesis suggests that within any culture, those who are more interdependent will resonate more strongly with in-group members relative to those who are more independent. Thus, we predict that relative to independents, interdependents will see in-group members as closer.

The complex literature relating self-construal to prejudice (cited by a reviewer of a previous version of this article) suggests a different prediction. Some research suggests that individualism increases prejudice (e.g., Biernat et al., 1996; Katz and Hass, 1988; Sears and Henry, 2005), and a few studies (e.g., Kleugel, 1990) suggest that within a collectivist culture there is a tendency toward lower prejudice and higher tolerance toward the out-group. If prejudice can be related to motor effort, then one might expect that interdependents (from collectivist cultures) would see out-group members as closer. However, our results suggest the opposite, and so we frame those results in terms of the cultural motor-effort hypothesis.

## STUDY 1: DISTANCE TO AMERICANS AS PERCEIVED BY AMERICANS

### DESIGN AND PROCEDURE

American participants ( $n = 33$ ) were first trained on estimating distances to a human target in terms of seconds needed to walk to the target. Besides inducing a motor-oriented perception of distance, using seconds also minimized any potential effect of the culture-specific distance measurement units (e.g., feet vs. meters) on the reporting of perceived distance to the target, a tack that was especially important in the second study, and used here to maintain the use of a uniform DV across the experiments. The training comprised three trials. In each, the participant estimated the time to walk to the experimenter, then actually walked up to her, and finally received feedback on accuracy of the initial estimate. The training distances in this stage were quasi-randomly selected by the experimenter.

Immediately after training, but in a different location, the participant made 36 distance estimates (three 12-trial blocks) to two Caucasian (i.e., in-group) confederates<sup>1</sup>. The confederates stood at marks along two (imaginary) axes that intersected where the participant stood to make the estimates. The marks on each of the axes were pre-set to be at six different distances from the intersection: the short-distance marks were at 6.77 and 8.77 m from the participant's location at the intersection, the medium distances were at 10.43 and 12.43 m, and the long distances were at 20.43 and 22.43 m. The use of two distances for each of the distance ranges was meant to discourage participants from copying earlier estimates in later trials.

On any given trial, the experimenter asked the participant to turn away from both axes, one of the two confederates would position herself at a mark, then the experimenter signaled to

<sup>1</sup>The three blocks corresponded to three types of barriers behind which the participants stood: a physical barrier (a fence), a symbolic barrier (a caution tape), and no barrier. This independent variable was included to test another hypothesis. Because the effect of barrier was not significant and did not interact with other variables, it will not be discussed further.

the participant to face the confederate. The participant was then immediately asked to estimate, in seconds, the time it would take her to walk to the confederate (half of the trials), or given a 2.5 foot-long stick and asked to estimate the number of sticks it would take her to touch the confederate. On the next trial, the same process repeated, except that the other confederate would position herself on another mark on the other axis. The assignment of the two confederates to the two axes, and the order of distance presentation (i.e., short, medium, or long) were independently counterbalanced within blocks and across participants. Finally, after completing their distance estimates, the participants filled out the Interdependence and Independence subscales of the Self-Construal Survey (SCS) (Singelis, 1994).

## ANALYSIS AND RESULTS

As expected, for these American participants the mean score on the independence subscale ( $M = 5.12$ ,  $SD = 0.73$ ) was greater than the mean score on the interdependence subscale ( $M = 4.60$ ,  $SD = 0.84$ ),  $t_{(28)} = 2.51$ ,  $p = 0.02$ . We used multi-level modeling (MLM) with maximum likelihood estimates of the parameters to take advantage of (a) the continuous nature of the six distances and the measure of cultural orientation, and (b) to obviate potential problems with the sphericity assumption. MLM is similar to regression in that it estimates regression parameters, however, maximum likelihood is used as the estimation procedure and estimated along with each parameter is its own standard error. Thus, the test for statistical significance is a simple  $t$ -test of the parameter divided by the standard error, although the degrees of freedom are often fractional because of the use of Welch-Satterthwaite estimates.

Separate MLMs were run for the two estimates of distance, namely number of seconds to walk and number of sticks. Four participants were dropped from the analysis of number of sticks, one for providing stick estimates more than 3  $SD$  below the mean and three for providing stick estimates more than 3  $SD$  above the mean.

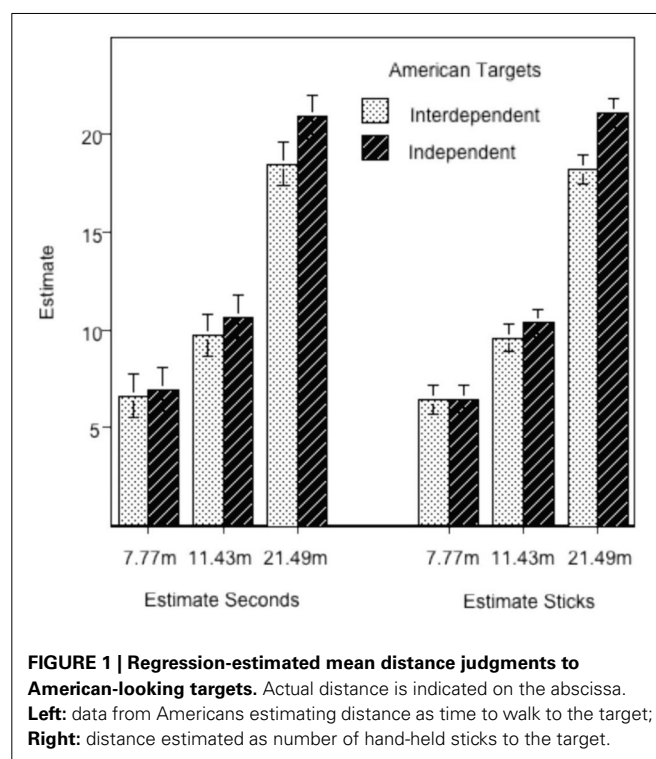
The participants' cultural orientation scores were computed as the ratio of their responses to the interdependent and independent subscales of the SCS (Int:Ind). As is recommended for regression analyses that involve interaction terms (Aiken and West, 1991), all of the independent variables were centered around their respective means.

**Table 1** contains the important results from the MLMs, and **Figure 1** plots the regression-estimated marginal means for the Time estimates in seconds (on the left) and Sticks (on the right) as a function of the actual distance. For both dependent variables, the effect of Distance was significant. (For Time, the parameter value of 0.858 indicates that the estimate grew by 0.858 s for each one meter increase in actual distance; likewise for Sticks, the parameter of 0.874 indicates an increase of 0.874 sticks for each meter of distance).

More importantly, our predictions were confirmed in the form of significant interactions of cultural orientation (Int:Ind) and Distance for both the Time estimate and the Sticks estimate. Rather than arbitrarily breaking the sample into those with interdependent and independent self-construals and losing the statistical power inherent in the continuous variable, we used the

**Table 1 | Parameter estimates (in seconds, upper panel) or stick-number estimates (lower panel) to walk to or touch American confederates.**

Factor	Parameter	Seconds		
		df	t	p
Int:ind	−3.000	29	−0.77	0.45
Distance	0.858	493	37.07	0.001
Int:ind × distance	−0.384	493	−3.00	0.003
Sticks				
Int:ind	−3.09	29	−1.23	0.23
Distance	0.874	493.02	40.9	0.001
Int:ind × distance	−0.47	493.02	−4.41	0.001



**FIGURE 1 | Regression-estimated mean distance judgments to American-looking targets.** Actual distance is indicated on the abscissa. **Left:** data from Americans estimating distance as time to walk to the target; **Right:** distance estimated as number of hand-held sticks to the target.

regression parameters to estimate means for interdependents and independents. The estimates for interdependents were obtained by using a value for the Int:Ind ratio 1  $SD$  above the mean Int:Ind ratio (Aiken and West, 1991). Likewise, the values for independents were obtained by using a value of the Int:Ind ratio 1  $SD$  below the mean of the Int:Ind ratio.

Turning to **Figure 1**, the statistical interaction becomes evident: interdependents, compared to independents, judge distance to in-group confederates as smaller. Furthermore, the difference between interdependents and independents grows with actual distance. This finding is consistent with our cultural-effort hypothesis. Namely, interdependents, compared to independents, spend more time interacting with their in-group and tuning their motor system toward those interactions. Then, because expected



motor effort contributes to distance estimation (Proffitt, 2006), interdependents judge distance as smaller than independents.

The statistical interaction (that the difference in judged distance between interdependents and independents increased with distance) is even more important than the main effect for demonstrating that the groups were using different measurement scales (Proffitt and Linkenauger, 2013). That is, when the unit of measurement used by one group (e.g., X amount of anticipated effort) is different from the unit used by the other group (e.g., 3X amount of anticipated effort), then the difference in the groups' estimates becomes larger with increased distance (an interaction). For example, suppose that Person A measures distance in feet, and Person B measures distance in yards. At a distance of one yard, the two measurements, 3 (feet) and 1 (yard), differ by 2. But at a distance of 5 yards, the two measures, 15 (feet) and 5 (yards), differ by 10. Thus, the interaction is strong evidence that the interdependents and independents are measuring distance using different scales, namely different amounts of expected effort. Nonetheless, it is important to demonstrate that this interaction is replicable, and that is one purpose of the next study.

## STUDY 2: DISTANCE TO ARABS AS PERCEIVED BY ARABS AND AMERICANS

### DESIGN AND PROCEDURES

Clearly, our novel findings in Study 1 need to be replicated and the cultural-effort hypothesis subjected to further test. In Study 2, we used Arab-looking confederates as targets: the confederates were chosen to have dark skin tone, and one of them wore a headscarf, or hijab. Furthermore, we sampled both Arab ( $n = 16$ )<sup>2</sup> and American ( $n = 42$ ) participants. All other aspects of the design and procedures were identical to those of the first study, except that the participants were asked to report their estimates only in terms of time (i.e., number of seconds) to walk up to the confederates.

We predicted that the effect of cultural orientation (Int:Ind) on the American participants' estimates would flip in direction relative to the effect in Study 1. That is, since the confederates were Arab-looking, and hence, out-group members, the interdependent Americans would overestimate the distance relative to the independent Americans. Because the interdependents have tuned their motor systems to interact with other Americans, they should expect greater effort in interacting with the Arab-looking confederates than the independent Americans who have a more broadly tuned motor system. In contrast, we predicted that the Arab participants' estimates to their in-group looking confederates would resemble that of the American participants in Study 1. Interdependent Arabs have a motor system finely tuned for interaction with their in-groups, and thus they should report smaller distance to the targets than the more broadly tuned independent Arabs.

<sup>2</sup>We had hoped to include a larger sample of Arab participants. Unfortunately given current political realities, many Arab students were not willing to participate in psychological research.

## ANALYSIS AND RESULTS

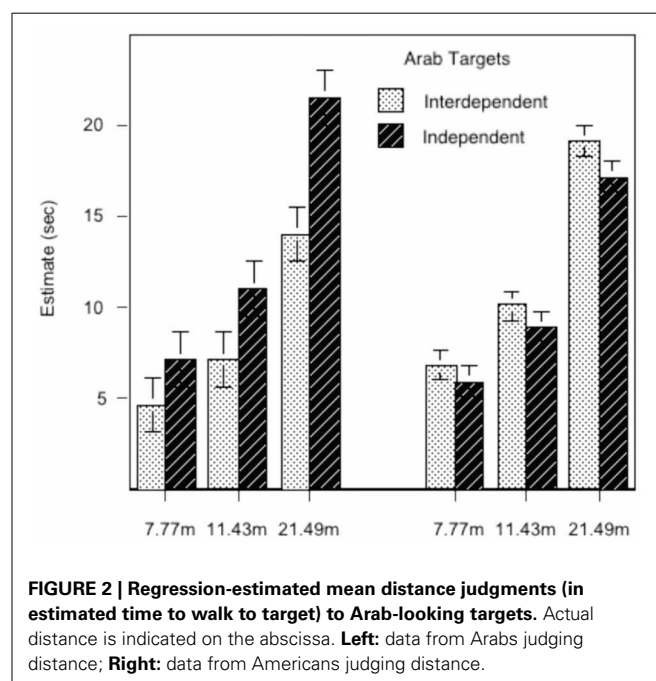
The data from three Arab participants were dropped for procedural errors, and the data from one American were dropped for providing an Int:Ind ratio more than 3 *SD* above the mean. As expected, the mean Int:Ind ratio was significantly higher for Arabs ( $M = 1.1$ ,  $SD = 0.17$ ) than for Americans ( $M = 0.98$ ,  $SD = 0.14$ ),  $t_{(52)} = 2.61$ ,  $p = 0.012$ .

An MLM analysis was run to examine the main effects and interactions of Distance, Int:Ind, and National culture (Arab, American). The results are reported in Table 2.

Figure 2 plots the regression-estimated marginal means of the participants' walking-time estimates as a function of the real distance. The predicted pattern of results was successfully obtained in the form of two interactions. First, there was an interaction of Culture (Arab vs. American) and Int:Ind on distance estimates. For the Arabs (bars on the left), the in-group (i.e., Arab-looking) confederates were perceived as closer by the interdependent than by the independent participants (distance in seconds estimated, respectively, at 1 *SD* above and below the mean Arab Int:Ind ratio). For the American participants, the

**Table 2 | Parameter estimates (in seconds) to walk to Arab-looking confederates.**

Factor	Parameter	df	t	p
Int:ind	0.688	53.99	0.19	0.74
Culture	-0.43	53.99	-0.33	0.85
Distance	0.813	1884.99	74.18	0.001
Culture × int:ind	18.6	53.98	2.42	0.02
Culture × distance	-0.068	1884.99	-2.55	0.01
Int:ind × distance	-0.018	1884.99	-0.25	0.80
Culture × int:ind × distance	1.236	1884.99	7.87	0.001



**FIGURE 2 | Regression-estimated mean distance judgments (in estimated time to walk to target) to Arab-looking targets.** Actual distance is indicated on the abscissa. **Left:** data from Arabs judging distance; **Right:** data from Americans judging distance.

pattern flips: the out-group (i.e., Arab-looking) confederates were perceived as farther by the interdependent than the independent subgroups (estimated at 1 *SD* above and below the mean American Int:Ind ratio). Second, this interaction was modified by actual distance to the confederate such that increasing the distance increased the size of the two-factor interaction. As with the first study, this interaction strongly implies the use of different measurement scales (e.g., expected amount of effort) associated with cultural differences.

## DISCUSSION

Contemporary psychology continues to be composed of diverse discourse communities that do not make substantial connection with the discipline as a whole. These diverse communities of psychologists, which have proliferated in rapid succession, increasingly work under different, often conflicting, conceptions of science (Hoshmand and Martin, 1994). . . In some cases, psychologists appear to be more interested in contributing to a subdiscipline or specialty than to psychology as a whole (Staats, 1983; MacIntyre, 1985). In this way, fragmentation has been, and continues to be, as much a part of psychology as any of its pragmatic definitional characteristics such as “the study of behavior” or “the study of cognition.” Indeed, there seems to be no evidence that psychology is united by any explicit conception or theoretical framework. (Yanchar and Slife, 1997, p. 236).

What is psychology? Is it a single, coherent scientific discipline awaiting transformation from the current preparadigmatic state into a more mature unified one? Or, is it a heterogeneous federation of subdisciplines that will ultimately fragment into a multitude of smaller, more specialized fields? This is, in essence, the “to be or not to be” question of the field (Henriques, 2004, p. 1207).

*Psychology is what I call a modern disunified science, with a plethora of diverse and unrelated scientific products but with little investment in unifying those products. The resulting disorganization of knowledge leads people such as Toulmin (1972) to consider psychology a “would-be science.” A science in the early stage of disunity does not have the full power of science, and it is not considered to be a full science. That power and that recognition await the beginning of the science’s advancement to unification. Psychology has not begun that arduous journey. That will happen inevitably, in my opinion.* (Staats, 2004, p. 273)

These critical citations do not stand alone. They concisely articulate a contentious meta-theoretical controversy that has been reverberating since the latter decades of the past century (Staats, 1983, 1991, 1999; Kimble, 1989; Sternberg and Grigorenko, 2001; Driver-Linn, 2003; Goertzen, 2008). Psychology is perceived by many as a “house divided,” a fragmented collection of sub-disciplines locked into pigeonholes of disparate theoretical paradigms and levels of construct specification, which makes an integrative understanding of behavior difficult. In fact, this apparent lack of common theoretical principles that spans the array of psychological sub-disciplines has led, in some extreme cases, to the reserved use of the label “scientific” in characterizing psychological inquiry (e.g., Koch, 1993).

In light of this last and serious implication, we present here among the first and most explicit empirical attempts to

counteract the disunity problem. We developed and experimentally illustrated an approach to unification (Glenberg, 2010): sensorimotor mechanisms can be exploited to traverse the cognitive, social, and cultural domains of behavior while sidestepping the incommensurable theoretical metaphors dominant in each of these territories. Consistent with this approach, the two studies reported here strongly point to the involvement of the motor system even in one of the most abstractly-framed areas of human behavior: culture.

By bringing together findings from the cultural and motor-simulation literatures, we predicted that people with interdependent self construals would anticipate needing less motor effort to interact with in-groups than with out-groups. In contrast, people with independent self construals would anticipate more similar motor effort to interact with in-group and out-group members. We took advantage of the visual signature of motor effort (Proffitt, 2006) to examine this cultural motor-effort hypothesis. Based on Proffitt’s work, we expected inflated reports of visual distance to be associated with greater expected effort.

Study 1 confirmed the prediction using two different means of distance estimation, estimated time to walk to a target and estimated number of sticks to the target. Relative to American independents, interdependent Americans reported a shorter expected time to walk to, and fewer sticks to touch, the American in-group confederates. Study 2 replicated and extended the effects by demonstrating that the interdependent Arab participants perceived their in-group Arab confederates as closer than did the independent Arabs, whereas the same Arab confederates were perceived as farther by the interdependent than by the independent American participants. In both studies, the difference between the estimates of the interdependents and independents grew with actual distance, lending further support to the psychological reality of the proposed cultural motor-effort construct.

These result sets are consistent with our prediction that several of the basic characteristics of the motor system (i.e., it scaffolds action recognition and intention-grasping through simulation; it functions predictively by projecting its future states; and it is sensitive to the cost of looming interactions) extend from the basic cognitive (i.e., visual distance perception, Proffitt, 2006), to the interpersonal (i.e., social support, Schnall et al., 2008), and into the domain of self-construal and inter-cultural contact. Importantly, one and the same bodily mechanism can explain these otherwise diverse human behaviors.

Our findings are not the only demonstration of the principle of embodied psychological unity we are trying to promote. In retrospect, many of the embodied-cognition findings may indirectly support the unifying potential of the bodily mechanisms. For example, the neural circuits responsible for the perception of somatic, visceral pain are (a) implicated in one’s own experience of social emotions of seclusion (Eisenberger and Lieberman, 2004), (b) resonate with the perceived pain of others (Immordino-Yang et al., 2009), and (c) this resonance is moderated by personality and cultural factors (Avenanti et al., 2010). As another example, the primary somatosensory cortex (that had long been considered to have a purely epistemic function) was recently found to (a) resonate vicariously with the perceived touch of others (Bolognini et al., 2011), (b) show moderated

activity based on the assumed gender of who applies the touch (Gazzola et al., 2012), and (c) shows higher resonance levels when the observed touch is at a cultural in-group's body (Xu et al., 2009). And third, circuits that represent comparative magnitude, intensity, and extent (i.e., spatial-cognitive functions; Dehaene et al., 2003) were found to serve the homologous social function of status and rank recognition and discrimination (Chiao et al., 2009). Yamakawa et al. (2009), using fMRI, showed that a common neural substrate located in the parietal lobe is implicated when participants judge the proximity of objects in the physical space as well as when they judged relationships of kinship of family members and closeness of friends.

The above results may, in fact, take the argument for embodied psychological unity (as exemplified in the current research) to a neurophysiological level. Rather than being a mere metatheoretical necessity, the contention that bodily mechanisms can serve multiple cognitive, social, and cultural functions may be reflective of a foundational principle for the functional and structural organization of the brain. Anderson (2010) presents extensive evidence that over both the phylogenic and ontogenic brain lifetimes, "neural reuse" is commonplace. That is, the same neural structures are re-used for progressively more advanced functions. Thus, much as we have argued that sensorimotor systems may underlie individual, social, and cultural behaviors, neural reuse may be a neurophysiological mechanism for how the brain responds efficiently to the cognitive, social, and cultural adaptive demands. In this way, neural re-use may underlie the re-use of sensorimotor mechanisms that we have demonstrated (see also Immordino-Yang et al., 2010).

There is also a body of literature in social psychology that is consistent with our findings. As one example, van Baaren et al. (2003) examined how interdependence and independence affects mimicry. Consistent with our notions of tuning and motor resonance, they report that interdependents produced more non-conscious mimicry. Although less strongly tied to the mechanisms we propose, there is also evidence that mimicry (produced by motor resonance, we suppose) also extends to positive social interactions beyond the dyad (Ashton-James et al., 2007).

Nonetheless, we acknowledge that much more research is needed to further validate the empirical unification approach proposed here. As noted by a reviewer of a previous version of this article, future research should employ designs that allow for a fully crossed cross-cultural investigation. Adding an American confederate to Study 2, for example, would permit examining the proposed cultural motor-effort hypothesis at the (national/cultural) group level, in addition to the cultural individual-difference level (i.e., self construals of interdependence and independence) examined here. Alternative interpretations should also be ruled out. For example, future studies should directly record the height and walking speed (toward culturally neutral, inanimate targets) of interdependents and independents to eliminate these two potential systematic confounds that could yield results similar to the ones reported here (although the reversal of the effect for Americans across the studies make this alternative unlikely). Furthermore, we need to develop a more explicit, mechanistic account of exactly how an anticipated increase in interaction could be used to scale distance. In much of Proffitt's previous

work, the connection is close and specific. For example, throwing a heavy ball increases perceived distance to a target when intending to throw, but not when intending to walk. However, in our research and in Schnall et al. (2008) social and cultural factors that are not specifically related to the effectors affect distance perception. Instead, social factors seems to have a generalized effect.

In conclusion, these results are consistent with the cultural motor-effort hypothesis, albeit with the limitations noted above and the possibility of alternative predictions related to self-construal and prejudice noted in the introduction. The results also suggest that the conceptual tools of embodied cognition can be used to help unify psychology by applying the same mechanistic account for behavior at the level of the individual, the social dyad, and the cultural group.

## ACKNOWLEDGMENTS

Experiment 1 was part of Alison Gibson's Senior Thesis. Tamer Soliman was support by a Ford Foundation Fellowship and Seed Funding from the College of Liberal Arts and Sciences, and Arthur M. Glenberg was partially supported by NSF grants 1020367 and 1324807. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

## REFERENCES

- Achten, J., Halson, S., Moseley, L., Rayson, M., Casey, A., and Jeukendrup, A. (2004). Higher dietary carbohydrate content during intensified running training results in better maintenance of performance and mood state. *J. Appl. Physiol.* 96, 1331–1340. doi: 10.1152/jappphysiol.00973.2003
- Aiken, L. S., and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
- Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* 33, 245–266. doi: 10.1017/S0140525X10000853
- Archer, D. (1997). Unspoken diversity: cultural differences in gestures. *Qual. Sociol.* 20, 79–105. doi: 10.1023/A:1024716331692
- Ashton-James, C., van Baaren, R. B., Chartrand, T. L., Decety, J., and Karremans, J. (2007). Mimicry and me: the impact of mimicry on self-construal. *Soc. Cogn.* 25, 518–535. doi: 10.1521/soco.2007.25.4.518
- Aslin, R. N., Jusczyk, P. W., and Pisoni, D. B. (1998). "Speech and auditory processing during infancy," in *Handbook of Child Psychology: Vol. 2, Cognition, Perception, and Language*, 5th ed., eds D. Kuhn and R. Siegler (New York, NY: Wiley), 147–198.
- Avenanti, A., Sirigu, A., and Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Curr. Biol.* 20, 1018–1022. doi: 10.1016/j.cub.2010.03.071
- Biernat, M., Vescio, T. K., Theno, S. A., and Crandall, C. S. (1996). "Values and prejudice: toward understanding the impact of American values on outgroup attitudes," in *The Psychology of Values: The Ontario Symposium on Personality and Social Psychology*, eds C. Seligman, J. M. Olson, and M. P. Zanna (Mahwah, NJ: L. Erlbaum Associates), 153–189.
- Blakemore, S., and Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia* 43, 260–267. doi: 10.1016/j.neuropsychologia.2004.11.012
- Bolognini, N., Rossetti, A., Maravita, A., and Miniussi, C. (2011). Seeing touch in the somatosensory cortex: a TMS study of the visual perception of touch. *Hum. Brain Mapp.* 32, 2104–2114. doi: 10.1002/hbm.21172
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007
- Casile, A., and Giese, M. (2006). Nonvisual motor training influences motor perception. *Curr. Biol.* 16, 69–74. doi: 10.1016/j.cub.2005.10.071

- Chiao, J. Y., Harada, T., Oby, E. R., Li, Z., Parrish, T., and Bridge, D. J. (2009). Neural representations of social status hierarchy in human inferior parietal cortex. *Neuropsychologia* 47, 354–63. doi: 10.1016/j.neuropsychologia.2008.09.023
- Coyle, E. (2004). Fluid and fuel intake during exercise. *J. Sports Sci.* 22, 39–55. doi: 10.1080/0264041031000140545
- Cutting, J., and Vishton, P. (1995). “Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth,” in *Perception of Space and Motion*, eds W. Epstein and S. Rogers (San Diego, CA: Academic Press), 69–117.
- Davis, J., Jackson, D., Broadwell, M., Queary, J., and Lambert, C. (1997). Carbohydrate drinks delay fatigue during intermittent, high-intensity cycling in active men and women. *Int. J. Sport Nutr.* 7, 261–273.
- Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506. doi: 10.1080/02643290244000239
- Desy, M. C., and Theoret, H. (2007). Modulation of motor cortex excitability by physical similarity with an observed hand action. *PLoS ONE* 2:e971. doi:10.1371/journal.pone.0000971
- Driver-Linn, E. (2003). Where is psychology going? Structural fault lines revealed by psychologists’ use of Kuhn. *Am. Psychol.* 58, 269–279. doi: 10.1037/0003-066X.58.4.269
- Eisenberger, N. I., and Lieberman, M. D. (2004). Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn. Sci.* 8, 294–300. doi: 10.1016/j.tics.2004.05.010
- Ekman, P. (1972). “Universals and cultural differences in facial expressions of emotion,” in *Nebraska Symposium on Motivation*, ed J. Cole (Lincoln, NE: University of Nebraska Press).
- Ekman, P., Friesen, W. V., O’Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *J. Pers. Soc. Psychol.* 53, 712–717. doi: 10.1037/0022-3514.53.4.712
- Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* 15, 399–402. doi: 10.1046/j.0953-816x.2001.01874.x
- Flege, J. E. (1988) Factors affecting degree of perceived foreign accent in English sentences. *J. Acoust. Soc. Am.* 84, 70–79.
- Fokes, J., and Bond, Z. (1984). “Patterns of English word stress by native and non-native speakers,” in *Proceedings of the Tenth International Congress of Phonetic Sciences*, eds M. Van den Broecke and A. Cohen (Dordrecht: Foris), 682–686.
- Forgas, J., Williams, K., and Laham, S. (2005). *Social Motivation: Conscious and Unconscious Processes*. Cambridge: Press syndicate of the University of Cambridge.
- Gatbonton, E. (1975). *Systematic Variations in Second Language Speech: a Sociolinguistic Study*. Unpublished Ph.D. thesis, McGill University.
- Gazzola, V., Spezio, M. L., Etzel, J. A., Castelli, F., Adolphs, R., and Keysers, C. (2012). PNAS Plus: primary somatosensory cortex discriminates affective significance in social touch. *Proc. Natl. Acad. Sci.* 109, E1657–E1666. doi: 10.1073/pnas.1113211109
- Glenberg, A. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 586–596. doi: 10.1002/wcs.55
- Goertzen, J. (2008). On the possibility of unification: the reality and nature of the crisis in psychology. *Theory Psychol.* 18, 829–852. doi: 10.1177/0959354308097260
- Henriques, G. R. (2004). Psychology defined. *J. Clin. Psychol.* 60, 1207–1221. doi: 10.1002/jclp.20061
- Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937. doi: 10.1017/S0140525X01000103
- Hong, Y., Morris, M. W., Chiu, C., and Benet-Martinez, V. (2000). Multicultural minds: a dynamic constructivist approach to culture and cognition. *Am. Psychol.* 55, 709–720. doi: 10.1037/0003-066X.55.7.709
- Hoshmand, L. T., and Martin, J. (1994). Naturalizing the epistemology of psychological research. *J. Theor. Philos. Psychol.* 14, 171–189. doi: 10.1037/h0091136
- Immordino-Yang, M., Chiao, J., and Fiske, A. (2010). Neural reuse in the social and emotional brain. *Behav. Brain Sci.* 33, 275–276. doi: 10.1017/S0140525X10001020
- Immordino-Yang, M. H., McColl, A., Damasio, H., and Damasio, A. (2009). Neural correlates of admiration and compassion. *Proc. Natl. Acad. Sci.* 106, 8021–8026. doi: 10.1073/pnas.0810363106
- Katz, I., and Hass, R. (1988). Racial ambivalence and American value conflict: correlational and priming studies of dual cognitive structures. *J. Pers. Soc. Psychol.* 55, 893–905. doi: 10.1037/0022-3514.55.6.893
- Kimble, G. A. (1989). Psychology from the standpoint of a generalist. *Am. Psychol.* 44, 491–499. doi: 10.1037/0003-066X.44.3.491
- Kleugel, J. R. (1990). Trends in Whites’ explanations of the Black-White gap in socioeconomic status, 1977–1989. *Am. Sociol. Rev.* 55, 521–525. doi: 10.2307/2095804
- Knoblich, G., and Sebanz, N. (2006). The social nature of perception and action. *Curr. Dir. Psychol. Sci.* 15, 99–104. doi: 10.1111/j.0963-7214.2006.00415.x
- Koch, S. (1993). “Psychology” or “the psychological studies?” *Am. Psychol.* 48, 902–904.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608. doi: 10.1126/science.1736364
- Liew, S. L., Han, S., and Aziz-Zadeh, L. (2011). Familiarity modulates mirror neuron and mentalizing regions during intention understanding. *Hum. Brain Mapp.* 32, 1986–1997. doi: 10.1002/hbm
- MacIntyre, R. B. (1985). Psychology’s fragmentation and suggested remedies. *Int. Newslett. Paradigmat. Psychol.* 1, 20–21.
- Markus, H., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224
- Matsumoto, D., and Kuppertsbusch, C. (2001). Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience. *Asian J. Soc. Psychol.* 4, 113–131. doi: 10.1111/j.1467-839X.2001.00080.x
- Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., et al. (2002). American-Japanese cultural differences in judgments of emotional expressions of different intensities. *Cogn. Emot.* 16, 721–747. doi: 10.1080/02699930143000608
- Messick, D., and Mackie, D. (1989). Intergroup relations. *Annu. Rev. Psychol.* 41, 40–45.
- Molnar-Szakacs, I., Wu, A. D., Robles, F. J., and Iacoboni, M. (2007). Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PLoS ONE* 2: e626. doi: 10.1371/journal.pone.0000626
- Muhlau, M., Hermsdörfer, J., Goldenberg, G., Wohlschläger, A. M., Castrop, F., Stahl, M., et al. (2005). Left inferior parietal dominance in gesture imitation: an fMRI. *Neuropsychologia* 43, 1086–1098. doi: 10.1016/j.neuropsychologia.2004.10.004
- Newell, A. (1980). Physical symbol systems. *Cogn. Sci.* 4, 135–183. doi: 10.1207/s15516709cog0402\_2
- Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently*. New York, NY: Free Press.
- Oishi, S. (2010). The psychology of residential mobility: implications for the self, social relationships, and well-being. *Perspect. Psychol. Sci.* 5, 5–21. doi: 10.1177/1745691609356781
- Oishi, S., and Kislind, J. (2009). “The mutual constitution of residential mobility and individualism,” in *Understanding Culture: Theory, Research, and Application*, eds R. S. Wyer Jr., C.-Y. Chiu, Y. Y. Hong, and S. Shavitt (New York, NY: Psychology Press), 223–238.
- Petroni, A., Baguear, F., and Della-Maggiore, V. (2010). Motor Resonance may originate from sensorimotor experience. *J. Neurophysiol.* 104, 867–871. doi: 10.1152/jn.00386.2010
- Proffitt, D. (2006). Embodied perception and the economy of action. *Perspect. Psychol. Sci.* 1, 110–122. doi: 10.1111/j.1745-6916.2006.00008.x
- Proffitt, D. R., and Linkenauger, S. A. (2013). “Perception viewed as a phenotypic expression,” in *Action Science: Foundations of an Emerging Discipline*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 171–197.
- Rizzolatti, G., Fogassi, G., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670. doi: 10.1038/35090060
- Robbins, J. M., and Krueger, J. I. (2005). Social projection to ingroups and outgroups: a review and meta-analysis. *Pers. Soc. Psychol. Rev.* 9, 32–47. doi: 10.1038/35054102
- Schnall, S., Harber, K., Stefanucci, J., and Proffitt, D. (2008). Social support and the perception of geographical slant. *J. Exp. Soc. Psychol.* 44, 1246–1255. doi: 10.1016/j.jesp.2008.04.011



- Schubert, T. W., and Semin, G. R. (2009). Embodiment as a unifying perspective for psychology. *Eur. J. Soc. Psychol.* 39, 1135–1141. doi: 10.1002/ejsp.670
- Schug, J., Yuki, M., and Maddux, W. (2010). Relational mobility explains between- and within-culture differences in self-disclosure to close friends. *Psychol. Sci.* 21, 1471–1478. doi: 10.1177/0956797610382786
- Sears, D. O., and Henry, P. J. (2005). Over thirty years later: a contemporary look at symbolic racism and its critics. *Adv. Exp. Soc. Psychol.* 37, 95–150. doi: 10.1016/S0065-2601(05)37002-X
- Shweder, R. (1996). “True ethnography: the lore, the law, and the lure,” in *Ethnography and Human Development*, eds R. Jessor, A. Colby, and R. Shweder (Chicago: University of Chicago Press), 15–52.
- Singelis, T. (1994). The measurement of independent and interdependent self-construals. *Pers. Soc. Psychol. Bull.* 20, 580–591. doi: 10.1177/0146167294205014
- Staats, A. W. (1983). *Psychology's Crisis of Disunity: Philosophy and Method for a Unified Science*. New York, NY: Praeger.
- Staats, A. W. (1991). Unified positivism and unification psychology: fad or new field? *Am. Psychol.* 46, 899–912. doi: 10.1037/0003-066X.46.9.899
- Staats, A. W. (1999). Unifying psychology requires new infrastructure, theory, method, and a research agenda. *Rev. Gen. Psychol.* 3, 3–13. doi: 10.1037/1089-2680.3.1.3
- Staats, A. W. (2004). The disunity-unity dimension. *Am. Psychol.* 59, 273. doi: 10.1037/0003-066X.59.4.273a
- Sternberg, R. J., and Grigorenko, E. L. (2001). Unified psychology. *Am. Psychol.* 56, 1069–1079. doi: 10.1037/0003-066X.56.12.1069
- Tanaka, S., Inui, T., Iwaki, S., Konishi, J., and Nakai, T. (2001). Neural substrates involved in imitating finger configurations: an FMRI. *Neuroreport* 12, 1171–1174. doi: 10.1097/00001756-200105080-00024
- Tanaka, A., Koizumi, A., Imai, H., Hiramatsu, S., Hiramoto, E., and de Gelder, B. (2010). I feel your voice: cultural differences in the multisensory perception of emotion. *Psychol. Sci.* 21, 1259–1262. doi: 10.1177/0956797610380698
- Toulmin, S. (1972). *Human Understanding*. Princeton, NJ: Princeton University Press.
- Triandis, H., Bontempo, R., Villareal, M., Asai, M., and Lucca, N. (1988). Individualism and collectivism: cross-cultural perspectives on self-in-group relationships. *J. Pers. Soc. Psychol.* 54, 323–338. doi: 10.1037/0022-3514.54.2.323
- van Baaren, R. B., Maddux, W. W., Chartrand, T. L., de Bouter, C., and Knippenberg, A. (2003). It takes two to mimic: behavioral consequences of self-construals. *J. Pers. Soc. Psychol.* 84, 1093–1102. doi: 10.1037/0022-3514.84.5.1093
- Wilson, M., and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychol. Bull.* 131, 460–473. doi: 10.1037/0033-2909.131.3.460
- Xu, X., Zuo, X., Wang, X., and Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *J. Neurosci.* 29, 85–25. doi: 10.1523/JNEUROSCI.2418-09.2009
- Yamakawa, Y., Kanai, R., Matsumura, M., and Naito, E. (2009). Social distance evaluation in human parietal cortex. *PLoS ONE* 4:e4360. doi: 10.1371/journal.pone.0004360
- Yanchar, S. C., and Slife, B. D. (1997). Pursuing unity in a fragmented psychology: problems and prospects. *Rev. Gen. Psychol.* 1, 235–255. doi: 10.1037/1089-2680.1.3.235

**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.

Received: 05 July 2013; accepted: 07 November 2013; published online: 29 November 2013.

Citation: Soliman T, Gibson A and Glenberg AM (2013) Sensory motor mechanisms unify psychology: the embodiment of culture. *Front. Psychol.* 4:885. doi: 10.3389/fpsyg.2013.00885

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2013 Soliman, Gibson and Glenberg. 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.



# Action scaling of distance perception is task specific and does not predict “the embodiment of culture”: a comment on Soliman, Gibson, and Glenberg (2013)

Andrew D. Wilson\*

School of Social, Psychological and Communication Sciences, Leeds Metropolitan University, Leeds, UK

\*Correspondence: DrAndrewDWilson@gmail.com;

Web: <http://psychsciencenotes.blogspot.co.uk>;

Twitter: @PsychScientists

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Arthur M. Glenberg, Arizona State University and University of Wisconsin-Madison, USA

Guy Dove, University of Louisville, USA

**Keywords:** Proffitt, embodiment, action-scaling, perceived distance

## A commentary on

### Sensory motor mechanisms unify psychology: the embodiment of culture

by Soliman, T., Gibson, A., and Glenberg, A. M. (2013). *Front. Psychol.* 4:885. doi: 10.3389/fpsyg.2013.00885

There is now extensive evidence to support James J. Gibson's (1966, 1979) hypothesis that our perception of the environment is scaled in terms of our ability to act on that environment; by its *affordances*. One strand of evidence comes from Proffitt, who has shown that changing a person's ability to act affects how they judge their ability to perform an upcoming task. The most famous example (Bhalla and Proffitt, 1999) showed that people judge hills to be steeper when they are wearing a heavy backpack. The hypothesis is that the backpack would increase the effort required to climb the hill, and thus we perceive the hill as more difficult to climb (see Proffitt and Linkenauger, 2013 for a recent review).

Soliman et al. (2013) applied Proffitt's particular sensorimotor mechanism to a cultural context. They asked participants to judge the distances between themselves and in- or out-group members. The logic is as follows:

- (1) Interacting with someone requires you to (unconsciously) mentally simulate things they do
- (2) We do this simulation ahead of time, in anticipation of an encounter

- (3) This simulation will require more effort with out-group members, because they do things differently to you
- (4) This increased simulation effort will affect your judgment of distance, as per Proffitt
- (5) Therefore out-group members should look further away

The results seem promising; for example, American and Arab participants who rated themselves with an interdependent self-construal did rate their in-group members as closer than did participants with an independent self-construal. Soliman et al. argue that these results support an analysis of culture within Proffitt's embodied framework, which would be interesting if true because it would let us talk about both perceptual and cultural effects within a unified framework.

There is a major problem, however. Proffitt's research very clearly shows that increasing the effort required to perform a task only affects distance perception related to that task. For example, making walking harder by fatiguing the legs increases the perceived distance if you plan to cross it by walking, but not if you plan to cross it by throwing (Witt et al., 2010). Soliman et al., however, claim that the increased effort of internally simulating the movements required to interact with an out-group member will increase the perceived distance to that person when that distance is to be traversed

by locomotion. Their effort manipulation has nothing to do with locomoting across the distance and thus, contrary to the framing of their paper, their results are neither predicted by nor explained by reference to Proffitt's action-scaling theory.

Task specificity is central to Proffitt's theory. In a recent debate, Firestone (2013) highlighted this because he believes this creates a problem for Proffitt; if distance perception for walking and throwing are calibrated to different scales, you cannot compare the two in order to choose the best way to cross that distance. Proffitt (2013) disagreed that this creates a problem but completely agreed that action-specific units are incommensurable in this way. He stated “An important finding across our studies is that the influence of an action unit—such as graspability—is evident only within its action boundary” (p. 477). This exchange is relevant because Proffitt is specifically challenged here on this point and comes out strongly and unambiguously in favor of task-specificity.

Soliman et al. rest their non-task-specific analysis on one paper (Schnall et al., 2008). Participants in this study judged hills as less steep when accompanied by or thinking about a friend. The claim here is that social support makes the hill appear more easily traversed, without any apparent recalibration of task-relevant effectors. Soliman et al. argue that this supports their hypothesis that an increase in upcoming social effort

(the hypothesized prospective internal simulation of the other person described in points 1–3 above) could alter distance perception. However, it is worth noting that Schnall et al. specifically caution that “it is too early to speculate on the degree to which these influences [the effects of physical vs. psychological states on slant judgments] share common underlying mechanisms or on what these mechanisms might be” (p. 1254). In addition, when we place this single result in the context of the rest of Proffitt’s extensive body of work repeatedly demonstrating strong task specificity, it actually seems more likely right now that the *only* way in which a friend could help make a hill look more climbable is by doing something that recalibrates the embodied hill-climbing system. Discovering what this something is would be an invaluable contribution to the unification Soliman et al. propose, but it remains to be done.

## SUMMARY

Soliman et al. (2013) claim that the increased mental effort required to simulate an upcoming encounter with an out-group member will make the distance to that person look more difficult to cross and thus the person will look farther away. They ground this hypothesis in Proffitt’s embodied action-scaling theory of perception, but Proffitt’s data supports a strong form of task-specificity that means his theory neither predicts nor explains the current results.

The current data (plus Schnall et al., 2008) may eventually motivate a less

task-specific version of Proffitt’s mechanism. For example, the interaction of self-construal with distance Soliman et al. find is consistent with the claim that the interdependent and independent groups are evaluating the distances using different metrics (see p. 4–5). But whether those metrics are effort based (overturning the otherwise extensive evidence in favor of task-specificity) remains to be confirmed.

I am personally all in favor of an embodied approach to unifying psychology, but as I have argued (Wilson and Golonka, 2013) this will require careful attention to the details of the relevant sensorimotor (perception-action) mechanisms so that we are sure we are connecting them to “higher level” cognition in ways that reflect how those mechanisms actually operate. This connection is simply not present in the target article, and the implication for Soliman et al. is that their data do not support their particular attempt to unify psychology with sensorimotor mechanisms.

## REFERENCES

- Bhalla, M., and Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1076–1096. doi: 10.1037//0096-1523.25.4.1076
- Firestone, C. (2013). How “paternalistic” is spatial perception? Why wearing a heavy backpack doesn’t - and couldn’t - make hills look steeper. *Perspect. Psychol. Sci.* 8, 455–473. doi: 10.1177/1745691613489835
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton-Mifflin.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Psychology Press.
- Proffitt, D. R. (2013). An embodied approach to perception: by what units are visual perceptions scaled? *Perspect. Psychol. Sci.* 8, 474–483. doi: 10.1177/1745691613489837
- Proffitt, D. R., and Linkenauger, S. A. (2013). “Perception viewed as a phenotypic expression,” in *Action Science: Foundations of an Emerging Discipline*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 171–197.
- Schnall, S., Harber, K., Stefanucci, J., and Proffitt, D. R. (2008). Social support and the perception of geographical slant. *J. Exp. Soc. Psychol.* 44, 1246–1255. doi: 10.1016/j.jesp.2008.04.011
- Soliman, T., Gibson, A., and Glenberg, A. M. (2013). Sensory motor mechanisms unify psychology: the embodiment of culture. *Front. Psychol.* 4:885. doi: 10.3389/fpsyg.2013.00885
- Wilson, A. D., and Golonka, S. (2013). Embodied cognition is not what you think it is. *Front. Psychol.* 4:58. doi: 10.3389/fpsyg.2013.00058
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2010). When and how are spatial perceptions scaled? *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1153–1160. doi: 10.1037/a0019947

**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.

Received: 19 February 2014; accepted: 23 March 2014; published online: 21 April 2014.

Citation: Wilson AD (2014) Action scaling of distance perception is task specific and does not predict “the embodiment of culture”: a comment on Soliman, Gibson, and Glenberg (2013). *Front. Psychol.* 5:302. doi: 10.3389/fpsyg.2014.00302

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 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.



# How intent to interact can affect action scaling of distance: reply to Wilson

Tamer M. Soliman and Arthur M. Glenberg \*

Department of Psychology, Arizona State University, Tempe, AZ, USA

\*Correspondence: [glenberg@asu.edu](mailto:glenberg@asu.edu)

Edited and reviewed by:

Guy Dove, University of Louisville, USA

**Keywords:** culture, self-construal, embodied cognition, distance perception, motor effort

## A commentary on

Action scaling of distance perception is task specific and does not predict “the embodiment of culture”: a comment on Soliman, Gibson and Glenberg (2013) by Wilson, A. D. (2014). *Front. Psychol.* 5:302. doi: 10.3389/fpsyg.2014.00302

Soliman et al. (2013) set out to demonstrate how the bodily level of analysis can unify explanations in psychology. Our argument was that common sensorimotor mechanisms underlie many of the behavioral phenomena that are currently segregated as cognitive, social, or cultural. Toward that end, we re-characterized a cultural construct—self-construal along the dimension of independence and interdependence (Markus and Kitayama, 1991)—as reflecting degree of interaction with ethnically diverse others.

According to our cultural motor-effort hypothesis, the interdependence-independence continuum is in part determined by tuning sensorimotor behavior through interactions. Interdependents tune vocal, gestural, expressive facial patterns, as well as interactions in greeting, eating, walking, dancing, praying (and so on) with members of their in-groups. In contrast, independents tune their interactions with a more ethnically-diverse set of people. Consequently, interdependents, more so than independents, would anticipate greater motor effort when interacting with out-groups (vs. in-groups) because of poor tuning. Furthermore, reasoning from Proffitt and Linkenauger (2013) as well as Schnall et al. (2008), anticipated motor effort should lead to increased distance judgments. Thus we predicted, and found, that interdependents judge distance

to in-group members as shorter than do independents.

Wilson (2014) questioned our application of Proffitt and Linkenauger and Schnall et al. As he notes, Proffitt’s data (although not data from Schnall et al.) suggest that effects of anticipated motor effort are restricted to particular motor systems. Hence, Wilson reasoned, the anticipated effort in interacting should not affect scaling of distance when planning to walk. Here we address Wilson’s reasoning by (a) pointing to several research projects that suggest leakage across motor systems rather than modularity, and (b) suggesting why previous data, importantly, Witt et al. (2004, 2010) did not observe this leakage.

As one example of leakage, consider data reported by Gentilucci et al. (2001). When reaching for a block, the larger the block, the wider people unintentionally open their mouths. In addition, the larger the block, the louder they pronounce syllables printed on the block.

Now consider in more detail retroactive motor contagion (RMC): the ubiquitous finding that if two action patterns are conjoined, planning of the second action influences planning of the first action. Demonstrations of RMC can be found in Adam et al. (2000), Khan et al. (2007, 2010), and Lajoie and Franks (1997).

The “end-comfort effect” can also be seen as a type of RMC. For example, the kinematics of the transport-to-grasp movement toward a bottle systematically vary depending on whether the bottle is later to be displaced to a different spot, is used to pour water into a glass, or is to be thrown away (Ansuini et al., 2008). Importantly, RMC can cross anatomical and neuro-representational boundaries

within the motor system and shows coordination across different effectors. van der Wel and Rosenbaum (2007), for example, asked participants to locomote to a table, grasp a bottle, and then move it to another spot that was either close to or far from its initial location. The initial motor pattern (i.e., locomotion) was found to be influenced by the distal motor pattern (i.e., object transport). Namely, a participant’s final step was on the side opposite to the direction of the forthcoming transport movement when that transport required one more step after grasping (see also Cockell et al., 1995 and Studenka et al., 2012).

Thus, modularity of the motor system at the anatomical and brain-representational level does not always hold at the functional level. Instead, conjoining two action patterns induces an informational flow across effectors and planned goals. Importantly, this influence holds whether one or different motor systems are involved in the sequence, and whether the goals planned are homologous (e.g., tapping followed by tapping) or different (i.e., locomoting then grasping). In short, these findings support our assumption that anticipated effort of interaction can affect anticipated effort to walk, and thereby affect distance judgments.

With the above as a backdrop, why then do Proffitt’s data (e.g., Witt et al., 2004, 2010) seem to suggest modularity? One possibility is based on a subtle difference between the design of our experiments and the Witt et al. experiments. In Witt’s experiments, the manipulation phase targets one motor system and then tests the effect of the manipulation on perceived distance as the participant intends to



perform another task. For example, adapting Proffitt and Linkenauger's (2013) terminology, participants are adapted while temporarily turned into throwing phenotypes, and then tested while in the walking phenotype. Typically, it was found that the visio-motor scale developed while in one phenotype did not transfer to the other: the thrower-turned-walker participants do not show effects of the earlier throwing manipulation on their distance judgments while walkers (Witt et al., 2004), and vice versa (Witt et al., 2010).

In our experiments, however, no behavioral phenotype was turned on, manipulated, switched off, replaced by another, and then examined. Instead, our participants were walker-then-interactor phenotypes throughout. That is, the phenotype we manipulated (the interactor phenotype) was (a) always turned on and (b) always conjoined with the walker phenotype. Thus, by virtue of being conjoined with the interaction system during simulation, the locomotion system was "contaminated" by the constantly-running interaction system. This conjoining led to the effort parameter values instantiated in the interaction system to diffuse into the parameters in the locomotion system. We captured the state of the latter parameters through visual-distance estimates, and we hypothesized that they function, by proxy, as indicators of the amount of effort experienced by the interaction system.

We believe that these subtle design differences render our original results and theoretical arguments immune to Wilson's critiques. Perceived motor effort to interact with in-groups and out-groups can still be a conceptually valid re-characterizations of the cultural construct of interdependence-independence. And, importantly, when viewed in light of the RMC effects, our results can be categorized as belonging to the same class of phenomena explained by Proffitt's

theoretical framework. We thank Wilson for providing the opportunity for us to develop this account in greater detail, and we look forward to tests of the proposal.

## ACKNOWLEDGMENTS

Tamer Soliman was support by a Ford Foundation Fellowship and Seed Funding from the College of Liberal Arts and Sciences, and Arthur Glenberg was partially supported by NSF grants 1020367 and 1324807. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

## REFERENCES

- Adam, J. J., Nieuwenstein, J. H., Huys, R., Paas, F. G., Kingma, H., Willems, P., et al. (2000). Control of rapid aimed hand movements: the one-target advantage. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 295–312. doi: 10.1037/0096-1523.26.1.295
- Ansuini, C., Giosa, L., Turella, L., Altoe, G., and Castiello, U. (2008). An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* 185, 111–119. doi: 10.1007/s00221-007-1136-4
- Cockell, D. L., Carnahan, H., and McFadyen, B. J. (1995). A preliminary analysis of the coordination of reaching, grasping, and walking. *Percept. Mot. Skills* 81, 515–519. doi: 10.2466/pms.1995.81.2.515
- Gentilucci, M., Benuzzi, F., Gangitano, M., and Grimaldi, S. (2001). Grasp with hand and mouth: a kinematic study on healthy subjects. *J. Neurophysiol.* 86, 1685–1699.
- Khan, M., Mottram, T., Adam, J., and Buckolz, E. (2010). Sequential movement with two limbs and the one target advantage. *J. Mot. Behav.* 42, 325–332. doi: 10.1080/00222895.2010.510544
- Khan, M. A., Mourton, S., Buckolz, E., and Franks, I. M. (2007). The influence of advance information on the response complexity effect in manual aiming movements. *Acta. Psychol. (Amst)*, 127, 154–162. doi: 10.1016/j.actpsy.2007.04.001
- Lajoie, J. M., and Franks, I. M. (1997). Response programming as a function of accuracy and complexity: evidence from latency and kinematic measures. *Hum. Mov. Sci.* 16, 485–505.
- Markus, H., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224
- Proffitt, D. R., and Linkenauger, S. A. (2013). "Perception viewed as a phenotypic expression," in *Action Science: Foundations of An Emerging Discipline*, eds W. Prinz, M. Beisert, and A. Herwig (Cambridge, MA: MIT Press), 171–197.
- Schnall, S., Harber, K., Stefanucci, J., and Proffitt, D. (2008). Social support and the perception of geographical slant. *J. Exp. Soc. Psychol.* 44, 1246–1255. doi: 10.1016/j.jesp.2008.04.011
- Soliman, T., Gibson, A., and Glenberg, A. M. (2013). Sensory motor mechanisms unify psychology: the embodiment of culture. *Front. Psychol.* 4:885. doi: 10.3389/fpsyg.2013.00885
- Studenka, B. E., Seegelke, C., Schutz, C., and Schack, T. (2012). Posture based motor planning in a sequential grasping task. *J. Appl. Res. Mem. Cogn.* 1, 89–95. doi: 10.1016/j.jarmac.2012.02.003
- Wilson, A. D. (2014). Action scaling of distance perception is task specific and does not predict "the embodiment of culture": a comment on Soliman, Gibson and Glenberg (2013). *Front. Psychol.* 5:302. doi: 10.3389/fpsyg.2014.00302
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: a role of effort and intent. *Perception* 33, 570–590. doi: 10.1068/p5090
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2010). When and how are spatial perceptions scaled? *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1153–1160. doi: 10.1037/a0019947
- van der Wel, R. P., and Rosenbaum, D. A. (2007). Coordination of locomotion and prehension. *Exp. Brain Res.* 176, 281–287. doi: 10.1007/s00221-006-0618-0

**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.

Received: 08 April 2014; accepted: 10 May 2014; published online: 05 June 2014.

Citation: Soliman TM and Glenberg AM (2014) How intent to interact can affect action scaling of distance: reply to Wilson. *Front. Psychol.* 5:513. doi: 10.3389/fpsyg.2014.00513

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Soliman and Glenberg. 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.



# Linguistic embodiment and verbal constraints: human cognition and the scales of time

Stephen J. Cowley\*

Language and Communication, Centre for Human Interactivity and the COMAC Cluster, University of Southern Denmark, Slagelse, Denmark

## Edited by:

Guy Dove, University of Louisville,  
USA

## Reviewed by:

Etienne B. Roesch, University of  
Reading, UK

Paul John Thibault, University of  
Agder, Norway

## \*Correspondence:

Stephen J. Cowley, Language and  
Communication, Centre for Human  
Interactivity and the COMAC  
Cluster, University of Southern  
Denmark, Sdr. Stationsvej 28, 4200  
Slagelse, Denmark  
e-mail: cowley@sdu.dk

Using radical embodied cognitive science, the paper offers the hypothesis that language is symbiotic: its agent-environment dynamics arise as linguistic embodiment is managed under verbal constraints. As a result, co-action grants human agents the ability to use a unique form of phenomenal experience. In defense of the hypothesis, I stress how linguistic embodiment *enacts* thinking: accordingly, I present auditory and acoustic evidence from 750 ms of mother-daughter talk, first, in fine detail and, then, in narrative mode. As the parties attune, they use a dynamic field to co-embodiment speech with experience of *wordings*. The latter arise in making and tracking phonetic gestures that, crucially, mesh use of artifice, cultural products and impersonal experience. As observers, living human beings gain dispositions to display and use social subjectivity. Far from using brains to “process” verbal content, linguistic symbiosis grants access to diachronic resources. On this distributed-ecological view, language can thus be redefined as: “activity in which wordings play a part.”

**Keywords:** cognitive linguistics, prosody, coordination, social interaction, distributed cognition, ecological psychology, enactivism, distributed language

“The important issue is not where cognitive processing begins and ends.”

(Vallée-Tourangeau and Vallée-Tourangeau, 2014).

## INTRODUCTION

Since it is beyond debate that living systems depend on metabolism, it can seem trivially true that cognitive activity draws on embodiment. To block any such view, the paper turns to *how* metabolism functions as “cognition emerges in ecological space and ecological time from the interactions of brain activity, motor actions, and artifacts” (Vallée-Tourangeau and Vallée-Tourangeau, 2014). In examining coordination in a multi-scalar ecology, it pursues Chemero’s (2009) thesis that agent-environment dynamics ground all cognitive activity. In language, representation is replaced by emphasis on how persons concert activity or, simply, come to act as observers<sup>1</sup>. The resulting skills underpin the paper’s thesis: language is activity based on symbiotic control of bodily movements that are perceived as “wordings.” Given phenomenal experience of iterated patterns, understanding connects the subjective to the impersonal or, alternatively, linguistic embodiment falls under partial control of a community’s verbal constraints. Humans thus live in social meshworks of families, groups, communities, and even nations—each with characteristic ways of using linguistic embodiment.

While all embrained species interlace action and perception, in human groups, some of the time, and to some extent, people use “self-directed” “representational acts,” mimetic forms of activity

that emerged millions of years before language (Donald, 1991). They arise under the control of one or more persons and contribute much to a community’s forms of life. Ways of embodying mimetic performance appear in knapping flint, kicking a ball around, dancing, or taking part in talk. In each kind of activity, the results connect a human lineage with histories of individuals, relationships and ways of exploiting cultural complexity. Crucially, they link local and situated events to the products of a group’s history. Unlike other primates, humans use extended cognitive systems. People draw on the past to alter later behavior: lived experience is enriched by using artifice and language to connect up the scales of time. Both distributed (Hutchins, 1995, 2014) and systemic (Cowley and Vallée-Tourangeau, 2013) approaches stress the multi-scalar nature of human cognition.

Culture enables people to tackle tasks by using structures that criss-cross temporal dimensions. A classic case is that of equipment that is designed to create images from deep-space (Giere, 2004). Human-technology aggregates enable the Hubble, for example, to connect the slow scales of physical evolution with the rapidity of light and mid-scales of embodiment and observation. In Giere’s terms, *distributed systems* link artifacts, measuring, and the doings of human individuals. In spite of supra-individual complexity, the Hubble was made by and for observers. This is necessary to distributed systems: they depend on persons who use resources, including languages, to interpret what is perceived. Observers connect embodied measures with verbal skills to animate systems whose temporal scope reaches beyond lived experience. In what follows, this perspective is applied to human language-derived skills. Bodily synergies enable people to track and construe vocal movements: however, linguistic embodiment

<sup>1</sup> People do not just engage with the world: they also treat what they perceive and how they act as making sense and/or meaningful (e.g., Maturana and Varela, 1994).

is also run through with affect. As people cooperate, compete, and otherwise coordinate, they set off nonce events that will be called *wordings*<sup>2</sup>. These phenomena arise as speech gestures constrain other aspects of linguistic embodiment. Crucially, they shape a richer phenomenal field: wordings co-occur with visible behavior and resonances as the vocal folds modulate the flow of air through changing vocal tracts.

People link linguistic embodiment with verbal and discursive patterns as they animate distributed systems. They do so both unthinkingly and as actors. While reiterating speech patterns, context can be used to attribute properties to so-called “words.” Eschewing appeal to representation or content, I thus liken the concept of language to the concept of mind. As proposed by Ryle (1949), I explain *belief* in mind and language without appealing to linguistic or mental systems “in the head.” Rather than posit dependence on neural dispositions, however, I stress the meshing of living embodiment with phenomenal experience. As a result it is possible to hypothesize that language is symbiotic: the phenomenal field is influenced, in part, by how people make and track phonetic gestures. By tracing the verbal to the phenomenal the paper rejects the code-metaphor<sup>3</sup>, or, in other terms, the view that “inner” systems process, generate or produce linguistic forms. Rather, embodiment links phenomenal experience to verbal patterns as, during ontogenesis, humans become actor-observers. In so doing, speaking and cooperating come under a degree of collective control. People gain skills in using a multi-scalar linguistic resource that allows embodiment to evoke impersonal products as people manage later events (cf. Hollan et al., 2000).

Since language is distributed across space and time, much can be learned from examining whole-body expression. To naturalize linguistic experience, one can thus begin with how bodily dynamics take on a verbal aspect. For, on this view, language arises as people coordinate and interpret events by linking movement with experience (Cowley, 2011b); the verbal is secondary and derived. In distributed terms, human cognition links interactivity (Cowley and Vallée-Tourangeau, 2013; Steffensen, 2013) with the normative sense-saturated flow of experience. Just as people learn to identify objects, events and situations, they learn to perceive reiterating phenomenal patterns, or wordings. As unique events, their sense results from experience with how a community uses phonetic gestures—events are heard against types that shape belief in words and languages. Humans become observers

by learning to hear and articulate wordings. Once these are used to inform coordination, linguistic embodiment can shape activity around cultural goals and tasks. As a result, language is necessarily symbiotic. First, as emphasized below, linguistic embodiment is affective experience or a flow of direct meaning making. Second, its phenomenal aspect links phonetic gestures to wordings that allow descriptions of what linguists usually call language. Given a strange duality, language extends experience as people manage actions that are described by, but do not reduce to, the working of linguistic form and semantic content. The duality of wordings and coordinated action allows human cognition to reach far beyond the body. This is because, in concerting across time, people mesh verbal patterns with skillful action. Language has a multi-scalar heterogeneity akin to that of music, pottery or scientific practice. Humans exploit the scales of time as people who co-opt and transform material and biological structures. As a result, a global meshwork constrains activity in a staggeringly complex social world.

## BRIEF OVERVIEW

My hypothesis is that linguistic embodiment and verbal constraints are symbiotic. Taking a distributed-ecological view, emphasis falls on, not linguistic forms (or content), but how cognitive resources are put to use in managing temporal experience. First, language is traced to the rapid or pico-scale dynamics that dominate linguistic embodiment: as shown below, it enacts measurable and observable bodily events. Second, parties are shown to use phonetic gestures, phenomenal experience and, given unending repetition, lay and linguistic *concepts* of language (*qua* verbal pattern). Further, while science cannot rest on faith in words, as argued by Sellars, Ryle, and Dennett (among others), just such beliefs underlie the social order and, thus, our accounts of human action. In tracing language to a symbolic-dynamical symbiosis, much can be shown to draw on co-embodiment. Accordingly, the core of the paper offers detailed description of events during 750 ms of dialog or languaging activity. Exploiting a pico-scale, mother and daughter coordinate by using voice dynamics that contribute to slower phonetic gestures (transcribed as [a:bene]). These dynamics connect phenomenal experience of a wording with innumerable accounts of a verbal pattern or second-order construct (that can be written “ah bene”). The relevant praxis evokes a history of non-local events that gives each party her own understanding of what occurs. My hypothesis is thus defended by detailed description of how human observers use the symbiotic nature of language. This clarifies what “goes on” (at least roughly)—human actors make it happen. As languaging beings, we use beliefs about language (and languages) based on concerting actions and talk as part of living within the many scales of time (Madsen and Cowley, 2014). In this way, people gain subjective experience of temporality that enables culturally defined time to be used in action and perception. Human cognition is fundamentally diachronic.

## LANGUAGE AND THE CONCEPT OF LANGUAGE

“Language is first and foremost a temporal process whose dynamics and effects result from activity by two or more contextually situated individuals” (Fusaroli and Tylén, 2014, p. 1). In viewing

<sup>2</sup>The paper builds on the Browman and Goldstein’s (1992) articulatory phonology in that the verbal aspect of speech is traced to how people make and track phonetic gestures (based on timing of articulatory movements in a moving vocal tract). Not only is there massive support for the view (see Fowler, 2010) but it abolishes mental representation. More controversially, I treat this as compatible with Port’s (2010) demonstration that people use rich phonetic memory—a single utterance-act evokes many phonetic exemplars. Given the ecological basis of such views, the paper’s perspective is deemed distributed-ecological.

<sup>3</sup>Appeal to the code view is shorthand for reducing linguistic activity to the systematic use of linguistic forms (whether cashed out in behaviorist or cognitivist terms). Variations on the negative argument are set out in Harris (1981); Linell (2005); Love (2004), and Kravchenko (2007). Interactivism offers a related perspective that derives from the context of cognitive science (see Bickhard, 2009).

it as a temporal process, language is allowed to permeate the scales of experience that bind people into a living meshwork that connects verbal patterns, social resources (e.g., money) and acting in ways that change the natural world. Much depends on cooperation between people who re-enact culture by using, for example, talk, texts, and output from language-machines. When relying on digital systems, human interactivity links multi-scalar dynamics to Shannon information. Importantly, computing depends on probabilities and, as Taylor (2012) shows so clearly, human talk also relies on statistical patterns. For Taylor, this confirms that form-meaning mappings use a mental lexicon. Denying the existence of a mental lexicon, the paper presents language as symbiotic. While disembodied *concepts* sustain intuitions and have enormous pragmatic value, their basis lies in, not individual minds, but using linguistic co-embodiment to grasp and sustain wordings.

Reified cultural templates (“languages”) dominated 20th century linguistics. This was because, just as people believe in tables and trees, they believe that wordings correspond to abstract objects (“words”). While having immense social and practical value, any such approach relies on lay views of language, intuitions or, as Wittgenstein (1957) prefers, “agreement in judgments.” In fact, like all folk beliefs, such views rely on reports of multi-scalar human coordination. In arguing against reducing language to a verbal aspect, I stress the use made of embodiment. As with other human activity, language exploits sense-saturated coordination: like colors or numbers, verbal patterns index highly socialized lived experience. While phonetic gestures lack the constancy of digits, like both these and color display, they use a history of synergies based on interpersonal coordination. On the view presented here it is precisely its amenability to description as both embodied and phenomenal that renders language possible. This is because, using appearances, grammatical, and statistical aspects self-sustain as living persons pass away over historical time. Phonetic gestures shape events that are lived as lexico-grammatical, pragmatic and probabilistic. Like number, language thus functions at a population level—in ways that change over historical time. However, in focusing on how phonetic gestures connect verbal constraints to linguistic embodiment, I stress how language lives through people or, indeed, how verbal patterns self-propagate through human coordination. This appears in a scale where people draw on cultural products to co-construct situations and lived experience. Pursuing this, I show how social, moral and linguistic products constrain movements by living beings or, in the terms of the paper, linguistic embodiment.

A naturalized linguistics begins with acoustic and kinetic measures—not reports about wordings (or “words”). Rather than favor “signs” over “substance,” language is traced to movement. Tracing the said to phonetic gesture, the paper presents 750 ms of talk during which a person utters [a:bene]<sup>4</sup>. Analysis presents

detail that shows why dialogical events are irreducible to intentions, phonetic gestures or experience of *a priori* types. While the intention is plain and [a:bene] describes gestures made, the speech is symbiotic: events arise as verbal pattern constrains linguistic embodiment. Acting in a dynamic field, contributing to the flow of talk changes the layout of affordances. An act of utterance is joint activity which evokes what can be called linguistic “symbols.” In so saying, I echo Howard Pattee (see Pattee and Rączaszek-Leonardi, 2012), who, as a micro-physicist, traced language, computation and DNA to dynamics constrained by “symbols” qua self-organized measuring systems. On this view linguistic embodiment co-occurs with self-set control parameters. Yet, there is also a contrast between language and computation/DNA. Whereas computers and cells are self-managing (viz. they use phylogeny/metabolism and physical laws/human programming), language—and its symbols—depend on living human beings. Linguistic embodiment, coordinated action, speech, and hearing, is bodily movement that connects up central nervous systems. Skills in linguistic action arise as people couple control of airstream mechanisms, vocal folds, and articulatory tracts with phonetic gesture. Individuals act to connect metabolism with wordings and changing perceptions. As a result, skills in coordinating linguistic embodiment become enmeshed with what is learned from speech—people language by linking subjective experience with a grasp of the impersonal.

All embrained species exploit what Alain Berthoz (2012) calls *perçaction*. Action meshes with perception as, inseparably, people actively perceive the world. In humans, however, *perçaction* is transformed by language. Once utterance-acts are heard as reiterating patterns (as utterance-types) people mesh *perçaction* with experience of phonetic gestures (or wordings). They perceive objects and hear what the folk call “words.” The skills are learned; babies make sense of coordinated activity long before they make or track phonetic gestures. They come to use vocalizations to manage behavior and, in behaving, manage caregivers; they use rudimentary observing to manage how parties act, move and vocalize. By the second year of life, children co-construct “lived situations” and negotiate ways of “going on.” In parallel to Berthoz, wordings open up *observation*<sup>5</sup>. Talk-in-interaction exploits utterance-acts that prompt multiple interpretations. In what follows, I focus falls on, not what such acts achieve, but their embodiment. I show how people use local control parameters (Pattee’s symbols) and, without knowing what they are doing, evoke the linguistic, moral, and institutional resources of a community. By so doing, metabolism drives language. Embodiment connects the scales of time as people attune action with perception in activities that depend on coordinating finely regulated vocal and non-vocal expression:

- As *perçaction* language is sensorimotor activity that draws on/gives rise to rich phonetic memory (and its imagistic equivalents).

<sup>4</sup>This moment of first order languaging was chosen for two reasons: first, its striking inventiveness is relatively independent of what is said—a daughter shows exquisite skill in trying to head her mother off her conversational path. Second, for Cowley, the case exemplifies an utterance act that enacts meaning directly as prosody begets prosody.

<sup>5</sup>Didier Bottineau (personal communication) suggested this extension to Berthoz’s (2012) work.



- As observation, language is sensorimotor activity that draws on/gives rise to phonetic gestures (and their visible equivalents).

Phonetic gesture no more reduces to phonetic memory than rich memories of voice-speech-and-action suffice to explain verbal pattern. An utterance act like [a:bene] influences social behavior whose dynamics *also* invite phenomenal experience (that is amenable to verbal description). Indeed, emphasis on event-experience symbiosis parallels Darwin's (1998) observation that language-expression is part-natural and part-artificial; in terms of the *Descent of Man* (Darwin, 1989), the ability to moderate natural sounds co-functions with mimetic abilities or, in his terms, imitation. Playing down both linguistic form and its derived artifacts (especially, texts and text-like "systems"), I pursue the Darwinian intuition by tracing verbal pattern to how phenomenal experience uses linguistic embodiment.

Instead of reducing [a:bene] to linguistic forms, its acoustic and audible features can be attributed to the linguistic embodiment of Italians<sup>6</sup>. Indeed, those familiar with the "bel paese" will find themselves using this pattern of phonetic gestures in acts of utterance<sup>7</sup>. From a distributed perspective, the gestures that constitute [a:bene] enact linguistic flow, a form of perçaction that permeates Italian ways of life. While [a:bene] can be said unthinkingly, its uttering can also invite observation, construal, and interpretation. In the conversation described, it serves, in the main, as part of family events. By contrast, this paper subjects the same event to analysis. This is possible because [a:bene] is symbiotic. It is at once:

- Rich phonetic (and visible) activity that is integrated in a family conversation.
- The issuance of an actor-observer (a wording) that can elicit observer-actor response.

Because an utterance act is symbiotic, [a:bene] connects scales of human action. It can shape affective flow, enact a relationship and reflect on a person's family roles. This is possible because,

<sup>6</sup>Nigel Love (personal communication) argues that, somehow, Italian speakers share considerable knowledge about how verbal patterns (the second-order constructs of Italian) are used. Use of [a:bene] is, he thinks, inseparable from whether the parties believe it consists in, say, one or two words. He and I differ on two parallel points: I claim that people are, at once, caught up in flow and, at times, use observation; he emphasizes deliberate speech and interpretation. Accordingly, I emphasize that [a:bene] is heard by tracking phonetic gesture as part of whole-body experience; he emphasizes that, for Italians, any instance of [a:bene] evokes abstract units that pertain to a community's speech. For the same reason, Love makes no distinction between "wordings" (*qua* nonce phenomena) and verbal patterns (*qua* constructs that describe community practice).

<sup>7</sup>Bechtel (2008) offers a book length account of how the biology of cognition can be explored mechanistically. On a distributed perspective, the verbal aspect of language is a *mode of organization* that constrains the workings of bodily *parts* and the biocultural *procedures* that serve to insinuate language into a range of resources (i.e., by linking face-to-face interaction to both texts and various kinds of language-machines).

like airborne synapses (Steffensen, 2013), its dynamics continuously enable and constrain how parties feel and act. Potential meanings—and a wording—trigger and result in a flow of phonetic gestures. The projecting, speaking/listening and gesturing of [a:bene] is *direct* meaning making (see Cowley, in preparation). On a distributed view, [a:bene] exemplifies sense saturated coordination. In making and responding to an utterance act, a mother and daughter are less concerned with construals (or "form") than the richness of coordinated whole-body expression. At this instant, phenomenal experience frames events: far from inviting interpretation, the wording triggers subjective anticipation. It is not to be described by non-local meaning but, rather, by the particular sense it has for each party. While further discussed below (in Section How [a:bene] Functions), far from reducing to truth-conditional acts (see Oaksford and Chater, 1991), human action draws on essentially subjective probability estimations (Madsen, 2014). Bayes' theorem is a normative description whose probabilistic estimations describe, not a brain's workings, but how a person anticipates. Linguistic experience—and interpretation—thus builds on concerted embodiment. Since subjectivity is inherently social, it has a central role in cognition that extends beyond the body.

Subjectivity uses embodiment in all forms of perçaction (e.g., looking). Cognitive events such as those based on saccading or looking depend on time-scales at and below awareness<sup>8</sup>. In language too, action uses concerted looking as affect links utterance acts while setting off resonances and damping-effects. Coordination links looking and talking as people establish consensual domains (Maturana, 1980) or, alternatively, develop shared discursive and other practices. In short, during coordinated action, people also gain the skills of human observers. For example, they may come to see the point of actions or, indeed, to grasp the sense of various ways of displaying intentions and attitudes (and the microsocial order). As a result, they share beliefs about language: they *picture* forms and meanings as part of the world. Concerting bodies while attending to wordings thus links perçaction with forms (and *concepts*). Just as with monetary values or musical offerings, people draw on probabilistic information. In an attested, mundane example, a suspect can reasonably refuse to give a policeman information when he has admitted his guilt. In these circumstances, not naming confederates is licensed by the normative order (see Edwards and Potter, 2005). Since this is legitimate and intelligent, the management of social roles must be deemed "cognitive." An observer draws on circumstances to decide what he need not say. Much the same applies to the unsaid. Since language is symbiotic, observers may focus on acting like a policeman, speaking English or (not) sounding Liverpudlian. The language flow is cognitive in that it affects the unfolding of lived experience. Thus, as with refusing to give information, a person may speak in order to sound educated: the cognitive influences social judgment. Indeed, since practices link human embodiment with the verbal, the artificial becomes social. As

<sup>8</sup>Emphasis on the pico-scale or how vocal and other gestures are made is characteristic of the distributed view. Pioneered by Cowley (1994), the approach is increasingly influential (for example, Thibault, 2011; Uryu et al., 2014).

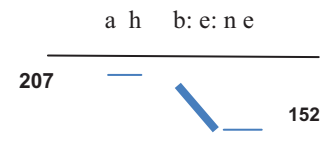
Hutchins (1995, 2014) shows, in settings like the cockpit, cognitive events are dominated by slow processes. In Wittgenstein's terms, the cockpit links language games with forms of life as pilots perceive *aspects* of the world. While cognition is enabled by embrained bodies, cultural resources prompt meaning making as an observer individuates what is important. By linking natural, social and material resources, the products of past events change later activity. A 4 million year history of self-directed representational acts (Donald, 1991) influences evaluation and learning as people self-improve and cooperate. Human motivations induce practicing and, thus, people gain fine control over vocalizations—the grounding of musical, mathematical and linguistic extensions to embodiment. The lived experience of language thus meshes with beliefs and conceptual tools that arise in a community's praxis.

## LINGUISTIC EMBODIMENT

Linguists link lay views with Saussure's authority to build linguistics on phenomenal experience of how phonetic gesture can be transcribed. Invoking abstract objects (e.g., words, generative grammar, conceptualization, or I-language), they invoke, not acts of speaking-while-hearing, but abstract types (e.g., utterances, sentences, constructions, usage-patterns). Emphasis on words and rules thus divides the linguistic from the non-linguistic. By fiat, sense-saturated coordination ceases to be language; linguistic embodiment ceases to be linguistic embodiment. If acknowledged, bodily dynamics are ascribed to modalities or paralinguistic and prosodic systems. Conversely, on a distributed perspective, language is sense saturated coordination whose neuro-social constraints sustain observing—activity in which wordings play a part (Cowley, 2011b). It emerges from the synergies and movements of linguistic embodiment that shape a flow of activity during which both macro and micro constraints affects what people do. People may speak and hear, for example, as part of a family: talk coordinates action (and vice versa). Linguistic embodiment has a role in constituting phenomenal experience that uses neural, microsocial, and cultural constraints. It enacts social activity and, paradigmatically, conversation. The claim is readily defended. First, talk is of pivotal concern to most people. Second, conversations ground skills that depend on language (e.g., flying planes, seduction, hunting). Third, talk is almost certainly the basis for the phylogenetic emergence of language—perçaction based on linking airstream mechanisms with control of the articulators. Not only is this a Darwinian view, but it allows cooperation and cognition to derive from coordinated movement. As with dance, music, and sport, language uses cultural and bodily constraints to social effect. Although literate people picture language as it “can be separated from its material expression” (Thibault, 2011, p. 2), this dubious surgery strips it away from lived experience. By leaving aside how people use utterance-acts, language is excised from the ecology: it is forgotten that “thinking depends as much on the environment of the thinker as it does on his or her brain” (Wells, 2006, p. 2).

Let us consider how co-activity draws on a single uttering of what can be described as [a:bene]. In offering a little detail about these 750 ms, I show two segments (see colons) exploit audible pico-scale lengthening. Whereas the initial “b” is striking, the long

[ε] vowel is typical of the speaker<sup>9</sup>. For the speaker's mother, the latter is thus unlikely to be perceptually salient; further, letter-spacing hints at other pico-scale timing (thus, “a h” is slow). Moreover, while hundreds of measures could be reported, the transcription picks out acoustic correlates of pitch on the first and last measurable vowels [Cowley's (1994); interchange (IF) and enjoining (EF) frequency]. Finally, there is a marked fall on the prominent syllable. (All measures are given in Hz).



As part of mother-daughter “thinking,” the utterance-act binds what precedes with what is likely to follow. Speaking [a:bene] is a “striking example of human inventiveness” (Cowley, 1994) where discursive practice uses human musicality. While the initial pitch (207 Hz) is near the daughter's norm (her mean IF is 215 Hz) and the prominent falling tone unexceptional (compare unmarked “oh good”), the act is striking. The lengthening of [b:] is an emblem of status that prefigures a “decisive” fall of half an octave (on *bene*). Indeed, this drops from about the daughter's mean IF a full standard deviation below her norm (152 Hz)<sup>10</sup>. Importantly, her speech rate *matches* her mother's almost perfectly: while her mother's rate is 240 ms per syllable, the daughter's is 250 ms. And, as emphasized in the Section “How [a:bene] Functions,” the “meaning” is also striking. Since thinking and social events are partly constituted by linguistic embodiment, details show more than sophisticated speech timing. Crucially, as phonetic gestures attune to her mother's voice, the voices create inter-individual patterns. Once these coordinated dynamics are noticed, one sees that, far from being paralinguistic, their musicality affects how the parties act, feel, and verbalize. Thus, [a:bene] is co-constructed or, in another idiom, an other-oriented act (Linell, 2009). Human dialogicality neither reduces to conventional use of form/meaning nor to typologies of speech act. No “pure” linguistic or cognitive model can show how mother and daughter coordinate. In Levinson's (1995) terms, this is interactional thinking: the sense of [a:bene] is enacted (i.e., not inferred from the context of “ah bene”).

Although amenable to separate analysis, the so-called modalities co-constitute speaking-while-listening or first-order languaging. Indeed, on the distributed-ecological view, interactivity shapes *experience* of talk. Its sense-saturated and normative

<sup>9</sup>While Lombard, the speaker usually uses a Northern version of standard Italian which is striking, in part, for marked use of geminate consonants: as a schoolteacher she often speaks in ways that contrast with the local dialect (where there are no geminates). Her unusual lengthening of the “b” on “bene” (as in central and Southern Italian speech) distances her from her mother's use of the dialect form “borsassa” (-assa is a widely used negatively charged suffix that denotes large size).

<sup>10</sup>Detail is offered in a paper (Cowley, in preparation) that uses this passage to show how linguistic embodiment (or first-order language) enacts *direct* meaning making. Details like those reported are ubiquitous in family conversations (see Cowley, 1994).

aspects enable bodies to coordinate peoples' feelings, thinking, and acting. Crucially, the dynamic field of experience involves more than phonetic gesture. Even if one leaves aside visible behavior, people use rhythmically based pico-scale dynamics based on modulating the air stream mechanism while making phonetic gestures. Saussure's error lay in dividing form from substance or, without argument, unzipping breathing from vibration of the vocal folds and how a changing vocal tract constrains phonetic gestures. Linguistic embodiment exploits a speaker's whole body movements. Indeed, it is a scandal that phenomenal experience is often blithely assumed to confirm the "reality" of a language-system. In ignoring linguistic embodiment, a focus on "form" echoes Cartesian dualism and debates about "representation." As a result, many 20th Century linguists ascribe language to the mind—echoing rationalist or empiricist debate. The radical nature of Chemero's view of embodiment is that agent-environment dynamics, not brains, become the basis of cognition. Language thus begins with pico-scale sensorimotor control that allows wordings (and inscriptions) to be derived from phonatory control, movement, and phonetic gesture. Indeed, it is a simple fact that people hear utterances as reiterations of the latter: given motor skills, the brain is trained through re-use<sup>11</sup>. Repetition of speech fragments in strategic social action attunes phenomenal experience to the movements that sustain and echo collective life. In lived experience, people thus draw on the past, invoke the future and exploit the impersonal. As with mother and daughter, meaning-making is direct and idiosyncratic. *Contra* Lyons (1977), language reduces to neither standardized, regularized nor decontextualized forms. While such models highlight the half-artificial, they overlook how bodies to move each other in a pico-scale. In fact, as with [a:bene], much uses what Abercrombie (1967) calls voice dynamics, continuous phonetic fluctuations that modulate the said. Prosody is thus redefined as "aspects of an individual's speech explicable neither in relation to word-based forms into which the speech can be analyzed, nor as part of the invariant auditory coloring that identifies an individual speaker's voice" (Cowley, 1994, pp. 6–7). As linguistic embodiment, meaning spreads as people exploit pitch, loudness, pace and so on. Like a cultural artifact or brain, human musicality serves as a cognitive resource. People show exquisite sensitivity to voices as they co-operate, talk, and manage emotion. Since voices serve in action, the results shape joint procedures and, thus, social events. In everyday life, dynamics connect wordings with circumstances, history and what is manifestly heard. The symbiotic coordination of "language" derives from, not verbal patterns, but bodily achievements: it is activity in which wordings play a part.

### HOW [a:bene] FUNCTIONS

Linguistic embodiment involves much more than phonetic gesture. In presenting a single "interact" (Linell, 2009), the case of [a:bene] shows how phonetic gesture can be subordinated to a pico-scale flow. However, it is also important to sketch how

coordinated thinking draws on the richness of lived experience (for more detailed description, see Cowley, 1998; Cowley, in preparation). In what follows, therefore, I place how the women act within a wider event trajectory. In so doing, I use transcription to build a narrative gloss:

M : Questi sono del tuo orto?	M. Are these from your garden?
A: Oeu me ne ha dato un po' la Rosa # ce ne ha dato una borsassa la Palmira	A: Mmm, she gave me a few did Rosa # she gave a ginormous bag to Palmira
M. Ah bene	M. Oh good

Briefly, having asked if the peas they are eating come from their garden, the daughter soon realizes that this is a mistake. Her mother begins to launch into a lament—they are not and, what's more, she only got a few while, worse still, Palmira was given a "ginourmous" bag of peas<sup>12</sup>. As the daughter says "ah bene," she attempts to control her mother. Thus, in terms of content-pattern OH GOOD is anomalous: giving a positive spin to events ("good that she gave you a bag"), the daughter seeks to deflect a train of thought. At the same time [a:bene] enacts how she feels—affect permeates gesture, wordings, and facial activity. Crucially, the act thus depends on pico-scale voice dynamics, the metabolic underpinning of language. It is in this sense that the importance of the phonetic detail lies in how the daughters' utterance-acts come to be suffused by her mother's co-presence. Broadly, the utterance-act is the thinking or, alternatively, speech enacts meaning.

Many linguists focus on how a person can perform the same acts over and over again. In prioritizing what Colunga and Smith (2008) term the *problem of stability*, they reduce prosody to patterns and, overlooking voice dynamics, emphasize discursive, and intra-utterance regularities that are said to generate rhythmic and tonal patterns (e.g., how tone groups map onto prominences and patterns of pitch, duration, and loudness associated with marked syllables). While said to be "communicative" (whatever that means), models of prosodic systems are powerless in clarifying function. This is because, in focusing on the recurrent, they make pico-scale voice dynamics "paralinguistic." The models disembodiment language by separating it from experience. On the distributed-ecological view, by contrast, the dynamics of language flow shape how parties "decide what to say." As shown in fine-grained analysis, extensive use is made of rapid bodily attunement, improvisation, and lived relationships. Indeed, this enacts most of what we call emotion, attitude and how people vary the deliberation (and inhibition) of social life. Of course, as an embrained species, humans use learning and repetition; yet, as observers, individuals also use particularities. Language exploits interactivity or, for Colunga and Smith, dynamics permitting us "to smartly do novel things that integrate the stabilities

<sup>11</sup>As multiscalar activity, language unites many kinds of network. It is likely that, with skills in speaking, listening and otherwise drawing on linguistic resources, people "rewire" or sculpt their brains as suggested by Anderson's (2010) hypothesis of neural reuse.

<sup>12</sup>As noted above "borsassa" is a dialect form that gives negative connotations to an object of some size. In order to render something of this, I have translated it as "ginormous."

of past experience with the idiosyncrasies of the moment” (2008, p. 175).

Metabolism reasserts itself as the talk continues through the 750 ms during which the mother finds a way of going on. In so doing, she pointedly pays no attention to the phonetic gesturing. Far from speaking up on the positive, she speaks as if her daughter had *said* nothing. Indeed, using the voice dynamics, she redoubles her complaint. Far from relying on interpretation, this is affective vocal expression: the parties co-enact flexible, adaptive behavior that alters neural processes, sets up priors, and shapes subjective experience. Not only do they know what Everett (2012) calls “the joy of language” but their speaking and moving is thick with sense. Phonetic gestures intermesh as people take each other’s measure—sensing how they are assessed. Interactivity affects feeling, thinking and acting in scales that are more rapid than phonetic gestures and audible shifts in tones of voice. While a micro-scale highlights what we articulate (syllables, tone-units/phrases, and utterances), much depends on people whose resonating voice dynamics set off sound patterns with variable probabilities<sup>13</sup>. Thus, attention shifts to functions that characteristically occur in 50–200 ms range: in this pico-scale interpersonal synergies are ubiquitous. Living language is grounded in, not wordings, but bodily movement. Of course, linguistic embodiment is no more than a necessary part of language—at times people choose to say, write and construe things with much more deliberation. That too must be considered.

## THE ELEPHANT IN THE ROOM

There is nothing exceptional about tracing human activity to *how* cognition emerges in ecological space and ecological time. Not only is this also true for sport and dance but, not surprisingly, it is also applicable to problem solving. As shown in experimental work, people use interactivity, sense-saturated coordination, that mediates embrained bodies, motor actions and artifacts (Ball and Litchfield, 2013; Vallée-Tourangeau and Vallée-Tourangeau, 2014). In talk, much depends on pico-scale dynamics (Cowley, 1994, 2010; Thibault, 2011; Steffensen, 2013; Cowley, in preparation). By implication, linguistic embodiment enables living subjects and communities to use musicality as language arises *beyond* the body. Next, therefore, I turn from lived dynamics to the non-metabolic. The point is that embodied musicality suffuses what are heard as physically-based patterns like “*ah bene*.” Pico-scale events shade phonetic gestures as utterance-acts become audible as utterances *of* something. To repeat the mantra, language is activity in which wordings play a part. However, one must be careful: wordings emerge as those familiar with, say, Italian forms of life hear phonetic gestures. While reported as linguistic types, these function, not as “forms,” but nonce events. Like numbers or colors, phonetic gestures are resources used in action. Indeed, linguistic embodiment gains its power from managing how one

languages in ways that evoke a community’s speech patterns. In distributed terms, second-order constructs (i.e., lexical, semantic, phonological, morphological, syntactic, pragmatic, and stylistic patterns) constrain what people do, feel and, thus, think. Further, much of the time, of course, people act “mindlessly”: as a result of a life-history, they adopt beliefs in the reality and power of wordings. In literate communities, these become associated with inscriptional forms such as “*ah bene*.”

As activity in which wordings play a part, language becomes insinuated into almost all areas of human life. In more formal kinds of talk, worship, text-messaging, for example, the verbal dominates. Indeed, without skill in perceiving wordings, there can be no human observers—and no language. In ontogenesis, skill in perceiving phonetic gestures as wordings arises from experience of interacts (Linell, 2009). On a distributed-ecological view, the perceptual skill arises in zooming out of a full-bodied situation. Thus, while utterance-acts link metabolism to local practices, they also come to be heard in a particular sense. While these can be ascribed to intentions, this is a second-order model. In the case of [a:bene], it is a fact that, for Italians, the act evokes hearings of “*ah bene*.” While further discussed below, I stress only that prompts and probes exploit meaning potentials (Linell, 2009). The daughter finds herself moved to stop her mother: her polyphonic [a:bene] links Italian ways of managing conflict with a mother-daughter relationship; it enacts her concerns, her perception of her mother and, indeed, a wish to be “positive.” The holistic nature of [a:bene] links neural synergies with habit as phonetic gesture binds various time-scales into a lived situation. As a chunk of behavior, its verbal aspect resembles a *real-pattern* (Dennett, 1991; for development, see Ross, 2000). Functionally, the utterance-act triggers pattern recognition that bears the hallmark of reward-based learning. Dennett (1991) compares this with how von Neumann machines use zip files; instead of applying a value to every bit of information, programs use compression. By analogy, hearing [a:bene] calls up senses that *may* be valid. Voice dynamics constrain experience and thus prompt anticipation. In the flow of talk, wordings—the phenomenal experience—index ways of proceeding. Just as a computer needs users or a cell an environment, language demands an observer. Like a zip-file, [a:bene] evokes compressed information if, and only if, a person finds a *perspective* from which the pattern can be used.

Wordings exist as they are perceived. As phenomenal experience, they carry a particular sense which has little in common with either the meanings or forms that linguists ascribe to verbal patterns (second-order constructs). Yet, as phenomenal experience, a wording calls up historically derived affordances as embodiment and circumstances prompt ways of going on. People use partial control over phenomenal experience to concert their movements. As in music or dance, they jointly manage how the rapid scale of linguistic embodiment resonates with historical events as, in pico-scales, neurophysiology enacts coordinated activity. The symbiotic nature of language thus connects two kinds of reward-based learning. On the one hand, like rats and wolves, people learn both individually and socially: they use exposure, probability judgments and rewards. On the other, people also use wordings in learning by observing. They notice,

<sup>13</sup>In many species, Hinde (1979) informs us, much depends on both the probabilities with which individuals do things and, at times, on how they are done. Though primate grooming is the classic case, birds also exploit both micro- and pico-scales—a clear case is that of how Alex the parrot taught himself to say “spool” (see Pepperberg, 2007).



and prompt each other to notice, aspects of the world. They use anomalies and turns of phrase; they pick up on attitudes, intentions and hidden parts of the environment (e.g., urgency, potential for use). Observation-based learning, it seems, is specifically human. It enables the use of abstract qualities (e.g., bene, red, strong) to connect the more intuitive to the more deliberate in, for example, telling a joke or proposing another round of beer. Phenomenal experience thus offers valuable ways of gauging how to act in the circumstances. Although dealing with [a:bene] is largely a matter of co-embodiment, the example represents relatively automatic talk. On many occasions, human intercourse depends on much closer attention to wordings. In stories, for example, wordings dominate narration. In other settings, they are even more basic—for example, they sustain writing-systems.

Wordings can be used in literal, poetic and hypothetical ways; in contrast to dealing with [a:bene], people can choose to rely on the “words that are actually spoken.” As argued elsewhere, much can be gained by taking a distance from the flow of talk by focusing on phonetic gestures and seeking to repress the unsaid. In so doing, people learn to take a language stance (Cowley, 2011a). While talk fluctuates between greater reliance on automaticity (as in the 750 ms of [a:bene]) and more careful use of a language stance, there are also intermediate modes of acting and attending. People can and do shift emphasis between the said, what they say and, indeed, the unuttered (“silent thoughts”). In its impersonal aspect, language opens up *other people's* experience (see Cowley, 2014). For example, if I allude to Endel Tulving, informed readers may evoke mental time-travel: in its verbal aspect, the inscription *Endel Tulving* compresses ways of going on. By opening up an impersonal past (using “priming”), the reader's embodiment anticipates what is likely to follow. Skills connect automated perception with how observers construe circumstances. Since wordings draw on (or resemble) real patterns, particulars can be perceived as types or “sames.” While real-patterns evoke reports of hearing (and seeing inscriptions), they are equally likely to affect how mother and daughter concert their speech. As they do so, people enact and display experientially-based modes of practical understanding<sup>14</sup>.

People use fine phonetic information to perceive, not an utterance act's particulars, but “salient” details that serve to anticipate circumstances. In making rapid judgments about [a:bene], both parties use compressed information associated with statistical experience. Far from using all available “information,” people note sensitivity to unmet expectations (or expected standards). As a result, combining real-patterns with voice dynamics (not to mention movements of face and gesture) contrasts with deliberate use of historically derived patterns. Whereas prosody is usually managed by ear, observers can shift attention between the said,

the words that actually spoken, and how they hope to sound. Wordings offer a degree of individual control: one can even speak impersonally. While consistent with folk wisdom, psychology, linguistics, and cognitive science are guilty of overlooking such phenomena. Avoiding dynamics, they trace phenomenal experience to skills in monitoring the said, projecting what is likely to be appropriate, and choosing how to “come over.” In turning to the verbal aspect of language, as opposed to how [a:bene] is co-embodied, groups can be said to draw on a “systemic” meaning potential (Halliday, 1985). Since this echoes the slow scales of history, to the extent that a subject grasps this potential, he or she can use wordings to recalibrate acting, thinking, feeling other forms of self-display.

Once language is recognized as symbiotic, one begins to rethink the verbal. First, since perceived wordings can be repeated and analyzed as parts and procedures, phonetic gestures allow both skilled hearing and strategic use of utterance-acts. From a distributed perspective, these shape perçaction—language is skilled action. By implication, mechanisms beyond the brain function as people use phenomenal experience to coordinate activity. Language thus sustains the people of a social meshwork: the verbal aspect of talk uses a time scale where people enact organized social practices (Enfield, 2011). However, [a:bene] also arises as a mother's voice moves her daughter to use cultural resources. Though hearing the same phonetic gestures, each party reacts differently. Linguistic symbiosis allows social factors to work through people whose interactions shape circumstances, relationships and the Italian life. Thus, while connotational meaning is precise, both women draw on experience of tens of thousands of similar cases: these constitute a fuzzy denotational meaning (OH GOOD). At this moment, however, this matters little. Events link the emotional interplay to circumstances and the mother ignores her daughter's “positive” move.

## VERBAL PATTERNS ARE PARTLY SHARED

Since people have similar experiences, verbal patterns come to be partly shared. Further, given rich phonetic (and visible) dynamics, circumstances influence how people assess and manage each other. In construing a single act of utterance, the women draw on how countless hearings of phonetic gestures enrich experience of both Italian forms of life and their relationships. In so doing, they attend to what can be written as “ah bene.” This verbal pattern can be described at the population or corpus level; even here, however, it is not purely verbal in that, among other things, it evokes attitudes and probabilities. In Italy, the pattern's penumbra thus sets off relatively predictable effects. When one experiences a wording that can be rendered as [a:bene], perception connects up with scales of time and thus cognition beyond the body. Far from using a shared lexicon, mental or social, the parties rely on making and tracking phonetic gestures. Given the symbiotic nature of human language, no more understanding is required. Rather than ascribe a causal function to verbal patterns, they are second-order constraints on lived experience. Like numbers or colors, wordings link indices of past events to circumstances or, in Maturana's terms, trigger connotations. Far from using tokens “in the head,” as (or like) real-patterns, wordings trigger events over

<sup>14</sup>The Dennett-Ross view of real-patterns recalls Pattee's symbols in that these too are measures that constrain dynamics in observable ways. However, in generalizing from DNA and computation to linguistics, Pattee views linguistic symbols as brain internal. By contrast on the Ross-Dennett view, while voice dynamics are public, cultural real-patterns (including wordings) evolve in populations: they encompass language, traditions, music, money, etc.—not to mention associated procedures and institutions.

which a person exerts some control. During talk, parties manage and inhibit their promptings, in part, by “choosing” what to say. People rely on activity in which wordings play a part to grasp and alter what they and others perceive, mean and say. Familiar ways of speaking/acting and associated probabilities offer some control over actions. The perceived—wordings, colored objects or analog/digital “representations”—need do no more than evoke an iterable pattern. Skills in conjuring up the audible or visible aspect of language thus ground what Wittgenstein (1980) came to call “certainty.” Human modes of life, and living bodies, enable one to make and accept utterance-acts such as: “My name is NN” or “I have never been to Bulgaria.” Indeed, people can even play philosophical games by making explicit judgments of whether or not it is appropriate to say, “That is a tree.” Crucially, such claims become transparent only to the perspective of an informed observer: in themselves, they are trivial. It may be *true* that a philosopher is pointing at a tree and saying what it is and yet, at the same time, appear quite pointless to act this way. If *that* is to be explained, human forms of life need to be traced back to interactivity: one must show how people become observers who, to an extent, share a perspective on phenomenal experience. By hypothesis, this is possible because of symbiosis between linguistic embodiment and the verbal. Utterance-acts evoke wordings that, via phonetic gestures, allow people to connect their co-embodiment with impersonal and shifting verbal patterns.

Language is typically approached from an observer’s perspective. In the terms used here, a person takes a language stance by attending to behavior, or products of behavior, that derive from phonetic gestures. From this stance, language can be discussed, re-described, formalized and, in slow time scales, transformed. In history, utterance-acts change in parallel with conceptual evolution. Slow events constrain how perception shapes individual experience: thus, while non-linguistic “thinking” appears in many embrained species (see Bermúdez, 2003), humans use new kinds of thought. Drawing on phonetic gesture, children hear wordings and, eventually, develop skills based on the language stance. This enables preferences to be connected with beliefs as children learn how, in various settings, things are done. As a result, they can develop ways of competing, coordinating and cooperating. They may discover, for example, that the same phonetic gestures allow a shirt, hair, or wine to be called “red.” In so doing, they gain access to the concept’s impersonal aspect (“redness”); however, like the taste of wine, the smell of hair and the look of a shirt, this is also subjective. Whilst the phenomenal moors certainty, social encounters permit a slow accumulation of conceptual understanding. Utterance-acts call up experience based estimations of how wordings will be heard or projected meaning-potential. As a person orients to others, they approach wordings as observers. While influencing talk, they also elicit construals as, over time, each person gains a sense of semantics. In the philosopher’s garden, people match judgments by connecting well-timed pointing to, for example, saying, “That is a tree.” Of course, philosophers often erroneously seek to “explain” this referential relation. Yet, on the distributed view, though languaging is subjective or connotational (but not private), communities also exploit collective and denotational meaning. As shown by the mother and daughter,

linguistic embodiment arises as concerted movements and voice dynamics shape a flow of social events. In the case described, while giving little attention to the words actually spoken (“ah bene”), the parties re-enact their relationship—as only Italians can. It is by virtue of the symbiotic nature of language that the parties grasp value-labels (e.g., BENE/GOOD) that serve to sustain a cultural lineage and, in so doing, a bundle of social practices.

## HUMAN COGNITION AND THE SCALES OF TIME

The paper shows that, in studying cognition, one can ask *how* embodiment functions. Applied to language, talk is seen as intrinsic to a history of interactions that connect sensorimotor activity, brains, and forms of human artifice. As in the mother-daughter exchange, linguistic embodiment connects parties across the scales of time. In this sense, language is symbiotic or, simply, linguistic embodiment connects movement with experience of wordings. There are, at least, two reasons for which the claim is non-trivial. First, symbiosis permits coordination between and within individuals: as this occurs, the women relate to each other. Using their voices, they concert expression and, thus, evoke meaning potentials that observers associate with verbal patterns. This leads to the second point. An evolutionary history links phonetic gestures with phenomenal experience such that wordings connect subjective experience with an impersonal aspect. During talk, people engage with each other and, to varying extents, use language reflectively. By taking a language stance, they can contribute to (or inhibit) debate about verbal and conceptual patterns. Affect and whole-body dynamics thus link human vocalization with impersonal experience that grants access to species-specific resources. Given linguistic symbiosis, cultural products can be re-used at later times (Hollan et al., 2000). In contrast to other primates, members of *homo sapiens sapiens* link embodiment to artifice as they shape relationships, institutions, and the cultural ecology.

The polyphony of language not only grants access to language machines and texts but it makes individuals part of a cultural heritage. People use this to draw on compressed information that pertains to the world beyond the body. This arises because, like zip files, phonetic gestures facilitate mental time-travel by calling up both personal and impersonal experience. As Merlin Donald saw, the evolution of a cognitive-cultural network transformed human intelligence: it made it possible to create and deflate possible worlds. Once verbal patterns are insinuated into embodiment, language can self-sustain in a collective or population domain. People live in language as they coordinate within a meshwork of bodies that link ecological space with ecological time. Children participate in distributed cognitive systems and, as Giere (2004) insists, they do so as human agents. By hypothesis, human cognition is transformed as linguistic symbiosis allows them to develop the skills of observers. Unlike other primates, the mother and daughter orient to [a:bene] and, in the space of 750 ms, co-construct a situation that re-enacts their relationship. Polyphony enables them to act strategically and cooperatively. By hypothesis, they use compressed Shannon information that is phonetic (e.g., durational), verbal (e.g., based on usage and discourse practices), and conceptual (e.g., exploiting semantic attributions). Using phonetic gestures, wordings sustain the ways of speaking used

in Italian communities. Humans thus use cultural resources to construct new kinds of temporal experience. It is because meaning making is ecosystemic that, for example, astronomers can explore the history of the universe. Further, as affective, interpreting beings, people link embodiment with wordings. Distributed systems enable living communities to build collective memories and specify possible futures. Brains are recalibrated as people use priors that sustain reasoning. Individuals become living subjects as embodiment connects people within a social meshwork.

Humans are strange. In most species that use social learning—rats, wolves, or elephants—collective intelligence centers on individuals. In humans, by contrast, much depends on an evolving cultural or impersonal domain. This is because, while based in embodiment, activity draws heavily on reports of how wordings contribute to experienced phenomena. People connect perception with skills based on using a language stance: they strategize, refine values, and develop social practices. For this reason, acknowledgement of linguistic symbiosis offers much to radical embodied cognitive science. Mental content is replaced by treating language as activity in which wordings play a part. People use embodied coordination together with phenomenal experience of wordings. They need, not neural representation, but dispositions that link neural resources to the world beyond the body. Human intelligence exploits diachronic agent-environment dynamics. Given linguistic symbiosis, perception enables human individuals to use wordings in observation. Individual lives can be regulated around impersonal resources. Not only do we conform to social practices, norms and beliefs but, crucially, artifice, and wordings enable individuals to self-configure. As Heidegger saw, *experience* of language makes humans distinct. As we exploit the accountable, meaning potentials arise—people come to believe in languages, minds, and mental content. Social life uses such beliefs, above all, to draw on past—mythical, lived, told and impersonal. Its imagined outcomes can thus be put to use in making futures. This allows experience to be recalibrated as when, for example, philosophers pursue enquiry by pointing at plants in a garden while uttering variations on “That is a tree.” Remarkably, language lays down markers for possible futures as people navigate ecological space and ecological time. Drawing on interactivity, history and wordings, each one of us becomes a living subject who, for a moment, exerts some control over who and what we become.

## ACKNOWLEDGMENTS

Special thanks to two anonymous referees for scrupulous attention to how the paper could be clarified, expanded and developed. I do hope that they prefer this version. The new draft also uses Nigel Love’s integrationalist counter arguments. His provocative view has led me to stress that wordings are, not verbal patterns, but nonce events: they arise in a dynamic field that is constituted, in part, by the experience of making and tracking phonetic gestures.

## REFERENCES

- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* 33, 245–266. doi: 10.1017/S0140525X10000853
- Ball, L. J., and Litchfield, D. (2013). “Interactivity and embodied cues in problem solving, learning and insight: further contributions to a “theory of hints,”” in *Cognition Beyond the Brain: Interactivity, Computation and Human Artifice*, eds S. J. Cowley and F. Vallée-Tourangeau (London: Springer), 223–240.
- Bechtel, W. (2008). *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge.
- Bermúdez, J. L. (2003). *Thinking Without Words*. Oxford: Oxford University Press.
- Berthoz, A. (2012). *Simplicity: Simplifying Principles for a Complex World*. Cambridge, MA: Yale University Press.
- Bickhard, M. H. (2009). The interactivist model. *Synthese* 166, 547–591. doi: 10.1007/s11229-008-9375-x
- Browman, C. P., and Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica* 49, 155–180.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
- Colunga, E., and Smith, L. B. (2008). Flexibility and variability: essential to human cognition and the study of human cognition. *New Ideas Psychol.* 26, 174–192. doi: 10.1016/j.newideapsych.2007.07.012
- Cowley, S. J. (1994). *The Place of Prosody in Italian Conversations*. Unpublished PhD dissertation, University of Cambridge.
- Cowley, S. J. (1998). Of turn-taking, timing and conversations. *J. Psycholinguist. Res.* 27, 541–571. doi: 10.1023/A:1024948912805
- Cowley, S. J. (2010). “Beyond symbols: interaction and the enslavement principle,” in *New Adventures in Language and Interaction*, ed J. Streeck (Amsterdam: John Benjamins), 47–70.
- Cowley, S. J. (2011a). Taking a language stance. *Ecol. Psychol.* 23, 185–209. doi: 10.1080/10407413.2011.591272
- Cowley, S. J. (2011b). *Distributed Language*. Amsterdam: John Benjamins.
- Cowley, S. J. (2014). The integration problem: interlacing language, action and perception. *Cybern. Hum. Knowing* 21, 53–65.
- Cowley, S. J., and Vallée-Tourangeau, F. (2013). “Systemic cognition: human artifice in life and language,” in *Cognition Beyond the Brain: Computation, Interactivity and Human Artifice*, eds S. J. Cowley and F. Vallée-Tourangeau (Dordrecht: Springer), 255–273.
- Darwin, C. (1998). *The Expression of the Emotions in Man and Animals*. Oxford: Oxford University Press.
- Darwin, C. (1989). *The Descent of Man and Selection in Relation to Sex*. London: Pickering.
- Dennett, D. C. (1991). Real patterns. *J. Philos.* 88, 27–51. doi: 10.2307/2027085
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.
- Edwards, D., and Potter, J. (2005). “Discursive psychology, mental states and descriptions,” in *Conversation and Cognition*, eds H. Te Molder and J. Potter (Cambridge, MA: Cambridge University Press), 241–259.
- Enfield, N. J. (2011). “Sources of asymmetry in human interaction: enchrony, status, knowledge and agency,” in *The Morality of Knowledge in Conversation*, eds T. Stivers, L. Mondada, and J. Steensig (Cambridge, MA: Cambridge University Press), 285–312.
- Everett, D. L. (2012). *Language: the Cultural Tool*. London: Profile.
- Fowler, C. A. (2010). Embodied, embedded language use. *Ecol. Psychol.* 22, 286–303. doi: 10.1080/10407413.2010.517115
- Fusaroli, R., and Tylén, K. (2014). Linguistic coordination: models, dynamics and effects. *New Ideas Psychol.* 32, 115–117. doi: 10.1016/j.newideapsych.2013.03.002
- Giere, R. N. (2004). The problem of agency in scientific distributed cognitive systems. *J. Cogn. Cult.* 4, 759–774. doi: 10.1163/1568537042484887
- Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Harris, R. (1981). *The Language Myth*. London: Duckworth.
- Hinde, R. A. (1979). *Towards Understanding Relationships*. London: Academic Press.
- Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput. Hum. Interact.* 7, 174–196. doi: 10.1145/353485.353487
- Hutchins, E. (1995). How a cockpit remembers its speeds. *Cogn. Sci.* 19, 265–288. doi: 10.1207/s15516709cog1903\_1
- Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philos. Psychol.* 27, 34–49. doi: 10.1080/09515089.2013.830548

- Kravchenko, A. V. (2007). Essential properties of language, or, why language is not a code. *Lang. Sci.* 29, 650–671. doi: 10.1016/j.langsci.2007.01.004
- Levinson, S. C. (1995). “Interactional biases in human thinking,” in *Social Intelligence and Interaction*, ed E. N. Goody (Cambridge, MA: Cambridge University Press), 221–260.
- Linell, P. (2005). *Written Language Bias in Linguistics, 2nd Edn.* London: Routledge.
- Linell, P. (2009). *Rethinking Language, Mind, and World Dialogically: Interactional and Contextual Theories of Human Sense-Making.* Charlotte, NC: Information Age Publishing.
- Love, N. (2004). Cognition and the language myth. *Lang. Sci.* 26, 525–544. doi: 10.1016/j.langsci.2004.09.003
- Lyons, J. (1977). *Semantics*, Vol. 1. Cambridge, MA: Cambridge University Press.
- Madsen, J. K. (2014). Approaching Bayesian subjectivity from a temporal perspective. *Cybern. Hum. Knowing* 21, 98–112.
- Madsen, J. K., and Cowley, S. J. (2014). Living subjectivity: time-scales, language and the fluidity of self. *Cybern. Hum. Knowing* 21, 11–22.
- Maturana, H. R. (1980). “The biology of cognition,” in *Autopoiesis and Cognition: the Realization of the Living*, ed H. R. Maturana (Dordrecht: Springer), 5–58.
- Maturana, H. R., and Varela, F. J. (1994). *The Tree of Knowledge: the Biological Roots of Human Understanding.* Boston, MA: New Science Library.
- Oaksford, M., and Chater, N. (1991). Against logicist cognitive science. *Mind Lang.* 6, 1–38.
- Pattee, H. H., and Rączaszek-Leonardi, J. (2012). *Laws, Language and Life: Howard Pattee’s Classic Papers on the Physics of Symbols with Contemporary Commentary.* Dordrecht: Springer.
- Pepperberg, I. M. (2007). Grey parrots do not always ‘parrot’: the roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Lang. Sci.* 29, 1–13. doi: 10.1016/j.langsci.2005.12.002
- Port, R. F. (2010). Rich memory and distributed phonology. *Lang. Sci.* 32, 43–55. doi: 10.1016/j.langsci.2009.06.001
- Ross, D. (2000). “Rainforest realism: a dennettian theory of existence,” in *Dennett’s Philosophy: a Comprehensive Assessment*, eds D. Ross, A. Brooks, and D. Thompson (Cambridge, MA: MIT Press), 147–168.
- Ryle, G. (1949). *The Concept of Mind.* New York, NY: Barnes and Noble.
- Steffensen, S. V. (2013). “Human interactivity: problem-solving, solution-probing and verbal patterns in the wild,” in *Cognition Beyond the Brain: Interactivity, Computation and Human Artifice* eds S.J. Cowley and F. Vallée-Tourangeau (London: Springer), 195–221.
- Taylor, J. R. (2012). *The Mental Corpus: How Language is Represented in the Mind.* Oxford: Oxford University Press.
- Thibault, P. J. (2011). First-order languaging dynamics and second-order language: the distributed language view. *Ecol. Psychol.* 23, 210–245. doi: 10.1080/10407413.2011.591274
- Uryu, M., Steffensen, S. V., and Kramsch, C. (2014). The ecology of intercultural interaction: timescales, temporal ranges and identity dynamics. *Lang. Sci.* 41, 41–59. doi: 10.1016/j.langsci.2013.08.006
- Vallée-Tourangeau, G., and Vallée-Tourangeau, F. (2014). The spatio-temporal dynamics of systemic thinking. *Cybern. Hum. Knowing* 21, 113–127.
- Wells, A. (2006). *Rethinking Cognitive Computation: Turing and the Science of Mind.* Basingstoke: Palgrave Macmillan.
- Wittgenstein, L. W. (1957). *Philosophical Investigations*, 2nd Edn. Oxford: Blackwell.
- Wittgenstein, L. W. (1980). *On Certainty.* Oxford: Blackwell.

**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.

Received: 02 May 2014; accepted: 08 September 2014; published online: 02 October 2014.

Citation: Cowley SJ (2014) Linguistic embodiment and verbal constraints: human cognition and the scales of time. *Front. Psychol.* 5:1085. doi: 10.3389/fpsyg.2014.01085 This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology.

Copyright © 2014 Cowley. 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.





# Embodied niche construction in the hominin lineage: semiotic structure and sustained attention in human embodied cognition

Aaron J. Stutz<sup>1,2\*</sup>

<sup>1</sup> Division of History and Social Sciences, Oxford College of Emory University, Oxford, GA, USA

<sup>2</sup> Department of Anthropology, Emory University, Atlanta, GA, USA

## Edited by:

Guy Dove, University of Louisville, USA

## Reviewed by:

Serge Thill, University of Skövde, Sweden

Guy Dove, University of Louisville, USA

## \*Correspondence:

Aaron J. Stutz, Division of History and Social Sciences, Oxford College of Emory University, 810 Whatcoat Street, Oxford, GA 30054, USA  
e-mail: astutz@emory.edu

Human evolution unfolded through a rather distinctive, dynamically constructed ecological niche. The human niche is not only generally terrestrial in habitat, while being flexibly and extensively heterotrophic in food-web connections. It is also defined by semiotically structured and structuring embodied cognitive interfaces, connecting the individual organism with the wider environment. The embodied dimensions of niche-population co-evolution have long involved semiotic system construction, which I hypothesize to be an evolutionarily primitive aspect of learning and higher-level cognitive integration and attention in the great apes and humans alike. A clearly pre-linguistic form of semiotic cognitive structuration is suggested to involve recursively learned and constructed object icons. Higher-level cognitive iconic representation of visually, auditorily, or haptically perceived extrasomatic objects would be learned and evoked through indexical connections to proprioceptive and affective somatic states. Thus, private cognitive signs would be defined, not only by their learned and perceived extrasomatic referents, but also by their associations to iconically represented somatic states. This evolutionary modification of animal associative learning is suggested to be adaptive in ecological niches occupied by long-lived, large-bodied ape species, facilitating memory construction and recall in highly varied foraging and social contexts, while sustaining selective attention during goal-directed behavioral sequences. The embodied niche construction (ENC) hypothesis of human evolution posits that in the early hominin lineage, natural selection further modified the ancestral ape semiotic adaptations, favoring the recursive structuration of concise iconic narratives of embodied interaction with the environment.

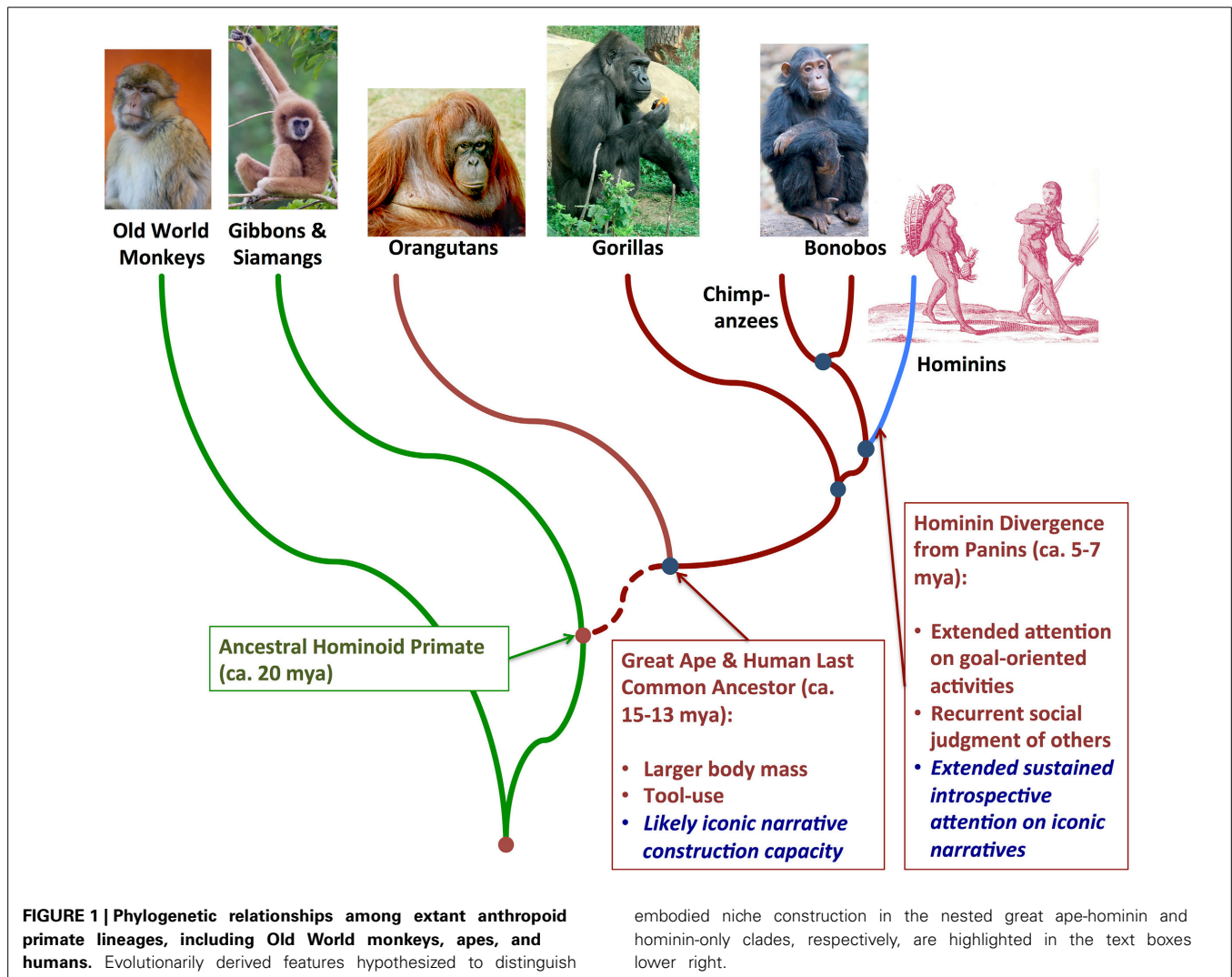
**Keywords:** embodied cognition, niche construction, hominin adaptation, co-evolution, iconic narrative, semiotics, bipedalism

## INTRODUCTION

Concepts of embodied cognition have been intensively developed and extensively treated in psychology, neuroscience, cognitive science, and the philosophy of cognition over the past 30 years (Humphrey, 1992; Clark, 1993, 1998, 2008; Damasio, 2003, 2008; Gallese and Lakoff, 2005; Rowlands, 2006; Dove, 2011). Although popular accounts have reached scholarly specialists in human evolution (e.g., Coward and Gamble, 2008), embodied cognition seems to be in its conceptual and theoretical infancy in paleoanthropology. Barton's phylogenetic comparative work—focusing on the neuroanatomical support for sensory-motor simulation in the primate order—stands out as a promising exception (Barton, 2012). Paleoanthropology's engagement with embodied cognition research could provide a needed comparative evolutionary perspective on what is unique about how the human body and body-environment interaction shape, facilitate, or constrain cognition.

Such an evolutionary approach highlights some fundamental questions. What aspects of embodied cognition might be

relatively evolutionarily primitive among terrestrial vertebrates? What aspects might be relatively derived among the (phylogenetically nested) primate, anthropoid, ape, and hominin lineages, respectively (**Figure 1**)? I suggest that recent theoretical developments in paleoanthropology, evolutionary biology, and ecology are especially amenable to—and would be scientifically strengthened by—embodied cognition research. Consideration of evolutionary niche construction dynamics (Odling-Smee et al., 2003), in particular, draws our attention to an emerging theoretical intersection, where we can explore how unique human phenotypes—including linguistic communication and symbolic representation—may have co-evolved with a niche significantly constituted by embodied interfaces (1) between the somatic and extrasomatic environments and (2) within the somatic environment itself. The embodied niche construction (ENC) hypothesis aims to complement but achieve more comprehensive explanation of language and sociality than recent proposals in paleoanthropology and linguistics (Deacon, 1998; Jackendoff, 1999; Dunbar, 2003, 2009; Tomasello, 2008). The ENC hypothesis



states that human capacities for symbolic mental representation, symbolic communication, and social cooperation emerged over the past ca. 5–7 million years through dynamic co-evolution with embodied cognition and environmental interaction. This occurred within a rather distinctive, very dynamically evolving ecological niche: one that is not only generally terrestrial in habitat—while being flexibly and extensively heterotrophic in food-web connections—but also defined by semiotic, structured and structuring embodied interfaces between the individual organism and the extrasomatic environment.

## BACKGROUND: TOWARD INTERDISCIPLINARY COMMON GROUND

This hypothesis article aims—perhaps naively—to prevail over potential interdisciplinary misunderstanding. I am a biocultural anthropologist (cf. Stutz, 2012) presenting a theoretical speculation about human cognition—and to a specialized cognitive science readership, at that. On the one hand, this is certainly like the student trying to lecture the teacher. On the other hand, it is precisely because I am aware of the importance of cognitive

science for studying the embodied dimensions of human experience, awareness, memory, and behavior that I would like to highlight potential areas of common ground. Here, I can make a strong argument that anthropology would benefit from recent theory and research in cognitive science and experimental psychology.

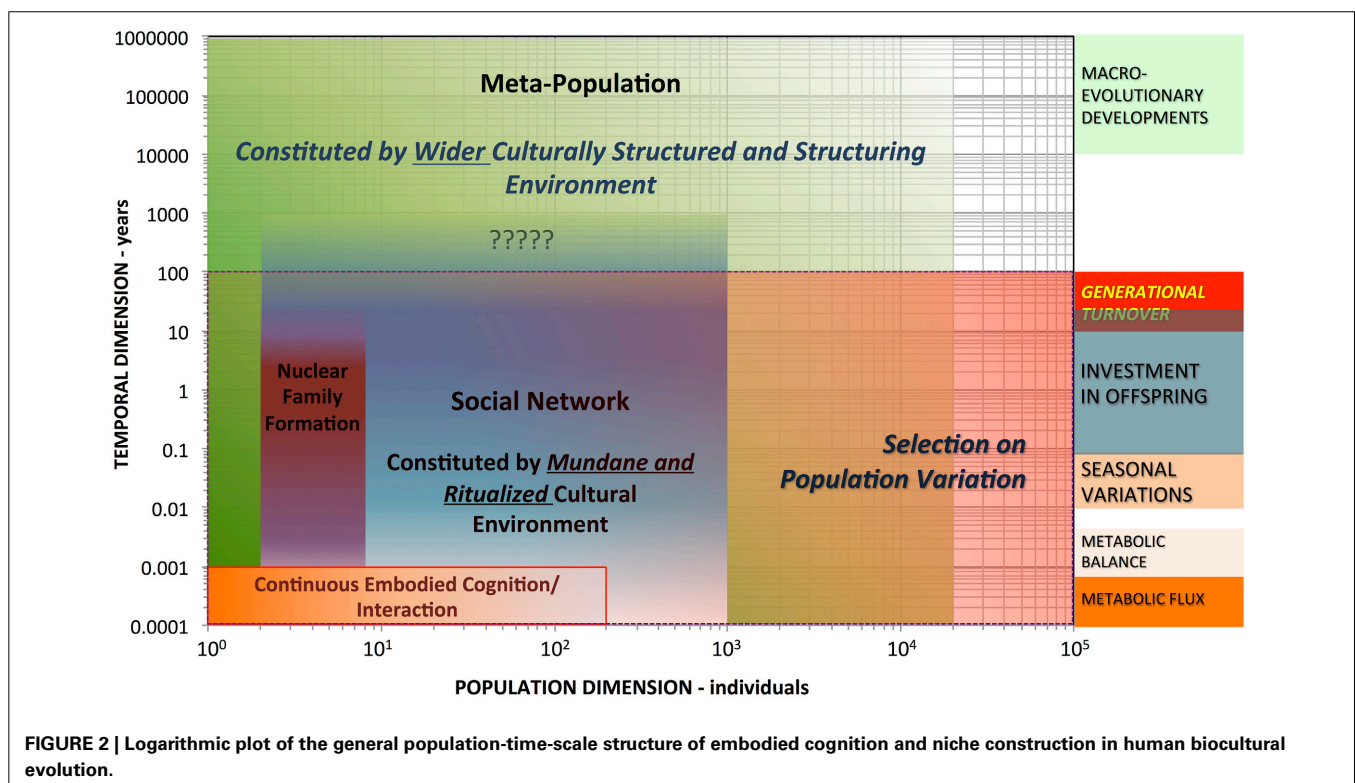
I would further suggest that—for the purpose of interdisciplinary bridge-building (or rebuilding, following mid-late Twentieth Century trends of academic disciplinary proliferation and divergence)—it is worth mapping important anthropological and linguistic terminology and concepts onto relevant psychological and cognitive science ones. For example, in cultural anthropological theory experience and action are widely understood as interrelated, practical, semiotically structured and structuring processes (Sahlins, 2000; Geertz, 2001; Ortner, 2006). There is potentially rich common ground with embodied cognition research, particularly surrounding continuous embodied interaction as a recursive, structured and structuring process that encompasses perception, complex associative learning, episodic memory, simulation and representation.

As with sociological and anthropological theories of practice—where action is always simultaneously symbolically and materially structured and structuring (Bourdieu, 1977; Giddens, 1984; Ortner, 1984)—embodied cognition and continuous environmental interaction is a phenomenon that challenges more than the problematic theoretical distinctions between mind and body. It also blurs distinctions between *representation* and *hierarchically organized, cybernetically regulated action sequences* (Rowlands, 2006; Pfeifer et al., 2007). I would further emphasize that embodied cognition breaks down the apparent boundary between *inter-individual social communication*, on the one hand, and *intra-individual, hierarchically structured feedback through the entire nervous system*, on the other. In fact, it is here that anthropological concepts point toward effectively augmenting recent embodied cognition theories of human language (Clark, 2008; Dove, 2011, 2012). And it is here that anthropology can help cognitive science tackle a key question: if embodied symbol concepts tend to be learned through multimodal sensory-motor experience and remain grounded through evoked sensory-motor simulation, how are abstract concepts successfully constructed and understood (Dove, 2011)?

With a distinctively anthropological point of departure, the ENC hypothesis can further, more comprehensively guide cognitive science research out of the laboratory, into the wide-open, culturally structured and structuring wild of human cognition (cf. Hutchins, 1996). As **Figure 2** schematically illustrates, hominin ENC is proposed to involve non-nested hierarchical system feedbacks from individual, temporally continuous embodied cognition and interaction (lower left) to a metapopulation constituting a cultural environment with elements persisting over

millennial timescales (e.g., language structures, technological traditions, and culturally modified landscapes). Ritualized (and ritualizing) cultural environments pervasively structure embodied experience and action. Examples are widely ethnographically and even archaeologically documented. Consider, for instance, how diverse communities handle death's dual crisis—involving both social loss and the emergence of an abject cadaver (Nilsson Stutz, 2003). The cross-cultural diversity in—and patterns of long-term prehistoric and historic change through—mortuary ritual dramatically highlights how human societies construct and practically manage life passages or crises (Nilsson Stutz, 2003). Everyday ritualized handling of embodied experience can also be powerfully transformative, including social interpretation of individual dream experiences—and representations of those experiences (Kracke, 2007). Such qualitative and cross-cultural comparative data support explaining how embodied perceptual representations and concepts could shape and sustain what Dove (2011) has called “dis-embodied” abstract concepts, which emerge from embodied interaction and experience. Abstract but socially useful concepts like being, death, hierarchy, identity, comparison, duty, purity, pollution, and sacred do tend to get materially grounded in embodied social techniques—through dramatic rituals, detailed myths, and production and interaction with artistic objects (Gell, 1998; Nilsson Stutz, 2003). At the same time, these very practices also produce unimaginably rich material for cognitively constructing associations (cf. Clark, 2008; Heyes, 2010a,b, 2012).

In general, the systemic feedbacks among individual memory construction, associative learning, mental planning and social decision-making, on the one hand, and prevailing cultural environments, on the other, combine to shape and constrain how



individuals cognize abstract concepts over time, dynamically associating them to the concrete (cf. Lévi-Strauss, 1962a). Still, the embodied cultural environment would seem to be overwhelmingly rich in potentially evocative associations. Thus, the embodied cultural environment would seemingly hinder the individual from sustaining her attention on any one given coherent embodied system of associated extrasomatic environmental stimuli, perceptual experiences, bodily sensations, and remembered experiences. With a prohibitively intricate range of possible, imaginable symbolic boundaries or associative relationships, how is it that we can even construct memory, make behavioral decisions, or achieve enduring attitudes or opinions? This consideration suggests framing the key question another way. If apparently abstract concepts are actually emergent features of complex systems of embodied association production (cf. Clark, 2008), how do such concepts remain robustly connected to a stable association structure? In thought and action, we regularly succeed in making such practical but effectively aesthetic choices (Gell, 1998; Agamben, 1999). I emphasize that theories of embodied cognition—considered from an anthropological perspective—suggest the following prediction. The emotional, proprioceptive, and interoceptive experiences that *result from* successfully constructing or reconstructing socially salient associations between present and past, general and specific, other and self, known and unknown themselves constitute an embodied, unconscious heroic narrative representation of *self successfully constructing a coherent, durable aspect of the world*.

In the remainder of this essay I hope to convince the reader that the ENC hypothesis focuses our own scholarly joint attention on embodied narratives—that is, not simply embodied cognitive simulations, but temporally-compressed, emotionally evocative representations of remembered or imagined events and experiences that would occur over longer time periods, often involving acquisition of durable, even timeless dispositions. Moreover, I hope to make the case that embodied narratives are complex because they are simultaneously adaptive phenotypes and part of our dynamically evolving niche. In this context, embodied narratives have been gradually transformed in the hominin (human) lineage, from private iconic constructions to socially shared, recursively elaborated and endlessly mashed up forms.

### EMBODIED NICHE CONSTRUCTION: SEMIOTICALLY STRUCTURED AND STRUCTURING COGNITIVE INTERFACES WITH THE ENVIRONMENT

When Dawkins described the “long reach of the gene” as its extended phenotype, he argued that DNA replicators drive ecological processes at multiple scales, from intra- to intercellular and from somatic to extrasomatic levels (Dawkins, 1982). The extended-phenotype concept was rhetorically compelling. It was also a heuristic corrective within the history of thought in evolutionary biology. Dawkins argued against overemphasizing or reifying the organism and environment—or even the nature of continuous intergenerational evolutionary change as *constantly* gradual. Yet, this replicator-centered perspective is just as theoretically insufficient for explaining biological evolution as are strictly organism, population, or ecosystem-centered views. Recent work on dynamic niche construction processes (Odling-Smee et al.,

2003) and multi-scalar complex ecological processes (Levin, 1992; Gunderson and Holling, 2001; Schneider, 2010) makes clear that biological systems characteristically exhibit resilient structures or equilibrium states at more than one scale, but change at a given scale or structure can have important dynamic feedback effects stretching beyond the local environment. Genetic variation in replicators may drive evolutionary competition and selection, but this is a relatively local ecological complex-systems process. Evolution involves dynamic feedbacks among replicator populations; their non-nested, hierarchically structured extended phenotypes; and their similarly non-nested hierarchically structured niches (Allen and Starr, 1982).

Moreover, there are multiple scales within and around the organism in which phenotypes have a *dual* systemic role. Phenotypes are not just subject to natural selection for fit to the prevailing environment. They are often also the very environments that influence their own fitness (Odling-Smee et al., 2003). In animals ENC occurs because the chordate body is interconnected with the world through complex thresholds, which constitute an integral part of the animal’s niche, even as the body is also adaptation to that niche.

It is especially relevant, then, that embodied cognition research emphasizes the dynamic multiscale, often parallel-channel process of an organism’s interaction with the environment (Clark, 1998, 2008; Damasio, 2000, 2003; Rowlands, 2006; Pfeifer et al., 2007). In fact, embodied cognition and environmental interaction may be understood as thoroughly intertwined, multidimensional and hierarchically structured processes in which:

- the bodily surface and the sensory organs constitute a complex, interactive threshold between the organism and surrounding environment—in which foraging, predation avoidance, resting, courtship and mating, and other social behaviors unfold; and
- the brain continuously interacts with:
  - the rest of the body—via homeostatic systems that regulate circulatory, metabolic/digestive, motor activity, bodily balance, immune/inflammation, endocrine, reproductive system, and alertness states; and
  - the extrasomatic environment—via sensory organs during joint cognition/bodily action processes; and
  - the subsystems within the brain itself—via:
    - connections among modality-specific perception systems;
    - connections between perceptual representation systems and more specific homeostatic systems, giving rise to feelings and awareness [e.g., metabolic and digestive information increasing alertness, generating an embodied feeling of hunger (cf. Damasio, 2003)]; and
    - simulation (Barsalou, 1999, 2009), memory construction, and other higher-order mental processes that use learned, iconic and indexical neural representations of embodied cognitive/interactive states (cf. Deacon, 1998: Chapter 3), which in turn can affect sustained mental or bodily attention on a cognized object.

With the exception of the simulation and memory construction component, this sketch of embodied cognition/interaction may



generally model behavioral processes in vertebrates and most invertebrates. That last component itself—capacity for simulation and learned iconic and indexical embodied representation—does encompass a range of phenotypes found across the vertebrates. In turn, this comparative zoological scope suggests that what Clark (2008:p. 42) has described as “profound” and “promiscuous” embodied engagement with the world is evolutionary quite ancient in the animal kingdom and widespread, especially in the terrestrial biosphere. Finally, as Deacon’s (1998: Chapter 3) symbolic-threshold model suggests, the higher-order cognitive capacity to recursively and indexically link symbols, symbols and iconic representations, and deictic symbol-index concatenations could have gradually evolved in the hominin lineage (see **Figure 1**), with hierarchical complexity increasing over 100’s of thousands of years, so that open-ended systems of symbolic representation incrementally pervaded the environments in which embodied interaction occurred.

It is worth pointing out that, here, Chomsky and colleagues’ recent general emphasis on symbolic recursion in language (Hauser et al., 2002, 2014) may be seen—at least potentially—as theoretically converging toward Deacon’s model, opening the way to a synthetic foundation for investigating the evolution of language as part of a much longer, gradual dynamic, in which embodied cognition and interaction has resulted from a very deep co-evolutionary process. Embodied cognition/environmental interaction has resulted from diversifying niche construction and phenotypic adaptation—via natural selection—across the animal kingdom. I thus suggest that it is particularly important to investigate how embodied cognition/environmental interaction dynamically evolved in the hominins as both niche and phenotype.

### HIGHER-LEVEL COGNITIVE REPRESENTATION IN EMBODIED COGNITIVE CONTEXT

Adaptive interfaces in most vertebrate animal niches constitute—and are constituted by—two non-nested hierarchically structured levels of embodied attention to the immediate surroundings. First, distributed embodied cognitive management of environmental interaction can be monitored and managed through higher-level neural connections in the brain—or even by the relatively distal spinal cord. In either case, environmental interaction effectively proceeds with little centralized higher cognitive interpretation and direction of bodily activity. This allows grasping, gross limb movements, management of torso posture, head movements, chewing, and dynamic gazing/visual scanning to occur efficiently, without constantly raising overall bodily alertness levels—and without overwhelming higher cognitive decision-making systems for determining the present focus of selective attention. Second, higher cognitive processes can take place at the same time as the substantially decentralized orchestration of behavior operates. Higher constructive cognitive processes would be able to manage modal and cross-modal learning and perception of relevant environmental objects, deixis with respect to those objects, memory construction, and—perhaps most importantly—tactical decision-making about changing short-term equilibrium goals.

In evolutionary perspective, this dual-level model of embodied cognition can aid our theorizing about vertebrate animal behavior. For example, it supports our hypothesizing about behaviors of common vertebrate species that thrive in habitats shared with humans. For example, I can reasonably speculate that embodied squirrel (genus *Sciurus*) cognition facilitates sustained manual grasping and rhythmic mastication of a food object, reducing embodied attention on the actual handling and eating behaviors, and thus opening the animal’s sensory focus on potential predators, possible mates, and territorial challenges from other squirrels. In general, we may expect that embodied dimensions of cognition have been important in animal evolution, because they may minimize the opportunity costs of selective attention on one fitness determinant—for example, foraging—at the expense of others—including predation risk, conspecific territorial challenges, or courtship and mating.

### SEMIOTIC SYSTEM CONSTRUCTION IN EMBODIED COGNITIVE CONTEXT

The dual-level structure of embodied cognition helps us to investigate the evolution of semiotic construction capacities—initially not for structuring social signaling, but for structuring individual learning, perception, and reasoning. Here, it comes into very sharp focus that “semiotic as logic” is sufficient for learning and memory construction—that is, the cognitive filtering, storage, and recall of relevant information in the animal’s ecological context. This is the case, even in the absence of adaptations for complex social communication. Evolutionary primitive cognitive construction of embodied icons and indices is arguably widespread in the mammalian and avian classes. Here, I use Peirce’s terminology of signs, following Deacon (1998: Chapter 3), emphasizing that perceived or otherwise cognitively evoked *icons*—as arbitrarily simplified representations of a more complex object—may be dynamically constructed over recurrent embodied interactions with the extrasomatic environment. Moreover, building on the semiotic framework presented in Gell (1998), Peirce’s definition of *index* implies that iconic representations also have indexical properties, intrinsically evoking embodied, remembered objects, including other iconic representations. Peirce (2012: Kindle location, 2018) stated, “An Index is a sign which refers to the Object that it denotes by virtue of being really affected by that Object.” The embodied—albeit higher, integrative—cognitive process of icon construction must involve indexical links to perceived objects, which are also deictically, indexically linked to bodily affective states. Thus, icons and their indexical connections form a learned, structured cognitive filter. This functions to convert highly complicated, often noisy extrasomatic and somatic stimuli into consistent, comprehensible body-environment interaction channels. The animal, then, learns to engage in bouts of behavioral activity involving habitual, substantially decentralized embodied cognition and interaction with the environment. At the same time, higher cognitive functions tune attention toward the recurrent abstraction of visual, auditory, or fine, focused haptic perceptions and memories. The resulting pre-linguistic semiotic constructions would thus be mental, private *icons* of more detailed perceptions or simulations. These could include both static imagery and narrative

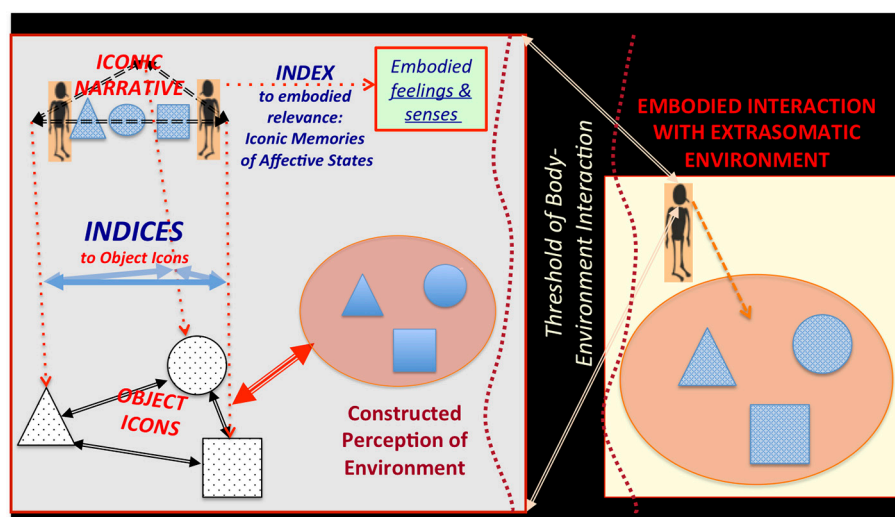
memory. Moreover, these icons would have a secondary semiotic function, *indexically* evoking proprioceptive, interoceptive, and tactile sensations and emotional states, metonymically tied to an iconic representation through prior experience and memory construction. Most basically, I suggest that the evolution of higher cognitive recursive functions that can abstract rich embodied experience and memory into iconic representations likely built on the embodied indexical connection between the following—also dual-level—system of cognition, with the second level further involving embodied indexical links among three particular aspects of cognition:

- visually, auditorily, or haptically mediated construction of extrasomatic objects as alienated from the body
- somatosensory states that affectively represent the condition of part or all of the body, including:
  - interoceptive states
  - proprioceptive states
  - alertness

Extending arguments presented in Rowlands (2006), I suggest that—within the context of dynamic experience that structures and is structured by icon-index semiotic systems—we can consider somatic states that are indexically evoked by visually, auditorily, or haptically shaped extrasomatic icons as *embodied icons*. Semiotic, structural linguistic, and anthropological theories of representation have long emphasized that signs are defined not only by the conventionally or logically defined relationship between signifier and signified. They are also defined by their formal relationships to other signs (Jakobson and Halle, 1956; Hockett, 1961; Lévi-Strauss, 1962a,b; Leach, 1976; Saussure, 2011;

Peirce, 2012). Following Deacon (1998: Chapter 3), I suggest that the cognitive ability to learn and construct indexical relationships among iconic representations of extrasomatic objects and embodied icons evolved gradually, supporting the recursive construction of concise iconic narrative representations (**Figure 3**). This would have set the stage for later hominin social manipulation of indexical connections to extrasomatic icons, through gesture and gaze-following (Tomasello, 2008).

In general, the semiotic structuration of embodied cognition is important when partially decentralized cognition and action affords the animal's short-term homeostatic behavior pattern in the extrasomatic environment (Gibson, 1979), freeing up attention toward receiving and decoding information that might imply the *relevance* (cf. Sperber and Wilson, 1995) of altering behavioral and affective homeostatic targets. I speculate that what may have become relatively evolutionarily derived in humans—already early in the divergence of the hominin lineage from that of the panins (see **Figure 1**)—was the cognitive capacity to construct iconic narratives, in which dramatically changing affective states are temporally contextualized in a representation of a problem (e.g., hunger during the search for food) and its resolution (satiation during feeding). Such iconic narratives should be considered as an emergent part of the embodied interface with the environment, helping the animal to sustain attention on a difficult-to-obtain goal (dragging a stone anvil to the base of a nut tree) or a social dilemma (accepting or rejecting a solicitation to engage in a social coalition). Thus, the recursive nature of constructing icon-index complexes is hypothesized to be an important evolutionary inheritance in hominins, subsequently modified by natural selection to support human symbolic thought and—eventually—communication.



**FIGURE 3 | Hierarchically constructed object icons and iconic narratives—indexically linked to embodied affective states and perceived objects in the extrasomatic environment—via body-environment thresholds.** Object icons are sufficiently formed through dynamic, recursive learning. Construction of indexical relationships among icons emerges through embodied interaction with the extrasomatic

environment. *Iconic narratives* are constituted by the indexical relationships among object icons and changes in embodied affective, proprioceptive, and interoceptive states. It is hypothesized that one of the most evolutionarily primitive iconic narrative genres—likely evolved in the hominin-great ape common ancestor—is that of heroically succeeding or tragically failing to construct an enduring aspect of the world.

The ENC hypothesis entails that evolution modified the ape “dual-level system” of embodied cognition (see above) in hominin prehistory. Partially decentralized management of locomotion, repetitive tool-making and tool-use gestures, grasping and carrying, and feeding—involving rhythmic or recurrent actions or sustained isometric postures—could be maintained over minutes or hours, while higher cognitive learning, construction, and perception could simultaneously support:

- introspective attention on semiotic signs, their objects, and evoked relationships to other signs
- monitoring of others
- communication, involving solicitation and engagement in bouts of joint attention

I speculate, then, that in early hominins, semiotic representation likely co-evolved with social monitoring and solicitation and sustained engagement in joint attention and interaction, prior to the evolution of spoken language.

### OVERVIEW OF EMBODIED NARRATIVES IN HOMININ NICHE CONSTRUCTION

It is an embodied cognition perspective that makes this possibility apparent. Self's frequent attention on participation in socially intense networks may be punctuated by highly focused, goal-oriented social interactions, technological engagement with the material environment, and ritualized motion sequences. All of these involve reduced alertness, temporarily shutting off embodied interfaces with the wider environment. Thus, managing a marriage alliance, butchering an animal, shaping a wooden digging stick, making a flint hand-axe, or building a hut may involve intense, narrowly selective embodied attention on a multi-step technological process or social interaction, resulting in a material product or negotiated relationship-state. The higher cognitive processes involved in such activities are particularly important for rapid niche construction. The ability to construct abstract, iconic representations from concrete visual, auditory, and haptic perception is an embodied behavioral adaptation. However, the representations themselves—along with their semiotic, indexically or metaphorically evoked connections to other learned signs, perceived objects and events, memories, and embodied mental simulations—become part of the niche, constituting a dynamic part of the interface between the body and the extrasomatic environment (cf. Clark, 2008; Dove, 2011).

The hominin embodied niche is hypothesized to have evolved to encompass a set of dual-level cognitive interfaces with the extrasomatic environment, where rhythmic or sustained static behavior patterns unfold in parallel with complex systems of indexically linked icons, facilitating social observation, action, judgment, and self-awareness. The evolution of bipedal locomotion in the hominin lineage illustrates the complex process of “semiotically constituted and constituting ENC,” examined in the following section.

### A CASE STUDY: BIPEDAL LOCOMOTION AS EMBODIED PHENOTYPE AND NICHE COMPONENT

Vertebrate locomotion involves a joint cognitive-behavioral system facilitating the animal's movement through its physical

habitat. In general, locomotion itself may be seen as an embodied cognition system, in which control of locomotion is partially—but significantly—decentralized across the central nervous system, peripheral sensory-motor subsystems, and musculo-skeletal subsystems. Central cognitive processing of sensory and other inputs from bodily homeostatic systems is usually minimized, first, through local oscillatory feedback in the limbs, and then through central nervous system management of small homeostatic neuromotor adjustments (Van de Crommert et al., 1998; Dietz, 2003, 2010; Ijspeert, 2008). Finally, the central nervous system supports monitoring a simple series of embodied indices of homeostatic exertion and equilibrium levels of bodily momentum in the immediate extrasomatic environment. These indicators mainly involve the sense of bodily balance, and departures from homeostatic ranges can trigger a cascading increase in local sensory-motor and overall central nervous system alertness, in order to respond to a sudden change in the body's interaction trajectory with the surrounding milieu. We can usually walk—or birds fly—without higher-order cognitive information-processing and decision-making about every heel strike or big-toe push-off—or wing flap. Thus, we can walk, chew gum, play air drums to an imagined tune—and seagulls can scan visually for other members of their flock, prey, and predators—while embodied, distributed cognition takes care of locomotion.

The embodied cognitive niche dimensions of bipedal locomotion are strongly shaped by the fact that, in terms of energy expenditure by the supporting musculo-skeletal and thermoregulatory systems, hominin two-legged walking or jogging uses caloric resources at a substantially lower rate than does great ape quadrupedal walking or running at the same pace (Leonard and Robertson, 1997; Sockol et al., 2007). Efficient, largely decentralized cognitive management of bipedal locomotion synergistically reinforces the biomechanical, energy-saving advantage of bipedal stride or jogging gait in a terrestrial open habitat. While undertaking long—and long-distance—bouts of bipedal locomotion, distributed cognitive management frees up other embodied cognitive systems for visual, auditory, and olfactory perception, semiotic construction, planning, sustained goal-oriented selective attention, and active communication (cf. Langdon, 2005).

### THE EMBODIED BIPEDAL TERRESTRIAL NICHE: REDUCED SOCIAL ALERTNESS DURING FORAGING

Hominin locomotion adaptations emerged from ca. 7–5 million years ago (mya) onward. Their evolution initially co-occurred with the phylogenetic divergence of our lineage from that of chimpanzees and bonobos (Won, 2004; Lovejoy, 2009; Lovejoy et al., 2009; Webster, 2009; Yamamichi et al., 2011). In early hominin populations directional selection modified the ancestral ape pattern of quadrupedal walking and vertical climbing (Sockol et al., 2007). Here, the evolutionary process would have favored locomotor phenotypes that were not only generally fit to the mosaic forest-grassland habitat features of East and northern Central Africa, but also minimized energy expenditure in that habitat (Wheeler, 1991a,b; Leonard and Robertson, 1997). The emergence of the genus *Australopithecus*, ca. 4 mya in East Africa, appears to have coincided with the evolution of “obligate bipedalism,” in which the anatomy supporting efficient stride is so specialized that it substantially limits

habitual arboreal climbing (Jungers, 1982; Latimer et al., 1987; Latimer and Lovejoy, 1990; Ohman et al., 1997; Haile-Selassie et al., 2010). Obligate bipedalism—involving an arched foot, non-opposable big toe, and a strongly disto-medially angled femur—is first documented among fossil traces of *Au. anamensis* (ca. 4.2–3.9 mya) and *Au. afarensis* (ca. 3.7–3.0 mya) in East Africa. In modern humans obligate bipedalism adapts our bodies highly efficiently to moving around, slowly but surely, in a terrestrial diurnal habitat, during continuous trips that may last hours and cover as much distance as the diameter of some wild chimpanzees' and gorillas' lifetime territories (ca. 30–50 km). Although australopithecines exhibited a range of vertical climbing and pedal locomotor grasping anatomy (DeSilva et al., 2013), skeletal support for bipedal locomotion in *Au. afarensis* (a fossil species most famously represented by the partial skeleton “Lucy,” specimen AL-288-1) had already evolved as an integrated adaptive system between ca. 4–3 mya, well fit to terrestrial activity (Haile-Selassie et al., 2010).

As such, early obligate bipedalism was a *phenotype* shaped by natural selection in a long-term process of niche-population co-evolution. Yet, as a *niche component*—that is, as an embodied interface with a terrestrial, relatively open habitat that housed a very diverse, heterotrophic food resource spectrum—obligate bipedal locomotion entailed new selective pressures. First, obligate bipedalism exposed early hominins to a range of large felid predators (Hart and Sussman, 2011). As a phenotype in a foodweb heavy with large carnivores, bipedal sprinting is no match for well-adapted quadrupedal running over short distances (Leonard and Robertson, 1997). It may be inferred—based on theoretical models and comparative data (Sussman et al., 2005; Hart and Sussman, 2011)—that during this critical period of niche-population co-evolution, natural selection would have favored pro-social behaviors for aggregation. Being part of a group not only enriches the individual's perceptual information about predatory threats, via indexical predator alert calls (Seyfarth et al., 1980; Zuberbühler, 2001; Stephan and Zuberbühler, 2008). It also simply reduces the likelihood that a given individual will be the one ambushed and captured by a lion or sabretooth cat hiding in the tall grass (cf. Hamilton, 1971). The bipedal embodied niche initially involved, at the very least, frequent bouts of reduced social attentiveness, coupled with increased alertness for indications of diverse predators—and of prey. Yet, when terrestrial foraging was successful, the immediate extrasomatic environment would have changed dramatically. If the discovered food resource was rich enough—whether it consisted of larger game or of carbohydrate-dense or fatty plant tissues—cooperatively transporting and defending that food resource would have tended to increase the group-members' average inclusive fitness. Here, distributed cognitive management of locomotion would have been especially important. Individuals would have had to carry bulky or heavy food packages, while suppressing attention to strong emotional-desire responses to hunger sensations, with heightened, rapidly shifting attention on immediate group members, predator risks, and near-term future possibility for satiating hunger in a safer aggregation locality.

## THE EMBODIED BIPEDAL TERRESTRIAL NICHE: COOPERATIVE OFFSPRING CARE AND INTERTWINED NARRATIVIZED SOCIAL IDENTITIES

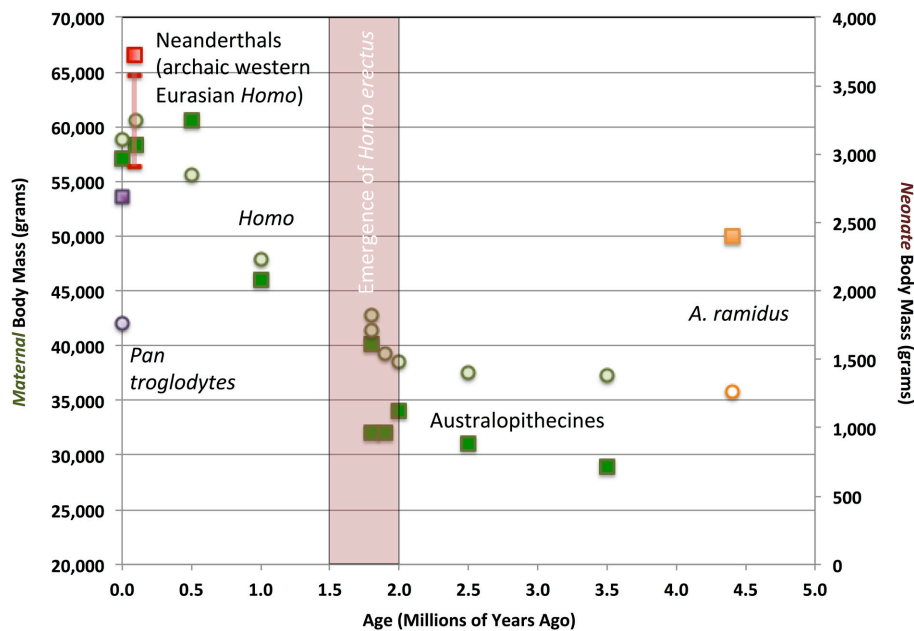
Obligate bipedalism—as an embodied niche interface system—thus conspicuously supported pro-social adaptations that went beyond simple gregariousness. In fact, there were profound social implications for hominin bipedal locomotion in a terrestrial, heterotrophic niche. Other things being equal, bipedal anatomy is not only relatively maladaptive for escaping large quadrupedal carnivores. For the adult female, it may be considered an evolutionary compromise—in comparative primate evolutionary ecological perspective—for giving birth (Tague and Lovejoy, 1986; DeSilva, 2011; Kurki, 2011, 2013; Wells et al., 2012) and carrying relatively helpless infants (Wall-Scheffler, 2012), while also supporting long bouts of bipedal walking. Pelvic morphology well suited for lowering the body's center of gravity and reducing mechanical effort (and short-term metabolic balance and long-term stress on muscles, bones, and associated connective tissues) during upright walking exhibits a limited pelvic aperture (Lovejoy, 1988; Rosenberg and Trevathan, 1995). Natural selection has shaped this functional anatomical compromise in the adult female hominin body, achieving efficient bipedal stride at the expense of more frequent, riskier obstetric complications giving birth to large-brained neonates. Given the evolutionary success of this compromise—measured in terms of extant human geographic range and biomass—it is clear that the typical population-level survival and reproductive success benefits of bipedal locomotion have more than made up for any morbidity and mortality risks associated with parturition through a bipedal pelvis.

The realized reproductive success associated in integral part with bipedalism is all the more remarkable, because the posture requires adults to use their arms to carry infants and young juveniles, raising the adult's center of gravity during locomotion. This imposes a relatively greater metabolic and stress cost in caring for very young offspring who are still developing sufficient fine and gross motor strength and coordination. Without a compensating phenotype, the “carrying cost” would increase risk of predation for mother and infant, alike. It would also reduce her foraging efficiency, slowing down food search rates, while burning more calories to search for food. The resulting metabolic deficit would also increase mortality risks for a lactating mother and her offspring. Virtually the only theoretically plausible behavioral phenotypic compensation that would have co-evolved with the bipedal embodied niche is alloparenting: cooperative offspring care.

Indirect, circumstantial evidence for alloparenting primarily consists of data on the metabolic costs and ecological risks otherwise imposed by carrying infants and young juveniles during bipedal locomotion. Further evidence comes from estimates of adult maternal and neonate body mass<sup>1</sup>. **Figure 4** depicts

<sup>1</sup>Maternal body mass may be estimated, even from fragmentary weight-bearing bones (especially the pelvis, femur, and tibia), based on extant human and ape population correlations between skeletal element morphology/allometric scaling and adult body mass (McHenry, 1992; Ruff, 2010; DeSilva, 2011). Neonate body mass estimates may be independently derived





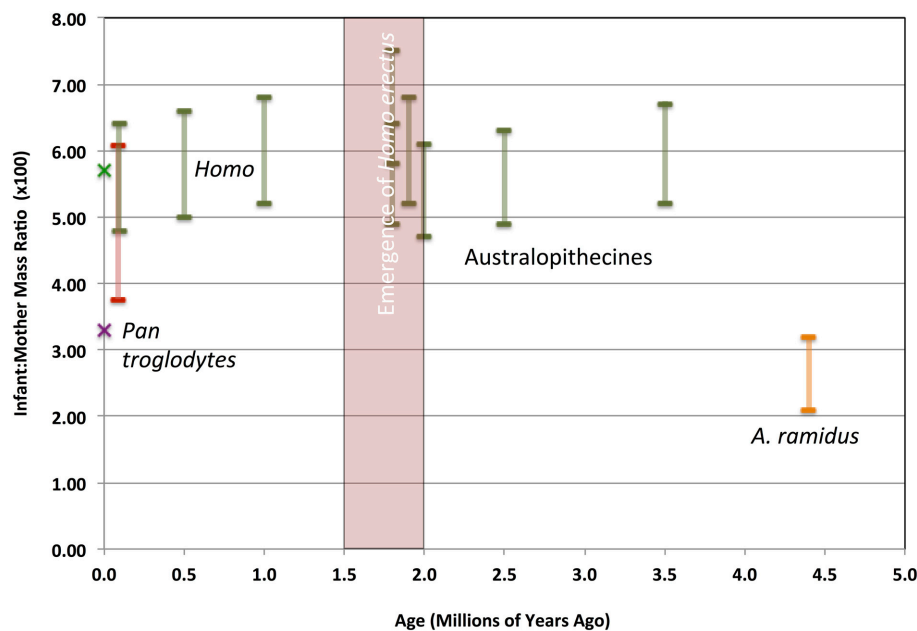
**FIGURE 4 | Evolutionary change in maternal and neonate body mass in the hominin lineage, compared with extant chimpanzees.** Maternal body mass averages measured from fossil and modern samples shown as squares. Neonate body mass averages measured from fossil and modern samples shown as circles. *Ardipithecus ramidus* shown in orange. *Australopithecus* and *Homo* samples shown in green, except for Neanderthals, shown in red. The Neanderthal female and neonate fossil samples raise the possibility that this late Pleistocene (ca. 200-40

thousand year old) western Eurasian population may have evolved a higher body-mass:brain ratio, compared to contemporaneous regional human populations (VanSickle, 2009; DeSilva, 2011). The Neanderthal neonate body mass estimate is thus shown as a conservatively estimated range. Modern chimpanzees mothers and neonates are shown in purple. Modern and fossil maternal body mass measurements and estimates are from VanSickle (2009) and DeSilva (2011), and neonate body mass estimates are calculated after the methods in DeSilva and Lesnik (2008).

variation in maternal and neonate biomass in hominin samples distributed across the last 4.4 million years of our lineage's evolution, mainly based on data from DeSilva (2011). The samples begin with *Ardipithecus ramidus*, fossil specimens of which document a hominin population lineage that had retained the ancestral anthropoid opposable big toe, along with other anatomical indicators of “non-obligate” bipedal locomotion and vertical climbing adaptations in a forested East African habitat (Lovejoy, 2009). In comparison, the later hominin samples from 4.0 to 2.0 mya include extinct or ancestral australopithecine species, whose fossil remains document early obligate bipedal adaptations. As DeSilva (2011) has emphasized, a substantial evolutionary reduction in adult female body mass occurred from *Ardipithecus* to *Australopithecus*. That such a shift occurred early in hominin evolution was not apparent prior to publication of the description of *A. ramidus* post-cranial anatomy (Lovejoy, 2009; Lovejoy et al., 2009). *A. ramidus* females (represented by the remarkably intact adult female individual “Ardi” whose estimated body mass is represented in Figure 4 by the orange square) had an adult size comparable to the typical level in living chimpanzees (ca. 50 kg, also shown in Figure 4, represented

by a purple square). Later, australopithecines and early members of the genus *Homo* inherited modified, reduced female body mass, with adults weighing only 30–40 kg. Yet, small adult female body size was linked to a substantial jump in encephalization, because head-size was as large or larger than that seen in apes and ardipithecines. Australopithecine and earliest *Homo* adult endocranial volumes (both sexes included) spanned ca. 400–800 cc. The lower end of this range overlaps with the upper end of the chimpanzee distribution. Thus, the slight australopithecine maternal body mass values would have been associated with relatively large—and large-brained—neonates. Exhibiting obligate bipedal anatomy, the small australopithecine/early *Homo* maternal body would have contributed to high neonate:maternal body mass ratios (Figure 5). This constitutes additional, albeit indirect, evidence that alloparenting co-evolved with bipedal locomotion. The actual maternal costs of carrying an infant while walking upright were greater than paleoanthropologists have previously thought. Evolutionary theory predicts that, other things being equal, early hominin mothers with either more evolutionarily ancestral (i.e., ape-like) or more derived (i.e., like large-bodied *Homo*) body mass and locomotor anatomical traits would have had a reproductive-success advantage in mosaic East African forest/grassland habitats. Yet, the smaller, more vulnerable, nutritionally precarious australopithecine maternal bodies should have faced stiff evolutionary competition. The australopithecine/early *Homo* pattern—small maternal bodies and high

from endocranial volume measurements, elegantly based on constant anthropoid proportions between maternal and neonate brain mass, coupled with great ape and human constant proportions between brain and body mass (DeSilva and Lesnik, 2008; DeSilva, 2011; Tüma and Brůžek, 2013).



**FIGURE 5 | Evolutionary change in neonate:maternal body mass ratios in the hominin lineage, compared with extant chimpanzees.** Samples are color-coded as in Figure 4. Fossil hominin samples are shown as estimate ranges. Extant human and chimpanzee sample means shown as Xs.

neonate:maternal body mass ratios—contradicts such standard evolutionary expectations.

This development may be explained by australopithecine co-evolution with alloparenting behaviors, associated with a female life-history strategy that gave up adult somatic mass, in order to maintain caloric and nutrient resource transfers to offspring during gestation and lactation. Small body size would have constituted a costly, honest signal that a mother actually needed assistance in carrying or provisioning offspring in order to maintain energy balance for herself and for her infant.

Arguably, then, the question is not whether alloparenting co-evolved with obligate bipedal locomotion and terrestrial foraging/costly infant-carrying in australopithecines. It is rather: What was the social structure of alloparenting, particularly in relation to courtship, mating, and possible cooperative foraging or food-sharing behaviors (Hrady, 2009)?

Here, I present a plausible theoretical claim for the following structure in early australopithecine social systems. Alloparenting would have fundamentally involved unrelated females—having transferred at maturity from their natal groups, as observed in living gorillas, chimpanzees, and bonobos—engaging in reciprocally altruistic assistance during parturition and early childrearing. Continuous honest signaling of maternal and infant need—in the embodied form of mother's small size—would have also favored reduced male aggression and male provisioning of offspring. This is because adult males would likely have faced a time and energy trade-off between provisioning mates and maintaining and defending a female harem. From this perspective alone, it may be argued that bipedal locomotion in a terrestrial, heterotrophic niche would have favored reduced

male aggression, evidence for which may be seen in the substantial evolutionary canine-size reduction seen in males and females alike, from ca. 6–4 mya (*Ardipithecus*) to ca. 4–1 mya (*Australopithecus* and early *Homo*) (Figure 6) (Haile-Selassie and WoldeGabriel, 2009). In fact, the most effective strategy for repeatedly obtaining additional calories, in the form of divisible “public-goods-like” food packages (cf. Hawkes et al., 1998; O'Connell et al., 1999, 2002; Hawkes, 2003) that could be shared with mates and (statistically likely) offspring—or at least brothers' or half-brothers' offspring—would be for males to cooperate, at least occasionally, in foraging, food transport, and food sharing.

Thus, three interrelated social behavioral patterns would have defined the embodied aspects of the obligate bipedal niche: adult female reciprocal cooperative alloparenting (and—possibly—midwifery); male cooperative foraging and food transport; and male provisioning of mates (Table 1). This behavioral nexus would have constituted a dynamic interface between the body and a socially intense, yet strongly ecologically structured extrasomatic environment.

The ENC hypothesis supports a key prediction: that long-term monogamous pair bonding would have also co-evolved with the bipedal niche, albeit hardly strictly driven by active male foraging, resource transport, and provisioning of females (cf. Lovejoy, 1981, 2009; see Table 1). To be sure, from the adult female's perspective, the socially intense bipedal interface would have partially structured—and been structured by—females' continuous honest signaling to males of precarious energy balance and risk of predation. Critically, this would have occurred across three key foci of adult female social attention:

- courtship
- soliciting food transfers or physical protection during gestation and lactation
- continued food transfers, protection, and learning support for weaned subadult offspring

Yet, the ENC hypothesis predicts that adult males and females would likely have developed complex pair-bonding relationships



**FIGURE 6 | Forensic reconstruction of a male *Australopithecus afarensis* adult.** The male and female *Au. afarensis* permanent dentition exhibits smaller canines than seen in *Ardipithecus*, suggesting evolution of honest signaling of reduced fighting ability. Photograph accessed from [http://upload.wikimedia.org/wikipedia/commons/2/22/Australopithecus\\_afarensis.png](http://upload.wikimedia.org/wikipedia/commons/2/22/Australopithecus_afarensis.png).

through recurrent bouts of food transfer solicitation, food transfers, courtship and sex, separating and reuniting, and mutual monitoring during group travel or nighttime aggregation. This social interaction and monitoring would have supported higher cognitive construction of mutually indexical, highly emotional iconic narratives (see **Figure 3**) of solicitation and resolution, jealousy and relief, and (often substantially prolonged) anticipation and satiation. Semiotic construction of simple mental iconic narratives would have been recurrently evoked through perceived embodied experiences. These iconic representations would have been constructed in part through indexically evoked similarities and contrasts with two other “australopithecine genres” of iconic narratives: those dealing with reciprocal social assistance—including midwifery—between adult females and those involving cooperative foraging and food transport among adult males. In the context of female-male pair-bonds, these evoked narratives would have indexically focused the couple’s joint attention on circumstances at hand that would afford actions—ranging from traveling together, to mutually caring for offspring, to grooming, and to sexual intimacy—that would reinforce the pair bond.

#### THE EMBODIED BIPEDAL TERRESTRIAL NICHE: INFANT HELPLESSNESS AND DIALECTICAL COGNITIVE CONSTRUCTION OF SELF AND OTHER

Obligate bipedal anatomy seems especially maladaptive for the australopithecine infant (other things being equal, of course). To be sure, cooperative social networks would have supported protecting, transporting, and provisioning infants. And in fact, this would have more than compensated for the fact that the especially vulnerable human infant is born with undeveloped, only potentially supportive anatomy for bipedal locomotion. The evolutionary loss of grasping toes in *Australopithecus afarensis* (ca. 4–3 mya) and later hominins (ca. 3 mya and onward), coupled with the mother’s habitual upright stance, would have rendered the infant especially dependent on the parent. The infant could do little—in terms of motor behaviors—to minimize risk of separation. Thus, accidental separation could result in accidental abandonment.

**Table 1 | Matrix of predicted social interactions among adult males, adult females, and juveniles in *Australopithecus afarensis*, based on the embodied niche construction hypothesis.**

	Column action toward row			Derived australopithecine/ <i>Homo</i> trait
	Juvenile	Adult female	Adult male	
Juvenile	Play/Socialization	Gestation Breast feeding Foraging transfers Protection	Foraging transfers Protection	Distributed and increased intergenerational transfers
Adult female	Affect hunger displays Honest signals of need	Reciprocal altruistic alloparenting Midwifery	Foraging transfers Protection Monogamous pair-bonding	Socially extensive cooperative breeding
Adult Male	Affect hunger displays Honest signals of need	Pair-bonding Honest signaling of need implying paternity certainty	Reduced mutual aggression Cooperative foraging, food transport and sharing	Cooperation instead of competition over: ➤ Food resources ➤ Mating opportunities Fatherhood as emergent trait

Moreover, the infant's embodied sensory-motor and visual interface with the upright bipedal mother entailed new challenges for learning in early childhood. Bipedal balance and locomotion likely imposed a steeper embodied cognitive learning curve than would the ancestral quadrupedal walking/grasping/climbing locomotor pattern. It is simply harder to learn to balance on two legs while standing upright than it is to balance on four limbs.

The mother's bipedal posture and gait also combined with her infant's very limited ability to grasp. This would have prevented the infant from habitually orienting its body in parallel to the mother, while lying on top of the mother's back. In apes and monkeys, the highly effective and sensorily rich activity of grasping onto mother's back synergistically allows the infant to feel the oscillatory rhythm of walking on all fours, while also resembling quadrupedal grasping during arboreal climbing. In australopithecines and early *Homo*, in contrast, the earliest experience of being carried during travel and foraging would have been deictically lateralized. Consequently, the obligate bipedal infant would have had a relatively distorted embodied experience of the rhythms of habitual locomotion.

### ***The embodied narrative of exerting agency***

Whereas riding on mother's back incrementally prepares the non-human primate infant to separate from the mother and actively explore, the human infant has a very different embodied learning experience. During the first months of life, the human infant registers and learns about the world through visual, auditory, gustatory, and olfactory inputs. Yet, gross motor balance and fine haptic inputs—which provide short-term, continuous feedback during active embodied learning in non-human primate infants—would have been severely limited for australopithecine and early *Homo* neonates. Thus, the initial period of hominin learning is frequently shaped by a passive embodied interface with the extrasomatic environment. Gross motor strength and balance gradually develop—through learning to lift up the head, roll, sit, crawl, stand, and walk. For the human infant, this embodied learning occurs in extrasomatic environments defined by a protective, usually taken-for-granted adult social network. Still, as I have underscored above, the infant cannot efficiently use embodied memory of earlier experience of being carried by an adult, as she begins independent gross motor learning. This early learning experience, then, can canalize construction of a heroic iconic narrative of exerting agency and achieving greater control over the environment.

### ***The embodied narrative of forming judgments***

Moreover, this iconic narrative would be indexically linked to another, contrasting early-life iconic narrative, in which the infant observes the extrasomatic environment over extended bouts, perhaps as long as an hour, at a distance. The infant is alert but passively secure, able to focus on objects and events around her. This iconic narrative would be one in which passive monitoring results in changing affect. The infant can look, hear, smell, taste, and judge.

### ***Overview of iconic narratives in the obligate bipedal niche: ideology and praxis***

The two main learned iconic narratives of the first months of life are then speculated to be those of exerting agency, on the one hand, and of judging the situation, on the other. I argue that the obligate bipedal niche itself would have encouraged the embodied learning of reflexive alertness in the first months of life, even prior to natural selection for derived higher cognitive phenotypic functions.

When the hominin juvenile has attained sufficient motor strength and balance to begin interacting with a wider environment over longer intervals, she experiences a major contrast between the more passive infant interface and the new, highly exciting but difficult juvenile one. The young juvenile's embodied capital now affords it the agency to seek out and integrate haptic and motor information with other sensory inputs. Here, new iconic narratives become indexically linked to the early-life narrative of agency, in turn, contrasting with the early-life narrative of affective judgment. This may be seen as the ontogeny of an individual dialectic between ideology—the semiotically constituted, narrativized representations of how one hopes or would like the world to be—and praxis—that is, what one does, or does not do, to actualize that ideology.

### **OVERVIEW OF EMBODIED NICHE CONSTRUCTION THROUGH THE BIPEDAL INTERFACE**

The bipedal niche would have minimally supported embodied pro-social bonding between juveniles and their maternal relatives, between adult females, and between adult males (see **Table 1**). The ENC hypothesis deepens existing ecological explanations of how obligate bipedalism evolved. This phenotypic system is traditionally seen strictly as adaptive to a primarily terrestrial heterotrophic niche in East and north Central African, initially among australopithecine populations, ca. 4–3 mya. The dynamic co-evolutionary model of reciprocal causation in niche and population change—based on niche construction theory—illuminates the possibility that obligate bipedalism as an *embodied niche* ontogenetically structured a semiotic interface with the extrasomatic environment. This interface mediated sustained embodied attention on prevailing social and material situations, through constructed, indexically interrelated iconic narratives.

The ENC hypothesis predicts that sustained introspective attention and more hierarchically complex semiotic systems co-evolved with bipedalism. Both semiotically structured cognition and bipedal locomotion were phenotypic adaptations and embodied niche interfaces. The complex niche-population co-evolutionary dynamic resulted in overall population fit to a much more open, yet dangerous terrain, but supported by a rich and diverse range of heterotrophic food resources. Thus, the ENC hypothesis further predicts selection for aggregating in large groups to avoid predation—especially from evening until morning. It also predicts alloparenting by older juvenile and adult female maternal allies, adult female mutual assistance during childbirth, and group fissioning during diurnal, omnivorous, and nonetheless gregarious—if not outright altruistic or synchronized cooperative—foraging, followed by occasional food transport and group fusion. The semiotic component of the embodied bipedal



niche is also predicted to have favored the evolution of more complex symbolic thought, which would have sustained the individual's attention on multi-step longer-term (at least hour-scale) planning, delaying gratification, and making social judgments in guiding solicitation of help, offers of assistance, and mobilization of cooperative activities.

## DISCUSSION

The hominin (human) and panin (chimpanzee and bonobo) lineages mutually evolved a reproductive barrier during the late Miocene period, ca. 7–5 mya, in Sub-Saharan Africa. The foodweb and habitat features that distinguished hominin niche construction from that of the panins included a spatially more extensive terrestrial setting and a broader heterotrophic prey spectrum. The case study on early australopithecine obligate bipedalism, ca. 4–3 mya, supports the argument that hominin niche construction further differed from that of the panins, increasingly involving semiotically structured cognition in the embodied bipedal interface. This interface facilitated adaptive behavior in an intensely social and complex material habitat.

According to the most recent paleoanthropological research, the final Pliocene and early Pleistocene periods—roughly 2.5–0.8 mya—encompassed even more dynamic co-evolutionary change in the embodied niche. The oldest known stone tools were used early in this time frame, with the emergence of the genus *Homo* following soon afterward (Kimbel et al., 1997; Semaw et al., 1997; Semaw, 2000; Domínguez-Rodrigo et al., 2005; Stout et al., 2005; Goldman-Neuman and Hovers, 2009; Kimbel, 2009). We can infer further change in the bipedal interface, which exhibits interrelated modifications in balance, static upper-limb loading, visual, and auditory components. Changes in the anatomy of the hand and thumb supported precision grip in early *Homo*. This adaptive modification would have been integral to making and using small stone tools with the thumb and other manual digits (Marzke, 1997; Marzke and Marzke, 2000). The resulting joint visual and fine manual interface would have structured bouts of goal-oriented, highly visually focused interaction with the material environment (Stout et al., 2008; Stout, 2011), suppressing embodied attention toward social or wider environmental information. Such sustained goal-oriented embodied attention would have contrasted with embodied experiences during which partially decentralized cognition supported locomotion or rhythmic, repetitive tool-making or use, while also maintaining attention on surrounding social events, facilitating social judgment. ENC thus favored temporally alternating interfaces involving social alertness suppression and very context-sensitive social hyper-alertness.

Hominin ENC was especially shaped by those behavioral traits—including bipedal locomotion, tool-making, resource transport, food processing and social food consumption (involving non-agonistic interactions during feeding, as well as active sharing)—whose long-term evolution already had come to define a distinctive adaptation-niche coevolutionary trajectory. These key behavioral phenotypes, whose evolution was well underway by the Middle Pleistocene period (ca. 780–130 thousand years ago [kya]), were also components of the hominin niche. The Middle Pleistocene timeframe is important as we begin to consider how

the ENC hypothesis might inform speculation, hypothesis formation, and research methods concerning the evolution of human language. Toward the end of the Middle Pleistocene, extant absolute brain volumes emerged in Sub-Saharan African anatomically modern humans (AMH) and western Eurasian Neandertals (ca. 200–130 mya) (Lee and Wolpoff, 2003). In addition, extant body proportions—including encephalization quotients—appeared in early AMH populations, after ca. 200 kya (Ruff et al., 1997). The embodied niche complex that co-evolved with the australopithecines and early genus *Homo*, then, was not dependent on a modern-sized brain or brain-body proportions. Evidence on skeletal growth rates and aging in fossil remains from the entirety of the Pleistocene era (ca. 1.80–0.01 mya) further suggests that initial adaptation to the early *Homo* embodied niche complex was not dependent on modern patterns of slow, delayed maturation and significant post-reproductive survival (Caspari and Lee, 2004, 2006; Smith et al., 2007a,b).

Here, the ENC hypothesis can elegantly explain the further coevolution of embodied niche with a hypothetical proto-language capacity (cf. Tomasello, 2008). The embodied dimension of the emergent hominin niche would have favored long-term selection for formally simple—and likely declarative but implicitly deictic—gestures or utterances that constituted symbolic representations. Such utterances would have been experienced as embodied sensory-motor concepts, whether they were initially verbal or brachial-manual gestures. Already defined by a complex web of visual, auditory, and motor-simulation associations that comprised a socially learned proto-language, the earliest verbal symbols would have further pointed to states of affairs in the environment, indexically tying together somatic and extrasomatic aspects of the prevailing milieu (Gallese and Lakoff, 2005; Tomasello, 2008). In turn, this would have synergistically reinforced learning complex foraging behaviors and social competence in networks sustained by strong reciprocity/reputation monitoring (Bowles and Gintis, 2004; Nowak and Sigmund, 2005). Linguistic communication would have shared a fundamental structural and functional similarity with omnivorous foraging, tool-making, and social cooperation (Stout et al., 2008). These embodied cognition/environment interfaces basically involve sustained attention on representation-dependent, goal-directed activity sequences, which may unfold over minutes or hours—with further potential for resuming attention on representation-dependent activities after intermittent breaks. Such embodied cognitive attention is hypothesized to reduce immediate sensory alertness levels, because the higher-level cognitive processes go beyond sensorio-motor cognitive simulation, involving more complex abstract narrative construction.

In light of the paleoanthropological evidence, the ENC hypothesis entails two mutually exclusive—but nonetheless plausible—evolutionary trajectories:

- Linguistic utterance and comprehension co-evolved gradually with the genus *Homo*'s embodied niche—following the establishment of the bipedal interface and the subsequent emergence of the manual precision/tool-making/tool-using interface—beginning as early as 3.0–2.0 mya (Schepartz, 1993; Deacon, 1998; Lieberman, 2013); or

- Spoken language rapidly evolved more recently (Chomsky, 1986), preceded by a long period of evolution in mental narrative construction and selective introspective attention—only later co-evolving with extant biological life history patterns of growth, maturation, aging and mortality (cf. Caspari and Lee, 2004, 2006; Smith et al., 2007a; Smith, 2013).

In either case, the evolutionarily derived hominin capacity for *narrativizing* simple iconic or symbolic semiotic representations would have evolved through by natural selection during the Pleistocene era. As argued throughout this hypothesis and theory article, it is predicted that the capacity to construct—and even recursively share representations of—embodied, iconically and symbolically represented narratives would have emerged via the dynamic bipedal and social monitoring/judgment interfaces that centrally define our terrestrial, extractive, and socially intensive extrasomatic environment.

### NARRATIVE REPRESENTATION AS EMBODIED TEMPORALITY

Theoretical approaches to human and non-human animal learning and behavior still strongly emphasize synchronic embodied representation or near-instantaneous cognitive feedback. Philosophers of cognition, experimental psychologists, and brain imaging experts have convincingly explained that embodied simulation connects affective and proprioceptive states or motor memories to extrasomatic stimuli (Barsalou, 1999, 2009; Damasio et al., 2004; Gallese and Lakoff, 2005; Heyes, 2010a,b; Dove, 2011; Man et al., 2012). Moreover, recurrent memory construction surely often shapes perceived extrasomatic phenomena or body-environment relationships, which specifically have a synchronic or short-term antecedent-consequent structure. Such embodied perceptual concepts may range from learning relevant, albeit complex figure-ground contrasts to more comprehensive, amodal learning of predator threat concepts or subtle indices of prey availability. Mapped onto Deacon's (1998) semiotic framework, such embodied concepts develop through bodily interfaces with the surrounding environment—specifically as *stable indexical relationship systems*. Cognized objects or bodily states immediately point to—or are pointed to by—other objects or somatic states.

If I am right that the embodied, private iconic narrative constitutes a key adaptive phenotype/embodied niche component in hominin evolution, then it would be especially important to consider the *narrativized temporality* of such indexical relationship systems. Quite different perspectives in recent human cognition scholarship have converged on highlighting embodied narratives and cognitive sensitivity to temporal duration and degrees of pastness (Hutto, 2008; Menary, 2008; Gallese, 2011; Panksepp and Biven, 2012). So far, I have emphasized two embodied iconic, non-linguistic narrative genres as important for the ENC hypothesis (for details, see section The Embodied Bipedal Terrestrial Niche: Infant Helplessness and Dialectical Cognitive Construction of Self and Other). The first is the exertion of agency, which is a narrative of self having—or failing to have—an effect on someone or something. The second is the achievement of judgment, in which self achieves—or fails to achieve—an unambiguous affective disposition concerning

someone or something else that she has been regarding in the extrasomatic environment. Such simple metacognitive narratives would have a scale-free, relative temporality: before, the situation was ambiguous or uncertain, but now, self has achieved a clear outcome. The relative chronological structure of these genres would reflect the most basic, temporally marked indexical relationship. On the one hand, a representation of the earlier, ambiguous past can point toward memory of reaching a clear outcome in the relatively more recent past. On the other hand, a prior representation of the remembered or imagined, more recent clear outcome could evoke the entire (possibly tragic, possibly triumphant) narrative arc, from earlier ambiguity to subsequent clarity. The very general, scale-free structure of the agency and judgment narratives has a series of remarkable implications:

- Iconic narrative representations may initially evoke the general, familiar feeling of attention on self's body-environment interaction, as it unfolds over short timescales, from fractions of a second to hours. Yet, they may recursively evoke a chain of chronologically ordered, nested, or longer continuous stories of agency or judgment. As such, the iconic narrative comprises the formal semiotic building block for representing diverse, complex pasts.
- The generic, elemental narrative form efficiently contextualizes and emphasizes diachronic changes in self's affective emotional states. As such, self's selective introspective or mindful attention on narrativized memories can autocondition complex learning. This may be especially important for developing competence at intricate, temporally extended, choppy, and contingent behavior sequences in the hominin terrestrial, extractive, heterotrophic, and intensely social niche.
- Simple iconic social judgment narratives have a particularly strong general conditioning effect on building complex stable indexical relationship systems. When self learns embodied understanding of a social other, her embodied cognitive construction process sets up a hierarchically constituted triadic association—one that is very familiar to structural anthropology, in which two terms are equated or opposed, and in turn, that dyadic relationship is metonymically indexed to a third, generally evocative term (Lévi-Strauss, 1963; Leach, 1976). Here, the socially interacting or regarding agent perceives that the other is either similar or dissimilar to herself, in some markedly evocative respect. Her embodied experience provides rich synchronic information, involving her own perceptions of the grounding environment in which the other is defined; her prevailing proprioceptive or interoceptive states; and memory content evoked by the situation at hand. This embodied information is material for reinforcing or metonymically tagging the experience that other is—or is not—like self. As an embodied concept that integrates perception of other with evoked, sensory-motor/emotional simulated experience, the understood relationship between self and other has its own rich potential to evoke—that is, to approach synchronic juxtaposition with—learned concepts that have similar experiential tags or indices.

- The auto-indexical connection between the narratives and the emotional changes embedded in them further recursively rewards introspection and the mental consideration of plausible and implausible associations, alike, so that narrativized embodied representations mediate changes in self's affective states over time.

In general, through recursive iconic narrative construction, embodied cognition can “pancake” or elide diachronic features of the narrative's content, yielding new synchronically represented associations among action sequences, changes in the extrasomatic environment, and changes in bodily affective states (Table 2).

### ICONIC NARRATIVES AS EVOLUTIONARY PRECURSORS TO SOCIO-LINGUISTIC CONSTRUCTIONS

The theoretically plausible—albeit preliminary and speculative—argument for the ENC hypothesis, presented above, also implies the following. As joint adaptive phenotype and embodied niche

component, *recursive iconic narrative construction* was evolutionarily ancestral to natural open language systems, with their fundamental feature of “double articulation.” Logically consistent patterns governing the orderly juxtaposition (that is, the articulation) of message features require that similar, substitutable elements—on the dual, nested levels of phonological regularities and syntactical structures—are available (Jakobson and Halle, 1956; Hockett, 1961; Lévi-Strauss, 1963; Saussure, 2011). Double articulation is an apparent formal and functional necessity for natural languages. More generally, though, interaction between message-feature *contiguity* and *similarity* makes double articulation possible (Jakobson and Halle, 1956; Lévi-Strauss, 1962a). The recursive interplay between contiguity and similarity is, in turn, sufficient to generate the openness of natural language, in which a finite set of signs may be combined, repeated, and substituted to generate potentially infinite expressions or representations (Chomsky, 1986; Hauser et al., 2002). Thus, contiguity-similarity interaction can structure and be structured by hierarchically nested message elements, in which

**Table 2 | Matrix of hypothetical emotional and affective trajectories of the two main proposed genres of embodied iconic narratives emerging in human evolution.**

		Object of Self's Attention in Narrative		
		<b>THING</b> (Tool or Food)	<b>SOCIAL OTHER</b>	<b>SELF</b>
Embodied Narrative Genre	<b>ATTEMPT TO EXERT AGENCY</b>	FAILURE: Hope, seeking, & high extra-somatic alertness → frustration, disappointment, anxiety	FAILURE or TRIUMPH: Motivating Emotions & Associations → Consequential Emotions & Associations	FAILURE: Hope & seeking <u>but</u> reduced extrasomatic alertness → frustration, disappointment, anxiety
		TRIUMPH: Hope, seeking, high alertness → satisfaction/satiation & reduced alertness	Highly context- and history dependent	TRIUMPH: Hope, seeking, <u>but</u> reduced extrasomatic alertness → satisfaction/elation & increased alertness
	<b>ATTEMPT TO REACH JUDGMENT</b> (embodied affective approach to ambiguity)	FAILURE: Seeking, possible anxiety & high extra-somatic alertness → continued anxiety, reduced alertness  TRIUMPH: Seeking, possible anxiety & high extra-somatic alertness → Satisfaction, increased alertness of introspective/extra-somatic environmental interface—  ➤ <u>recursively emerging introspective awareness that attempting to exert agency is relevant</u>		

one or more element-levels incorporate coherent, independent messages. In other words, double articulation at the phoneme and lexico-grammatical levels may be understood as an instance of our more general capacity for embodied cognitive recursion—a capacity that emerged gradually in hominin evolution, beginning well prior to language evolution itself.

## FROM THEORY TO PRACTICE: ANTHROPOLOGICAL PERSPECTIVES ON TESTING THE ENC HYPOTHESIS

The ENC hypothesis has particular potential to connect general anthropological (including ethnographic), paleoanthropological, cognitive science, and comparative experimental psychological approaches to human cognition and embodied experience. My aim in this article has been to outline the ENC hypothesis in some detail and attempt to establish its relevance for explaining the evolution of human cognition over very long timeframes. This simply reflects my paleoanthropological research specialization. However, my hope is that the ENC hypothesis can guide collaboration among anthropologists, cognitive scientists, and experimental psychologists, redefining and expanding theories, evolutionary perspectives, and observational and experimental designs. We should be able to evaluate whether the ENC hypothesis provides a more reliable, comprehensive explanation of how abstract or “dis-embodied” concepts (Dove, 2011)—as dynamic features of our semiotically structured worlds—might arise from, yet remain grounded in indexical associations with embodied sensory-motor representations and iconic memories. Perhaps the most concrete, testable prediction of the ENC hypothesis is that structured measurable changes in affect, emotional state, and sensory attentiveness should occur over brief time periods, as subjects experience and focus their introspective attention on embodied iconic narrative representations (see Table 2). More broadly, we should be able to measure how culturally contextualized narratives and relationships, rituals, learning tasks, skilled artistic or craft production, and responses to scenarios involving culturally relevant power structures influence neural activity, vital rates, pupil dilation, and hormone levels within and between study groups defined by biological life history stages and culturally relevant identities.

## CONCLUSION

If the embodied cognition theoretical framework explains behavioral, central nervous system, and conceptual or representational phenomena better than strictly computational brain models—and certainly better than “disembodied” theories of mind (as argued by Clark, 2008 and Dove, 2011)—then paleoanthropological inquiry would benefit from embodied cognition research. Such interdisciplinary borrowing would facilitate investigating how unique-derived hominin brain anatomy and behavior patterns evolved, potentially helping to demystify the prehistoric emergence of language, symbolic representation, and the conscious human mind (Barton, 2012). In this article I have argued that ENC in the hominin lineage has involved a distinctive, semiotically structured and structuring interface between the body and the extrasomatic environment. This interface is constituted by narratives that are at once embodied and semiotically constructed, at once cognitive adaptations and embodied niche

components. As we expand our perspective to view embodied cognition and interaction with the environment as both phenotype and niche, the ENC hypothesis can help to clarify the long-term evolutionary process through which human biology, semiotically structured worlds, and embodied experiences have emerged.

## ACKNOWLEDGMENTS

I thank special issue editor Guy Dove for his encouragement, patience, and critical comments as I developed the initial abstract and various revisions of this article. I am especially grateful for the critical and very constructive comments from Caroline VanSickle, Liv Nilsson Stutz, and an anonymous reviewer.

## REFERENCES

- Agamben, G. (1999). *The Man Without Content*. Stanford, CA: Stanford University Press.
- Allen, T. F. H., and Starr, T. B. (1982). *Hierarchy: Perspectives for Ecological Complexity*. Chicago, IL: University of Chicago Press.
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behav. Brain Sci.* 22, 637–660. doi: 10.1017/S0140525X99532147
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1281–1289. doi: 10.1098/rstb.2008.0319
- Barton, R. A. (2012). Embodied cognitive evolution and the cerebellum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2097–2107. doi: 10.1098/rstb.2012.0112
- Bourdieu, P. (1977). *Outline of a Theory of Practice*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511812507
- Bowles, S., and Gintis, H. (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theor. Popul. Biol.* 65, 17–28. doi: 10.1016/j.tpb.2003.07.001
- Caspari, R., and Lee, S.-H. (2004). Older age becomes common late in human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10895–10900. doi: 10.1073/pnas.0402857101
- Caspari, R., and Lee, S.-H. (2006). Is human longevity a consequence of cultural change or modern biology? *Am. J. Phys. Anthropol.* 129, 512–517. doi: 10.1002/ajpa.20360
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Westport, CT: Greenwood Publishing Group.
- Clark, A. (1993). *Associative Engines: Connectionism, Concepts, and Representational Change*. Cambridge, MA: MIT Press.
- Clark, A. (1998). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension: Embodiment, Action, and Cognitive Extension*. New York, NY: Oxford University Press.
- Coward, F., and Gamble, C. (2008). Big brains, small worlds: material culture and the evolution of the mind. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1969–1979. doi: 10.1098/rstb.2008.0004
- Damasio, A. (2003). *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Orlando, FL: Houghton Mifflin Harcourt.
- Damasio, A. (2008). *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY: Random House.
- Damasio, A. R. (2000). *The Feeling of what Happens: Body and Emotion in the Making of Consciousness*. Orlando, FL: Houghton Mifflin Harcourt.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., and Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition* 92, 179–229. doi: 10.1016/j.cognition.2002.07.001
- Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*. Oxford; New York: Oxford University Press.
- Deacon, T. W. (1998). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York, NY: W. W. Norton & Company.
- DeSilva, J. M. (2011). A shift toward birthing relatively large infants early in human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1022–1027. doi: 10.1073/pnas.1003865108



- DeSilva, J. M., Holt, K. G., Churchill, S. E., Carlson, K. J., Walker, C. S., Zipfel, B., et al. (2013). The lower limb and mechanics of walking in *Australopithecus sediba*. *Science* 340:1232999. doi: 10.1126/science.1232999
- DeSilva, J. M., and Lesnik, J. J. (2008). Brain size at birth throughout human evolution: a new method for estimating neonatal brain size in hominins. *J. Hum. Evol.* 55, 1064–1074. doi: 10.1016/j.jhevol.2008.07.008
- Dietz, V. (2003). Spinal cord pattern generators for locomotion. *Clin. Neurophysiol.* 114, 1379–1389. doi: 10.1016/S1388-2457(03)00120-2
- Dietz, V. (2010). “Organization of human locomotion: proprioception and quadrupedal coordination,” in *Encyclopedia of Behavioral Neuroscience*, eds G. F. Koob, M. L. Moal, and R. F. Thompson (Oxford: Academic Press), 490–496. Available online at: <http://www.sciencedirect.com/science/article/pii/B9780080453965002153> (Accessed April 25, 2014).
- Dominguez-Rodrigo, M., Rayne Pickering, T., Semaw, S., and Rogers, M. J. (2005). Cutmarked bones from Pliocene archaeological sites at Gona, Afar, Ethiopia: implications for the function of the world's oldest stone tools. *J. Hum. Evol.* 48, 109–121. doi: 10.1016/j.jhevol.2004.09.004
- Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Front. Cogn.* 1:242. doi: 10.3389/fpsyg.2010.00242
- Dove, G. (2012). Grammar as a developmental phenomenon. *Biol. Philos.* 27, 615–637. doi: 10.1007/s10539-012-9324-4
- Dunbar, R. I. M. (2003). THE SOCIAL BRAIN: mind, language, and society in evolutionary perspective. *Annu. Rev. Anthropol.* 32, 163–181. doi: 10.1146/annurev.anthro.32.061002.093158
- Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Ann. Hum. Biol.* 36, 562–572. doi: 10.1080/03014460902960289
- Gallese, V. (2011). Embodied simulation theory: imagination and narrative. *Neuropsychanalysis* 13, 196–200. doi: 10.1080/15294145.2011.10773675
- Gallese, V., and Lakoff, G. (2005). The Brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310
- Geertz, C. (2001). *Available Light: Anthropological Reflections on Philosophical Topics*. Princeton, NJ: Princeton University Press.
- Gell, A. (1998). *Art and Agency: An Anthropological Theory*. Oxford: Clarendon Press.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Orlando, FL: Houghton Mifflin.
- Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley, CA: University of California Press.
- Goldman-Neuman, T., and Hovers, E. (2009). “Methodological considerations in the study of oldowan raw material selectivity: insights from A. L. 894 (Hadar, Ethiopia),” in *Interdisciplinary Approaches to the Oldowan Vertebrate Paleobiology and Paleoanthropology*, eds E. Hovers and D. R. Braun (Springer), 71–84. Available online at: [http://link.springer.com/chapter/10.1007/978-1-4020-9060-8\\_7](http://link.springer.com/chapter/10.1007/978-1-4020-9060-8_7) (Accessed February 14, 2014).
- Gunderson, L. H., and Holling, C. S. (eds.). (2001). *Panarchy: Understanding Transformations in Human and Natural Systems*. Washington, DC: Island Press.
- Haile-Selassie, Y., Latimer, B. M., Alene, M., Deino, A. L., Gibert, L., Melillo, S. M., et al. (2010). An early *Australopithecus afarensis* postcranium from Woranso-Mille, Ethiopia. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12121–12126. doi: 10.1073/pnas.1004527107
- Haile-Selassie, Y., and WoldeGabriel, G. (2009). *Ardipithecus Kadabba: Late Miocene Evidence from the Middle Awash*. University of California Press.
- Hamilton, W. D. (1971). Geometry for the selfish herd. *J. Theor. Biol.* 31, 295–311. doi: 10.1016/0022-5193(71)90189-5
- Hart, D., and Sussman, R. W. (2011). “The influence of predation on primate and early human evolution: impetus for cooperation,” in *Origins of Altruism and Cooperation Developments in Primatology: Progress and Prospects*, eds R. W. Sussman and C. R. Cloninger (New York, NY: Springer), 19–40. Available online at: [http://link.springer.com/chapter/10.1007/978-1-4419-9520-9\\_3](http://link.springer.com/chapter/10.1007/978-1-4419-9520-9_3) (Accessed February 14, 2014).
- 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
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M., Watumull, J., et al. (2014). The mystery of language evolution. *Lang. Sci.* 5, 401. doi: 10.3389/fpsyg.2014.00401
- Hawkes, K. (2003). Grandmothers and the evolution of human longevity. *Am. J. Hum. Biol.* 15, 380–400. doi: 10.1002/ajhb.10156
- Hawkes, K., O'Connell, J. F., Jones, N. G. B., Alvarez, H., and Charnov, E. L. (1998). Grandmothering, menopause, and the evolution of human life histories. *Proc. Natl. Acad. Sci. U.S.A.* 95, 1336–1339. doi: 10.1073/pnas.95.3.1336
- Heyes, C. (2010a). Mesmerising mirror neurons. *Neuroimage* 51, 789–791. doi: 10.1016/j.neuroimage.2010.02.034
- Heyes, C. (2010b). Where do mirror neurons come from? *Neurosci. Biobehav. Rev.* 34, 575–583. doi: 10.1016/j.neubiorev.2009.11.007
- Heyes, C. (2012). Grist and mills: on the cultural origins of cultural learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2181–2191. doi: 10.1098/rstb.2012.0120
- Hockett, C. F. (1961). Linguistic elements and their relations. *Language* 37, 29–53. doi: 10.2307/411248
- Hrdy, S. B. (2009). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge, MA: Harvard University Press.
- Humphrey, N. (1992). *A History of the Mind*. New York, NY: Harper Collins.
- Hutchins, E. (1996). *Cognition in the Wild*. Cambridge, MA: The MIT Press.
- Hutto, D. D. (2008). *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons*. Cambridge, MA: MIT Press.
- Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: a review. *Neural Netw.* 21, 642–653. doi: 10.1016/j.neunet.2008.03.014
- Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends Cogn. Sci.* 3, 272–279. doi: 10.1016/S1364-6613(99)01333-9
- Jakobson, R., and Halle, M. (1956). *Fundamentals of Language*. Oxford: Mouton & Co.
- Jungers, W. L. (1982). Lucy's limbs: skeletal allometry and locomotion in *Australopithecus afarensis*. *Nature* 297, 676–678. doi: 10.1038/297676a0
- Kimbel, W. H. (2009). “The Origin of Homo,” in *The First Humans—Origin and Early Evolution of the Genus Homo Vertebrate Paleobiology and Paleoanthropology*, eds F. E. Grine, J. G. Fleagle, and R. E. Leakey (Springer), 31–37. Available online at: [http://link.springer.com/chapter/10.1007/978-1-4020-9980-9\\_4](http://link.springer.com/chapter/10.1007/978-1-4020-9980-9_4) (Accessed February 14, 2014).
- Kimbel, W. H., Johanson, D. C., and Rak, Y. (1997). Systematic assessment of a maxilla of Homo from Hadar, Ethiopia. *Am. J. Phys. Anthropol.* 103, 235–262.
- Kracke, W. H. (2007). “To dream, perchance to cure: dreaming and shamanism in a Brazilian indigenous society,” in *Explorations in Psychoanalytic Ethnography*, ed J. Mimica (New York, NY: Berghahn Books), 106–120.
- Kurki, H. K. (2011). Pelvic dimorphism in relation to body size and body size dimorphism in humans. *J. Hum. Evol.* 61, 631–643. doi: 10.1016/j.jhevol.2011.07.006
- Kurki, H. K. (2013). Bony pelvic canal size and shape in relation to body proportionality in humans. *Am. J. Phys. Anthropol.* 151, 88–101. doi: 10.1002/ajpa.22243
- Langdon, J. H. (2005). *The Human Strategy: An Evolutionary Perspective on Human Anatomy*. New York, NY: Oxford University Press.
- Latimer, B., and Lovejoy, C. O. (1990). Metatarsophalangeal joints of *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* 83, 13–23. doi: 10.1002/ajpa.1330830103
- Latimer, B., Ohman, J. C., and Lovejoy, C. O. (1987). Talocrural joint in African hominoids: implications for *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* 74, 155–175. doi: 10.1002/ajpa.1330740204
- Leach, E. (1976). *Culture and Communication: The Logic by which Symbols are Connected. An Introduction to the Use of Structuralist Analysis in Social Anthropology*. Cambridge University Press.
- Lee, S.-H., and Wolpoff, M. H. (2003). The pattern of evolution in Pleistocene human brain size. *Paleobiology* 29, 186–196. doi: 10.1666/0094-8373(2003)029%3C0186:TPOEIP%3E2.0.CO;2
- Leonard, W. R., and Robertson, M. L. (1997). Comparative primate energetics and hominid evolution. *Am. J. Phys. Anthropol.* 102, 265–281.
- Levin, S. A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. *Ecology* 73, 1943–1967.
- Lévi-Strauss, C. (1962a). *La Pensée Sauvage: Avec 24 Illustrations Dans le Texte*. Paris: Plon.
- Lévi-Strauss, C. (1962b). *Le Totémisme Aujourd'hui*. Paris: Presses Universitaires de France.
- Lévi-Strauss, C. (1963). *Structural Anthropology*. New York, NY: Basic Books.
- Lieberman, D. (2013). *The Story of the Human Body: Evolution, Health, and Disease*. New York, NY: Random House LLC.
- Lovejoy, C. O. (1981). The origin of man. *Science* 211, 341–350. doi: 10.1126/science.211.4480.341

- Lovejoy, C. O. (1988). Evolution of human walking. *Sci. Am.* 259, 118–125. doi: 10.1038/scientificamerican1188-118
- Lovejoy, C. O. (2009). Reexamining human origins in light of *Ardipithecus ramidus*. *Science* 326, 74–74e8. doi: 10.1126/science.1175834
- Lovejoy, C. O., Suwa, G., Spurluck, L., Asfaw, B., and White, T. D. (2009). The pelvis and femur of *Ardipithecus ramidus*: the emergence of upright walking. *Science* 326, 71–71e6. doi: 10.1126/science.1175831
- Man, K., Kaplan, J. T., Damasio, A., and Meyer, K. (2012). Sight and sound converge to form modality-invariant representations in temporoparietal cortex. *J. Neurosci.* 32, 16629–16636. doi: 10.1523/JNEUROSCI.2342-12.2012
- Marzke, M. W. (1997). Precision grips, hand morphology, and tools. *Am. J. Phys. Anthropol.* 102, 91–110.
- Marzke, M. W., and Marzke, R. F. (2000). Evolution of the human hand: approaches to acquiring, analysing and interpreting the anatomical evidence. *J. Anat.* 197, 121–140. doi: 10.1046/j.1469-7580.2000.19710121.x
- McHenry, H. M. (1992). Body size and proportions in early hominids. *Am. J. Phys. Anthropol.* 87, 407–431. doi: 10.1002/ajpa.1330870404
- Menary, R. (2008). Embodied narratives. *J. Conscious. Stud.* 15, 63–84.
- Nilsson Stutz, L. (2003). *Embodied Rituals and Ritualized Bodies: Tracing Ritual Practices in Late Mesolithic Burials*. Acta Archaeologica Lundensia Series, Vol. 46. Stockholm: Almqvist & Wiksell International.
- Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437, 1291–1298. doi: 10.1038/nature04131
- O'Connell, J. F., Hawkes, K., and Blurton Jones, N. G. (1999). Grandmothering and the evolution of *Homo erectus*. *J. Hum. Evol.* 36, 461–485. doi: 10.1006/jhev.1998.0285
- O'Connell, J. F., Hawkes, K., Lupo, K. D., and Blurton Jones, N. G. (2002). Male strategies and plio-pleistocene archaeology. *J. Hum. Evol.* 43, 831–872. doi: 10.1006/jhev.2002.0604
- Odling-Smee, F. J., Feldman, M. W., and Laland, K. N. (2003). *Niche Construction: The Neglected Process in Evolution*. Princeton, NJ: Princeton University Press.
- Ohman, J. C., Krochta, T. J., Lovejoy, C. O., Mensforth, R. P., and Latimer, B. (1997). Cortical bone distribution in the femoral neck of hominoids: implications for the locomotion of *Australopithecus afarensis*. *Am. J. Phys. Anthropol.* 104, 117–131.
- Ortner, S. B. (1984). Theory in anthropology since the sixties. *Comp. Stud. Soc. Hist.* 26, 126–166. doi: 10.1017/S0010417500010811
- Ortner, S. B. (2006). *Anthropology and Social Theory: Culture, Power, and the Acting Subject*. Durham, NC: Duke University Press.
- Panksepp, J., and Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions* (Norton Series on Interpersonal Neurobiology). New York, NY: W. W. Norton & Company.
- Peirce, C. S. (2012). *Charles S. Peirce, Selected Writings*. Wiener Dover Publications.
- Pfeifer, R., Bongard, J., and Grand, S. (2007). *How the Body Shapes the Way we Think: A New View of Intelligence*. Cambridge, MA: MIT Press.
- Rosenberg, K., and Trevathan, W. (1995). Bipedalism and human birth: the obstetrical dilemma revisited. *Evol. Anthropol. Issues News Rev.* 4, 161–168. doi: 10.1002/evan.1360040506
- Rowlands, (2006). *Body Language*. Cambridge, MA: MIT Press.
- Ruff, C. (2010). Body size and body shape in early hominins—implications of the Gona Pelvis. *J. Hum. Evol.* 58, 166–178. doi: 10.1016/j.jhev.2009.10.003
- Ruff, C. B., Trinkaus, E., and Holliday, T. W. (1997). Body mass and encephalization in Pleistocene *Homo*. *Nature* 387, 173–176. doi: 10.1038/387173a0
- Sahlins, M. D. (2000). *Culture in Practice: Selected Essays*. Brooklyn, NY: Zone Books.
- Saussure, F. de (2011). *Course in General Linguistics*. Transl. by: W. Baskin, eds P. Meisel and H. Saussy. New York, NY: Columbia University press.
- Schepartz, L. A. (1993). Language and modern human origins. *Am. J. Phys. Anthropol.* 36, 91–126. doi: 10.1002/ajpa.1330360607
- Schneider, D. C. (2010). *Quantitative Ecology: Measurement, Models and Scaling*, 2nd Edn. Burlington, MA: Elsevier.
- Semaw, S. (2000). The world's oldest stone artefacts from gona, ethiopia: their implications for understanding stone technology and patterns of human evolution between 2.6–1.5 million years ago. *J. Archaeol. Sci.* 27, 1197–1214. doi: 10.1006/jasc.1999.0592
- Semaw, S., Renne, P., Harris, J. W. K., Feibel, C. S., Bernor, R. L., Fesseha, N., et al. (1997). 2.5-million-year-old stone tools from Gona, Ethiopia. *Nature* 385, 333–336. doi: 10.1038/385333a0
- Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28, 1070–1094. doi: 10.1016/S0003-3472(80)80097-2
- Smith, T. M. (2013). Teeth and human life-history evolution\*. *Annu. Rev. Anthropol.* 42, 191–208. doi: 10.1146/annurev-anthro-092412-155550
- Smith, T. M., Tafforeau, P., Reid, D. J., Grün, R., Eggers, S., Boutakiout, M., et al. (2007a). Earliest evidence of modern human life history in North African early *Homo sapiens*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6128–6133. doi: 10.1073/pnas.0700747104
- Smith, T. M., Toussaint, M., Reid, D. J., Olejniczak, A. J., and Hublin, J.-J. (2007b). Rapid dental development in a Middle Paleolithic Belgian Neanderthal. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20220–20225. doi: 10.1073/pnas.0700747104
- Sockol, M. D., Raichlen, D. A., and Pontzer, H. (2007). Chimpanzee locomotor energetics and the origin of human bipedalism. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12265–12269. doi: 10.1073/pnas.0703267104
- Sperber, D., and Wilson, D. (1995). *Relevance: Communication and Cognition*, 2nd Edn. Cambridge, MA: Blackwell.
- Stephan, C., and Zuberbühler, K. (2008). Predation increases acoustic complexity in primate alarm calls. *Biol. Lett.* 4, 641–644. doi: 10.1098/rsbl.2008.0488
- Stout, D. (2011). Stone toolmaking and the evolution of human culture and cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1050–1059. doi: 10.1098/rstb.2010.0369
- Stout, D., Quade, J., Semaw, S., Rogers, M. J., and Levin, N. E. (2005). Raw material selectivity of the earliest stone toolmakers at Gona, Afar, Ethiopia. *J. Hum. Evol.* 48, 365–380. doi: 10.1016/j.jhev.2004.10.006
- Stout, D., Toth, N., Schick, K., and Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1939–1949. doi: 10.1098/rstb.2008.0001
- Stutz, A. J. (2012). Culture and politics, behavior and biology: seeking synthesis among fragmentary anthropological perspectives on hunter-gatherers. *Rev. Anthropol.* 41, 23–69. doi: 10.1080/00938157.2012.649183
- Sussman, R. W., Garber, P. A., and Cheverud, J. M. (2005). Importance of cooperation and affiliation in the evolution of primate sociality. *Am. J. Phys. Anthropol.* 128, 84–97. doi: 10.1002/ajpa.20196
- Tague, R. G., and Lovejoy, C. O. (1986). The obstetric pelvis of A.L. 288-1 (Lucy). *J. Hum. Evol.* 15, 237–255. doi: 10.1016/S0047-2484(86)80052-5
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge; London: MIT Press.
- Tüma, P., and Brůžek, J. (2013). “Neanderthal neonatal brain size estimating: validity of DeSilva and Lesnik (2008): Models and new proposals,” in *Poster Presentation* (Honolulu, Hawaii). Available online at: [http://www.paleoanthro.org/static/posters2013/Tuma+Bruezek\\_2013.pdf](http://www.paleoanthro.org/static/posters2013/Tuma+Bruezek_2013.pdf)
- Van de Crommert, H. W. A. A., Mulder, T., and Duysens, J. (1998). Neural control of locomotion: sensory control of the central pattern generator and its relation to treadmill training. *Gait Posture* 7, 251–263. doi: 10.1016/S0966-6362(98)00010-1
- VanSickle, C. (2009). Re-sampled human cranial capacities compared to würm neandertals. in *Poster Presentation* (Chicago, Illinois).
- Wall-Scheffler, C. M. (2012). Energetics, locomotion, and female reproduction: implications for human evolution. *Annu. Rev. Anthropol.* 41, 71–85. doi: 10.1146/annurev-anthro-092611-145739
- Webster, M. T. (2009). Patterns of autosomal divergence between the human and chimpanzee genomes support an allopatric model of speciation. *Gene* 443, 70–75. doi: 10.1016/j.gene.2009.05.006
- Wells, J. C. K., DeSilva, J. M., and Stock, J. T. (2012). The obstetric dilemma: an ancient game of Russian roulette, or a variable dilemma sensitive to ecology? *Am. J. Phys. Anthropol.* 149(Suppl. 55), 40–71. doi: 10.1002/ajpa.22160
- Wheeler, P. E. (1991a). The influence of bipedalism on the energy and water budgets of early hominids. *J. Hum. Evol.* 21, 117–136.
- Wheeler, P. E. (1991b). The thermoregulatory advantages of hominid bipedalism in open equatorial environments: the contribution of increased convective heat loss and cutaneous evaporative cooling. *J. Hum. Evol.* 21, 107–115.

- Won, Y.-J. (2004). Divergence population genetics of chimpanzees. *Mol. Biol. Evol.* 22, 297–307. doi: 10.1093/molbev/msi017
- Yamamichi, M., Gojobori, J., and Innan, H. (2011). An autosomal analysis gives no genetic evidence for complex speciation of humans and chimpanzees. *Mol. Biol. Evol.* 29, 145–156. doi: 10.1093/molbev/msr172
- Zuberbühler, K. (2001). Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behav. Ecol. Sociobiol.* 50, 414–422. doi: 10.1007/s002650100383

**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.

Received: 14 February 2014; accepted: 12 July 2014; published online: 01 August 2014.  
Citation: Stutz AJ (2014) Embodied niche construction in the hominin lineage: semiotic structure and sustained attention in human embodied cognition. *Front. Psychol.* 5:834. doi: 10.3389/fpsyg.2014.00834

This article was submitted to Cognitive Science, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Stutz. 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



### FAST PUBLICATION

Average 90 days  
from submission  
to publication



### COLLABORATIVE PEER-REVIEW

Designed to be rigorous –  
yet also collaborative, fair and  
constructive



### RESEARCH NETWORK

Our network  
increases readership  
for your article



### OPEN ACCESS

Articles are free to read,  
for greatest visibility



### TRANSPARENT

Editors and reviewers  
acknowledged by name  
on published articles



### GLOBAL SPREAD

Six million monthly  
page views worldwide



### COPYRIGHT TO AUTHORS

No limit to  
article distribution  
and re-use



### IMPACT METRICS

Advanced metrics  
track your  
article's impact



### SUPPORT

By our Swiss-based  
editorial team