Scientometrics
https://doi.org/10.1007/s11192-018-2811-x

Exploratory mapping of theoretical landscapes through word use in abstracts

Pablo Contreras Kallens[1,2] · **Rick Dale**[3]

Received: 15 December 2017

Abstract
We present a case study of how scientometric tools can reveal the structure of scientific theory in a discipline. Specifically, we analyze the patterns of word use in the discipline of cognitive science using latent semantic analysis, a well-known semantic model, in the abstracts of over a thousand academic papers relevant to these theories. Our results show that it is possible to link these theories with specific statistical distributions of words in the abstracts of papers that espouse these theories. We show that theories have different patterns of word use, and that the similarity relationships with each other are intuitive and informative. Moreover, we show that it is possible to predict fairly accurately the theory of a paper by constructing a model of the theories based on their distribution of word use. These results may open new avenues for the application of scientometric tools on theoretical divides.

Keywords Latent semantic analysis · Cognitive science · Text analysis · Theoretical issues

Introduction

Cognitive science is the interdisciplinary study of mind. It is about 50 years old, depending on how you count (Bechtel and Graham 1998). Its theoretical progress may be predicted by this youthful age—it has seen the rise of fall of various conflicts, and has not yet established a consensus identity in its aims and foundations. This is to be expected in a young field, of course. Its conflict often derives from pointed, yes-no theoretical questions: Can we analyze the brain much like a computer? Or does that metaphor ultimately fail, and need contributions from other ideas? Is the mind fully encapsulated in the cranium, or does it also fundamentally involve the body? Should social processes play a critical role in our theories of the mind? These theoretical points of conflict have been famously lamented

✉ Pablo Contreras Kallens
pc684@cornell.edu

1 Department of Psychology, Cornell University, Ithaca, NY 14853, USA

2 Cognitive and Information Sciences, University of California, Merced, USA

3 Department of Communication, University of California, Los Angeles, USA

Published online: 02 July 2018

early in its history (Newell 1973). To a great extent, these debates are ongoing. Our goal in this paper is to test whether the tools of scientometrics can help quantify this theoretical landscape of cognitive science. We develop a semantic model from a small corpus of abstracts, and explore the quantitative relationship among different perspectives. We hope that a scientometric approach can facilitate grounding the theoretical discussion in operationalized patterns of scientific behavior. We show that indeed the theories of cognitive science may be usefully quantified in this manner. In what follows, we first provide further background on this issue, and summarize our computational approach. Following this, we take a first step towards quantifying the theoretical landscape of cognitive science.

Background and motivation

Cognitive science has changed considerably since the relative dominance of the information-processing approach that characterized the 1980s and 1990s (Von Eckardt 1995). Core assumptions of the discipline have been put into contention, such as the extent to which computation as a metaphor for mental activity is useful for its explanation (Van Gelder 1995), the role of representation (Chemero 2011) or the relevant unit of analysis for its study (Hollan et al. 2000). This tension has sprung from the appearance of new theoretical views that challenge some or all of the traditional assumptions that characterized the discipline (Clark 1998). Moreover, the situation has manifested itself through substantive disagreements that have not yet been resolved, such as how much we consider the environment to be part and parcel of the cognitive system (Adams and Aizawa 2010; de Oliveira and Chemero 2015), the role of the body in mental simulation (Zwaan 2014), and the implications of the effective use of dynamical systems tools in explaining cognition (Bechtel and Abrahamsen 2006). Despite several proposals for overcoming this theoretical quandary (Dale 2008; Edelman 2008; Louwerse 2011; McCauley and Bechtel 2001; Yoshimi 2012; Zwaan 2014), cognitive science remains relatively fractionated theoretically.[1]

Our goal in this paper is to explore a possible quantitative framing of this problem that can act as a supplement to qualitative and argument-based accounts of the theoretical landscape of cognitive science. We use bibliometric data, specifically word use in abstracts, to map the theoretical perspectives and their relationships. We test a *latent semantic* model of the field, observing how it represents the different features of these theories, how they relate to each other, how the semantic space underlying word use relates to theories, and how these different semantic dimensions could be characterized. With this, we aim to take a first step toward a quantitative understanding of cognitive science's conceptual landscape.

Theories of cognitive science

The theoretical landscape of cognitive science

As stated above, our goal is to explore the theoretical divisions of cognitive science using quantitative tools. To motivate this exploration, an important first step is to identify some intuitive notions of this theoretical landscape. Here we identify some common traditions that have framed historical and more recent cognitive science. Although there are many

[1] For an example of such an overarching division, see the exchange between McClelland et al. (2010) and Griffiths et al. (2010).

taxonomies of the different theoretical approaches currently coexisting within the discipline, the general movement questioning core assumptions that appeared mainly during the 1990s is often labeled as *embodied cognition* (Rowlands 2010; Shapiro 2010; Ziemke 2003; Gomila and Calvo 2008). The broad movement that confronts or minimizes the relevance of embodiment can be identified with the label *cognitivism* (Varela et al. 2017, p. 37; Haugeland 1978). However, within these broad perspectives there exist several more specific stances on the key issues of cognition. For the present study, we focus on 8 perspectives, providing a brief description and canonical references for each:

- The **symbolic** approach to cognitive science is often regarded as the longest-standing modern perspective, in which cognition is interpreted as computations over internal states imbued with semantic properties (Fodor 1975; Haugeland 1981).
- **Connectionism** represents cognition as emerging from the computational activity of a brain-inspired network of interconnected nodes, with parallel distributed processing and neural network models serving as the primary mechanisms of this tradition (Rumelhart et al. 1987; Feldman and Ballard 1982; Smolensky 1988).
- **Bayesian** approaches have grown prominent in the past decade. These approaches model cognition as rational probabilistic inference (Tenenbaum et al. 2011; Hohwy 2013; Oaksford and Chater 2009; Tenenbaum et al. 2006).
- **Embodied cognition** is often an umbrella phrase used to refer to several more specific approaches in varying degrees of disagreement with cognitivism (Wilson 2002), particularly the remaining approaches on this list. It can also refer to a more specific approach that focuses on showing how cognitive processes depend on bodily factors (Barsalou 1999; Gibbs 2005).
- **Distributed cognition** focuses on processes that happen at a scale larger than the individual, and treats these processes as an important part of cognition (Hollan et al. 2000; Hutchins 1995; Cowley 2011).
- **Enactive** approaches put an emphasis on the self-organization of organisms, and how actively integrating with the environment generates meaningful albeit much simpler internal processes than those typically posited by traditional symbolic approaches (Di Paolo 2005; Stewart et al. 2010).
- **Dynamical models of cognition** posit that dynamical systems, and their modeling through differential equations, are a useful model for the organization of cognition, complementing, or even replacing, computation (Beer 1995; Thelen and Smith 1996; Spivey 2008).
- **Ecological psychology**, heavily influenced by the work of (Gibson 1979), contends that "cognition" includes both the organism and its environment. The structuring of an environment, and the close coupling of organisms to their environmental niches, allow for the direct perception of possibilities for action. This approach often seeks to explain behavior without appealing to internal representations (Kugler and Turvey 2015; Richardson et al. 2008; Gibson and Pick 2000; Michaels and Carello 1981).

By focusing on these particular traditions, the theoretical panorama of cognitive science appears much more complicated than what the overarching narrative of new "embodied" theories versus old "cognitivist" theories suggests. Instead, the theoretical landscape of cognitive science is articulated by different concepts and stances on core problems in the discipline. There have been attempts to organize these different strands. As an example, Chemero and Silberstein (2008) organize the discipline as a tree around several branches, according to the answers they provide to questions such as the role of representation and the correct unit of analysis for cognitive science. Other works have focused on specific issues, such as the role of the individual (Wilson 2004), the content of representations (Anderson 2003), the proper contribution of the body for cognition (Wilson 2002; Shapiro 2010), or the differences in explanatory schemes used by different theories (Stepp et al. 2011). Nevertheless, much discussion of the theoretical landscape of cognitive science has remained qualitative, often involving intuitive prescriptions rather than quantitative explorations.

Much like most areas of science, the literature of cognitive science now grows faster than any individual can track. This has resulted in theoretical analysis relying on more philosophically-oriented writing. This could help to exacerbate a growing divide between the *abstract, philosophical characterizations* of the disagreement, and its manifestation in the broader *empirical work*. Moreover, purely qualitative approaches to describing a theoretical landscape could be less convincing than approaches that use quantitative data. Qualitative analyses in the philosophical and theoretical literature are almost always carried out by devotees of one or another perspective. This makes it difficult to remain conceptually impartial, and may neglect the rather more concrete ways that theories play out in the empirical work that the theories are meant to motivate.

The goal of the present paper is to develop a more quantitative approach to this issue, based on the patterns of word use in documents produced by these different theories. Though at this stage our approach and results might be regarded as preliminary, we show that a quantitative analysis of texts can shed light on the relationships among these theoretical frameworks able to act as a supplement to other, more qualitative analyses.

Word use as a window to theory

Our analysis is based around word use—more specifically, their patterns of co-occurrence—on scientific texts produced by authors identifiable with different theories inside of cognitive science. Apart from this being the standard methodology for the kind of semantic model we use (Landauer et al. 1998), it also relates to prominent historical characterizations of theoretical landscapes inside of scientific disciplines. For instance, in *The Historical Meaning of the Crisis in Psychology*, Vygotsky (1997) analyses the then-current state of his discipline. According to his view, psychology was stagnating due to a deep stalemate between competing theories with substantive differences in the conception of their object of study. The most relevant part of his argument for our purposes is that one of the manifestations of this *theoretical* crisis is *changes in word use*. According to Vygotsky, different perspectives employ different words, as words are the "end point and not the starting point of the investigation" (1997, p. 285). Word use can be seen to reflect "the highest principles" (1997, p. 288) that a theory upholds, acting as a "tentacle" (1997, p.286) to grasp how a theory conceptualizes and explains a fact.

A similar view can be seen in Kuhn's later work in the concept of a scientific community's *lexicon* (Kuhn 2000a). With this term, Kuhn refers to a subset of terms whose meaning is interdependent, that together characterize the taxonomy to which a particular approach is committed. When different approaches to the phenomena of a discipline redraw these taxonomic boundaries, they put forward changes to the lexicon of the discipline. Word use is at least to some degree characteristic of the theoretical landscape of a discipline, and crises and revolutions within a discipline can be signaled by the "violation or distortion of a previously unproblematic scientific language" (Kuhn 2000b, p. 32).

Both approaches share the view that word use can reveal underlying features of a discipline; and, moreover, that these underlying features reflect substantive disagreements over key issues. Different theoretical perspectives could be correlated with different patterns in the choice of words the authors that espouse them make. Thus, a way forward for complementing qualitative discussion of theoretical landscapes is a data-driven analysis of the word use of the documents produced by the different theories that constitute the discipline.

As mentioned before, our approach to word use is through a semantic model that uncovers the semantic space underlying a set of documents. Latent semantic analysis (LSA) assumes that, although noisy, the distribution of words used in a document references a space of concepts that can be statistically retrieved from word co-occurrence information (Deerwester et al. 1990). Furthermore, against a backdrop of historical metatheory from Vygotsky and Kuhn, a method like LSA may help map words into a semantic space so that we can determine their theoretical import. Indeed, such a representation of words may then be used to map the theories themselves. This is our goal here.

LSA represents the set documents using a vector space model (Salton et al. 1975) in an asymmetrical matrix in which each paper is characterized by the frequency with which each unique word in the set is used in it. The document by term matrix is then subjected to singular value decomposition (SVD) that statistically groups word contexts and the patterns in which they occur (Landauer et al. 1998). This procedure creates a new set of dimensions that encodes statistical relationships across the whole set. Each dimension of the reduced dimensionality matrix has a different weight in each document, reflecting the concepts that a document expresses. More concretely, in contrast with the previous document by term matrix, documents in the resulting matrix are characterized by their score in each of the dimensions resulting from the SVD, which constitute the semantic space. The most relevant feature of LSA for our purposes is that the new dimensions reflect higher-order properties of the documents than what word frequency does. Therefore, previously veiled similarity or dissimilarity relationships between documents surface, allowing us to infer much more accurately the grouping structure of a given set of documents (Landauer and Dumais 1997).

There are precedents of the use of LSA to model the semantic space of sets of scientific documents. In their seminal work, Deerwester et al. (1990) use LSA to represent two different sets of scientific documents and compare its rate of successful retrieval of these documents from the index with the performance of other methods of co-word analysis. They found that LSA could be successfully used to uncover the structure of sets of scientific documents. More recently, Blatt (2009) shows that a LSA can represent four different disciplines using a purely bottom-up procedure on a set of documents. The clusters that these semantic representations form can then successfully classify and differentiate documents from these disciplines. Paxton (2015) used a semantic model on author-provided keywords to visualize the relationships between different terminology used to refer to inter-personal coordination in abstracts from cognitive science. Finally, Alhazmi et al. (2017) applied LSA to a database of neuroimaging results, showing how the relationship between cognitive functions and brain areas can be modeled as an underlying semantic space.

Our interest lies in exploring the *theoretical* composition of a scientific discipline, in contrast with the issues about which these theories take substantively different stances. While the latter has been the focus of meaning-based approaches to modeling science, the former has been explored primarily through citation analysis (Leydesdorff 1998). This use of citation analysis was delineated early by Garfield et al. (1970) as an aid to studying the history of a discipline, an issue or a theory. Moravcsik and Murugesan (1979) also discuss how citation patterns mirror scientific revolutions, specifically, in their case, in theoretical physics. More recently, Marx and Bornmann (2010, 2013) have used citation analysis as data with which theories of scientific change can be tested. All these suggest that scientometric data can potentially be used to model the theories that exist within a discipline. However, can theoretical features of science be observed in data other than through citation analysis?

Our approach is an attempt to apply a semantic-based tool to the modeling of different theories inside cognitive science. We performed a LSA on a corpus of papers that belong to different theories. Our aim was to show that differences in theoretical stances are significantly related to differences in patterns of word use. For this, first we tested if the patterns of word use of the different theories have significantly different features using simple measures of inner coherence and similarity to other theories. Secondly, we examined the clustering of different theories based on how similar the words used in the documents belonging to them are. Thirdly, we tested the accuracy of a model that predicts the theory of a document based on the patterns of word use exhibited by the documents of each theory. In all of these tests, we found that different theories have significantly different patterns in word use, that the similarities of these patterns matches the intuitive relationships between these theories, and that statistical models of these patterns are specific enough of each theory that they can be fairly predictive of the theory to which a paper pertains. We finalize with a tentative exploration of the terms that characterize each theory based on the statistical models of their location in the semantic space.

Methods

Data collection

Our data consists in abstracts obtained from Thomson Reuters' Web of Knowledge (WOK) using keywords related to these theoretical perspectives within cognitive science.[2] The database was queried for each keyword and a date range from 1975 to 2016, and the results were sorted by citation count. We filtered the results by discipline and document type (see "Appendix 1"). We downloaded the first 100 complete citation information results for each keyword. This resulted in a dataset of 1000 abstracts. Though small, it is a focused sample of the literature with WOK keywords that reflect our theoretical set. It also made it easy to ensure that the papers were relevant to the domains under investigation, namely, the theories of cognitive science.

Next, the abstracts were cleaned in the following way:

- All words were set to lower case.
- The occurrences of the keywords used to obtain them from WOK (Table 1) were removed.
- Possessive, punctuation and copyright marks and notices were removed.
- Using the *nltk* Python package (Bird et al. 2009), all words not recognized as *adjectives*, *nouns* or *verbs* were removed.
- The remaining words were lemmatized using the *nltk* package version of the *WordNet Lemmatizer*.

This procedure yielded a list of 9014 unique terms. These were used to construct a document-by-term matrix (DbT) in which each of the 1000 abstracts were represented as a

[2] All raw data, scripts used for analysis, and analyzed data used for generating the figures are available at http://github.com/contreraskallens/ExploratoryMapping.

Table 1 Keywords used to look for abstracts and their related theory using the "topic" methodology of WoK

Theory	Keywords
Classical computational	*Act-R Computational cognitive model*
Connectionism	*Connectionism "Parallel Distributed Processing"*
Dynamical systems	*Dynamical Systems Cognition*
Ecological psychology	*"Ecological Psychology"*
Embodied cognition	*Embodied Cognition*
Bayesian cognition	*Bayesian Cognition*
Distributed cognition	*"Distributed Cognition"*
Enactivism	*Enactive Cognition*

Quotes were used only when indicated

row stating the frequency with which each term appeared in that document. We removed terms which appeared in only 5 or fewer documents. Furthermore, we used a custom list to remove hand-picked terms corresponding to stop words and stylistic words (e.g. "Furthermore"), and variations of the theory words that the *nltk* package missed. We then removed all empty documents resulting from that procedure. This procedure returned a DbT of 964 $\times$ 1912.

Because of the distortion that document length introduces into some of the measures of similarity that we will use later,[3] we examined the differences in abstract length among the different theories. A one-way ANOVA showed that there are no significant differences in the length of the abstracts by theory, $F(7, 956) = 1.699, p > 0.1$. Furthermore, a regression analysis of the length of documents as predicted by the different theories showed that only *ecological* is significantly different from the intercept (*Bayesian*), $B = -10.719, SE = 3.805, t = -2.817, p < 0.01$. This single difference of *ecological* does not account for the systematic results we report below.

Each column of the DbT was further prepared by weighting each cell using the log-entropy procedure presented by (Martin and Berry 2007, p. 38) term-wise (see "Appendix 2"). Along with the the log normalization, this resulted in an increase of the weight of terms which appear in fewer documents (and are thus more informative), and a decrease in the weight of terms which appear in more documents.

Finally, the resulting matrix was subjected to a SVD procedure using the base *R* function (R Core Team 2017). This results in three different matrices: one that weighted every document by the new dimensions (964 $\times$ 964), one that weighted every term by the new dimensions (1974 $\times$ 964), and one that contains the singular values (964 $\times$ 1). We obtained the *loadings* matrix of both document and terms by multiplying their respective matrices with a diagonal matrix constructed from the singular values. We only saved the first 200 dimensions of the loadings: that is, a 964 $\times$ 200 reduced document loadings matrix, and a 1912 $\times$ 20 reduced unique term loadings matrix.

These matrices of lower dimensionality have been shown to encode semantic relationships by finding regions of shared variance across word co-occurrence patterns. In the following analyses, we utilize relatively small subsets of the $D = 200$ filtered dimensionality from the LSA output. We report this dimensionality D in each case.

[3] We thank an anonymous reviewer for pointing out this limitation.

Results

Self-similarity of theoretical areas

Because we are interested in theoretical conflict, we used the LSA model to test how theories relate in this statistical landscape. This can be tested in two ways. We can explore how internally coherent theories are relative to themselves; we can also explore how externally similar they are to abstracts under other theories.

In a first analysis, we examined how similar articles within a theory are to themselves. For example, are word usage patterns among authors in the *connectionism* abstract set more similar to each other, than, say, *symbolic* abstracts are to their own topic? This would reflect the "internal consistency" of the topics, and we can compare the measure across theories to determine if one or more of them is considerably more (or less) internally consistent.

To do this, we took each abstract from a category, and generated the cosine similarity between its lower-dimensional representation, using dimensions (from now on, D) 1 through 10 ($D = 10$) and all the other abstracts from that category. For each abstract, this produces a mean cosine. Cosine varies from -1 to 1, with 1 reflecting perfect consistency in word usage. By obtaining a cosine of articles within its own group, each domain is assigned a "self-similarity" score based on these comparisons.

After doing this, we used a regression analysis to determine if the theory of the abstracts predicts the variance in these self-similarity scores. Since all abstracts are independent data points, we used a basic ordinary least squares (OLS) regression analysis, with a multinomial factor of abstract category (as listed in Table 1).

Table 2 shows the output from this regression model. This table uses the *Bayesian* category as the reference category because of the alphabetical ordering of the theories. The coefficients represent the self-similarity score of each theory in comparison with *Bayesian*, and whether the difference between their score and the score of *Bayesian* is significant.

The regression model is significant overall, $F(7, 956) = 91.11$, $p < 10^{-10}$. It suggests that the variance in this self-similarity measure is accounted for quite substantially by these 8 categories. The adjusted R^2 of the model is 0.396, suggesting that a rather high amount of variance of word-use consistency is accounted for by theoretical topic.

The average cosine score for *Bayesian* topics is seen as the intercept in Table 2 ($B = 0.672$). The regression model reveals that all topics except for *distributed* significantly exceed or are lower than this baseline. The topics that significantly exceed the

Table 2 Results of self-similarity regression

Topic	B	SE	t
Bayesian (intercept)	0.672	0.008986	74.793***
Connectionism	-0.157	0.0112	-14.117***
Distributed	0.00599	0.0127	0.47
Dynamical	-0.0403	0.0127	-3.172**
Ecological	0.0331	0.0127	2.599**
Embodied	-0.083	0.0126	-6.565***
Enactive	0.0752	0.0128	5.870***
Symbolic	-0.0766	0.011	-6.973***

*$p < 0.01$, **$p < 0.001$, ***$p < 0.0001$

Bayesian cosine value are *ecological* (predicted to be greater by $B = +0.0331$), and *enactive* ($B = +0.0752$). The topics that are significantly lower than *Bayesian* are *connectionism* ($B = -0.157$), *dynamical* ($B = -0.0403$), *embodied* ($B = -0.083$), and *symbolic* ($B = -0.0766$).

By these measures, the most self-similar topic is *enactive*, at a mean cosine of approximately 0.747. The lowest is *connectionism*, with a mean self-similarity cosine of 0.515. In sum, we find considerable variation across theories in how internally consistent they are on this landscape. Relatedly, we find that theoretical identity across the whole set greatly predicts that internal consistency.

Qualitative inspection of theoretical clusters

To test if there are different relationships of similarity between theories, we built a distance matrix using of the relations between theories. For this, we first obtained averages of the cosine of the vector representation of each abstract and the vector representation of the each of the abstracts of the theories being considered (including its own). This gives each abstract a set of eight mean cosines. The abstracts were then grouped by the theory they belong to, and the eight mean cosines were averaged. Thus, each theory is assigned eight different cosines, one for each theory, including itself, representing the similarity of each of its members to each of the members of each theory. This similarity matrix is presented in Fig. 1 for $D = 10$, that is, using dimensions resulting from the SVD 1 through 10. In it, there are clear differences in the similarity between theories: *Bayesian* and *symbolic* are more similar among themselves than they are to other theories, and the same is true for *distributed*, *enactive*, *dynamical* and *ecological*. *Embodied* is also similar to the latter;

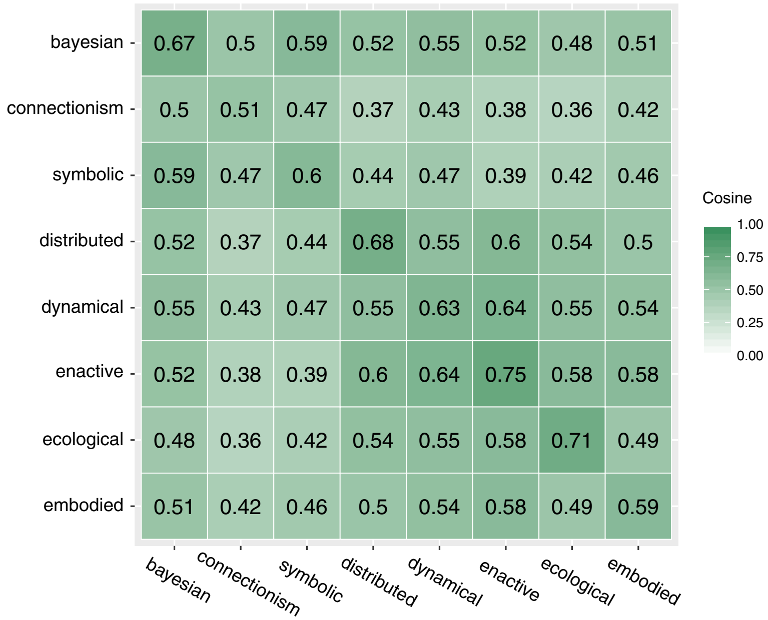

Fig. 1 Heatmap of the similarity matrix between theories for $D = 10$. Color shows the mean cosine between topics. Each cell has the mean cosine of documents of the theory of the row and the one of the column. Theories have been sorted to better visualize clustering. (Color figure online)

however, curiously, *Bayesian* has a slightly higher similarity to it than *ecological* and *distributed*. Although similarity with other theories for *Bayesian* is high for all of them, there are clear clusters on the heatmap: *connectionism* and *symbolic* have higher similarity with each other than they do with *distributed, enactive* and *ecological*. Likewise, they have the lowest mean cosines for *distributed, dynamical, enactive* and *ecological*.

To further explore these relations, the similarity matrix was transformed into a distance matrix, and then subjected to a hierarchical clustering analysis, using the average linking method of the *hclust()* function in *R*. Figure 2 is a dendrogram built with the results of the hierarchical cluster analysis. The original dendrogram was cut into clusters using the *Dynamic Tree Cut R* Package (Langfelder et al. 2008). This pattern is relatively stable from $D = 6$ to at least $D = 200$. The dendrogram shows two main tree branches: on the one side, *connectionism*, *Bayesian* and *symbolic*; on the other side, *ecological*, *distributed*, *enactive*, *dynamical* and *embodied*. Interestingly, the dendrograms for values immediately lower than $D = 6$ have *Bayesian* as part of the opposite cluster.

As implied by our definition of these theoretical approaches above, many cognitive scientists would predict relationships of this kind (for an example, see Chemero and Silberstein 2008). Thus, there seems to be a clear and intuitive clustering of the different theories consistent with the situation of the description of the theoretical landscape of the discipline provided in Sect. 2.1, and distinct from what would result from an unstructured space (see "Appendix 3").

Predicting topics from semantic vectors

Based on the structure found on the results from the hierarchical clustering analysis, we set out to determine if the values in each dimension resulting from the SVD of each abstract could be used to predict the theory of a given abstract. For this, we evaluated each of the first 80 dimensions of the result of the SVD, and ordered them for each topic based on how effective values in each dimension were in correctly predicting that topic using a binary logistic regression of the scores of that theory when pitted against all other theories. The

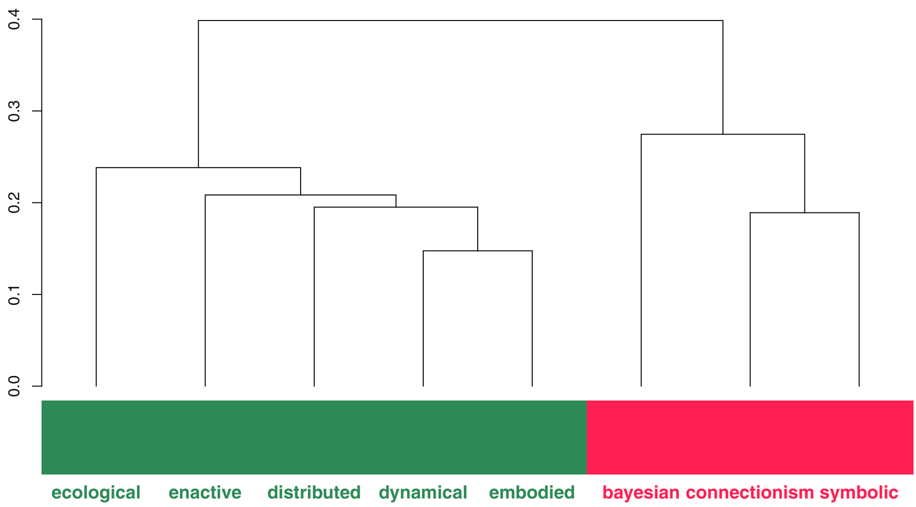

Fig. 2 Dendrogram built using the distance matrix of theories for $D = 10$

choice of the number of dimensions to be evaluated was based in predictive performance (see "Appendix 4"). With this, we could then evaluate not only how effective the lower dimensions obtained from the SVD are in predicting the theory of an abstract, but also how many of them are needed to obtain relevant results.

The procedure for theory prediction works in the following way. For any given value of D, a generalized linear model (GLM) is built for each theory by using the first D dimensions that best predict it, as described above. Each model is trained by using a training set of random abstracts of each theory, constituting 70% of the total abstracts of that theory. After that, each model is presented with the remaining abstracts. We collect the probability of the prediction that each model gives for the abstract, and treat the highest one as the final prediction. We also measured the *performance* of the models, that is, the percentage of times in which the GLM with the highest probability was the same as the theory of the abstract presented to the models. The results are mean aggregations of 10,000 iterations of the process.

First, we swept through values of D from 2 to 50, to see how many dimensions were needed to reliably predict the topic of the presented abstracts. Surprisingly, performance is considerably high even for low values of D. For $D = 5$ (Fig. 3), all theories have a performance considerably higher than 12.5%. The highest values are *connectionism*, with 58.3%, *symbolic* with 56.8$ and *Bayesian* with 54.2%. The mean performance for this value of D is 50.8% Although the value of D with the highest performance is different for every theory, the overall performance seems to stabilize around $D = 10$, with only slight gains or losses with the addition of further dimensions. Peak mean performance is achieved at $D = 20$, with a performance of 64.6%. Figure 4 shows the performance for each theory at this value of D. The highest performing model is *ecological*, with 72.5% correct predictions, closely followed by *distributed* with 70.4% and *enactive* with 69.4%. Every theory has a correct prediction rate of at least 60%, except for the *dynamical* model, whose

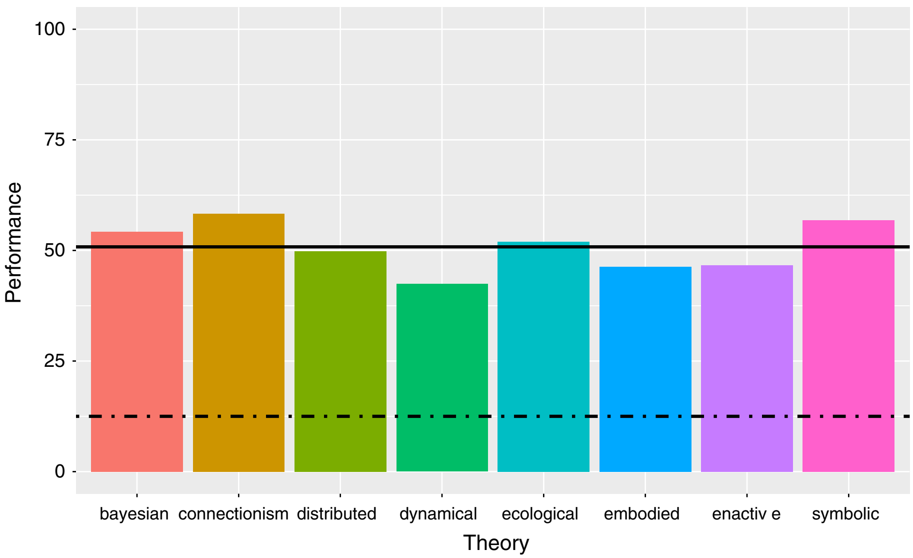

Fig. 3 Mean performance of the GLM of each theory using 5 dimensions for each. Solid line shows the average performance. Dotdashed line shows chance of 12.5% performance. Results aggregate over 10, 000 iterations

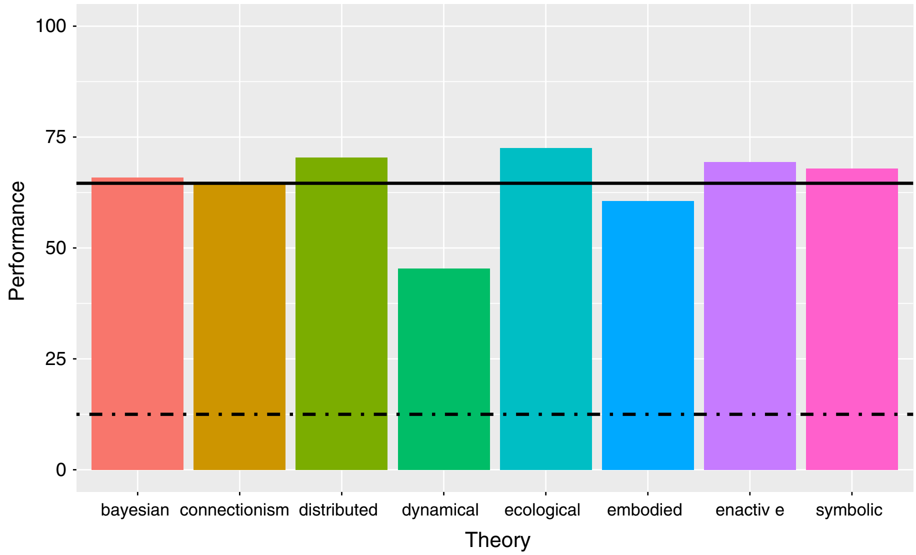

Fig. 4 Mean performance of the GLM of each theory using 20 dimensions for each. Solid line shows the average performance. Dotdashed line shows chance of 12.5% performance. Results aggregate over 10, 000 iterations

performance peaks at 47.6% for $D = 16$. To see the prediction values of all theories across all values of D tested, see “Appendix 5”.

It is also possible to evaluate the level of clustering of the different theories explored. Using both the correct and incorrect answer, a topic confusion matrix can be built for each D, showing what percentage of times a paper of topic A was predicted as being a paper of topic B. Taking into consideration the performance across values of D shown above, Fig. 5 shows the resulting confusion matrix for a value which yielded relatively good results across all topics, $D = 20$. The confusion of prediction shows that the performance of the models reflects the clustering observed in Fig. 2. For *symbolic*, topics within its cluster, *Bayesian* and *connectionism*, account for 18.9% of the remaining predictions, summing a combined 86.8% when added to the percentage of correct predictions. For *enactive*, predictions outside of its observed cluster, *symbolic*, *Bayesian* and *connectionism*, account for only 1.1% of the predictions, and for *ecological* they account for 3.3% of them. Furthermore, a relatively low score like the one for *dynamical*, is offset by the high accuracy of predictions within its cluster: 87.2%.

This latter case is more clearly observable in lower values of D (Fig. 6). At $D = 6$, the dimension in which the dendrogram observed in Fig. 2 becomes the predominant pattern, the clusters are more manifest. *Enactive* stands 44.9% of performance; however, if taken alongside with *dynamical*, *embodied*, and *ecological*, and to a lesser extent *distributed* and *distributed*, the performance increases to around 99.1%. *Bayesian* has a performance of 60.1% and is predicted erroneously as a member of the other cluster only 16.1% of the time. A similar phenomenon can be observed in all of the theories.

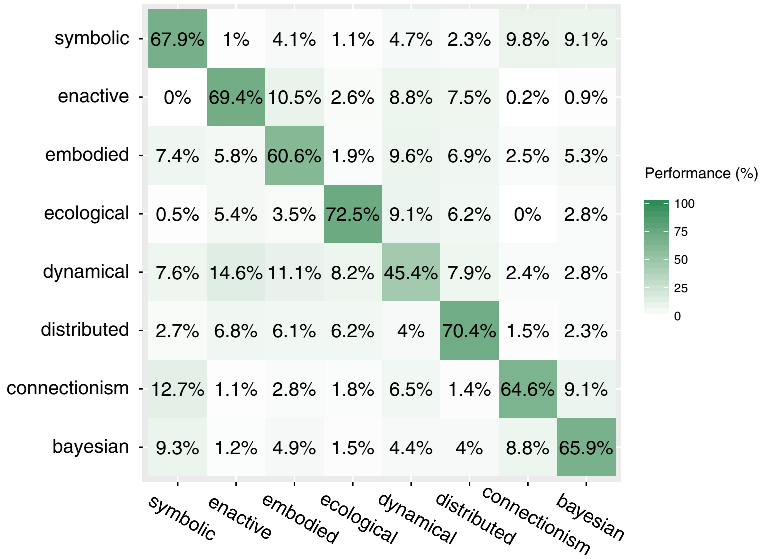

Fig. 5 Confusion matrix for $D = 20$. Results aggregate 10, 000 iterations

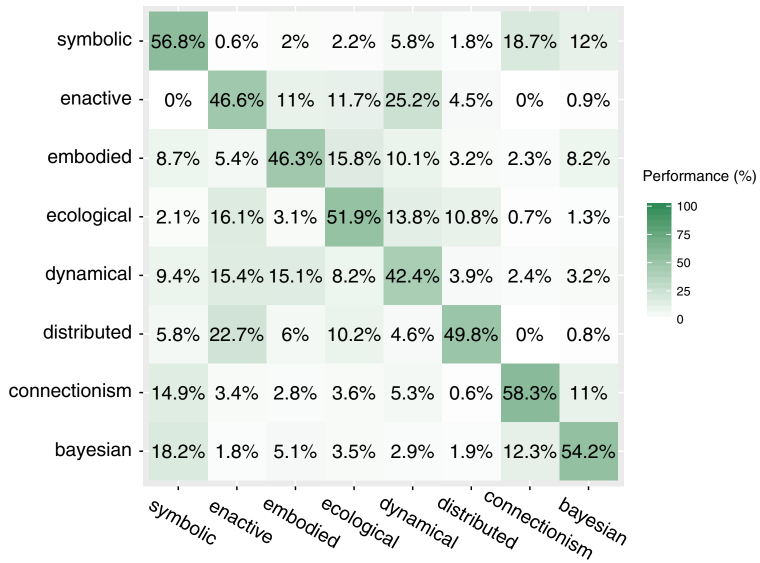

Fig. 6 Confusion matrix for $D = 5$. Results aggregate 10, 000 iterations

Replication

Lastly, we tested our prediction methodology on a new set of papers. We downloaded an additional 100 abstracts (per theory) using the same keywords and search filters. Thus, we obtained a replication dataset of 967 papers with 9243 unique terms. These papers were subjected to the same clean-up procedure as those in the original dataset, resulting in 926 remaining documents with 1926 unique terms. After removing all words from these abstracts that did not appear in the original data set, we projected these new documents onto the original, reduced-dimensionality word-space using the folding-in method Berry et al. (1995). This results in a set of matrices of the scores of each new document on the dimensions obtained by reducing the original document-by-term matrix through SVD. In other words, we used exactly the same LSA model in the prior analyses without retraining on the new data, and tested how well that first model predicts these unseen data.

The prediction results by topic again show a considerably better than chance performance for each topic with relatively small values of D. Figure 7 shows the prediction performance for $D = 5$. The cross prediction, even at this low dimensionality, is considerably higher than chance. The highest performing models are *symbolic* with 52.8% and *distributed* with 51.0%. Even the lowest performing model, *dynamical*, sits at more than 20% higher than chance. The mean performance of the models at this value of D is 45.9%.

Peak mean performance for the cross prediction is lower than using only the original dataset, at 53.3% for $D = 17$. Figure 8 shows the performance for the different theories at this dimensionality. *Ecological* is by far the best performing model, achieving 65.9%. All other models sit close to 50% performance except for *dynamical*, which is again the lowest one at 38.8% (see "Appendix 5")

Similarly, the clustering observed in the original data is expressed in the prediction confusion matrix (Fig. 9). Clusters still account for more than half of predictions in all of the topics in low dimensionality (here, $D = 5$). In the case of *enactive*, out-of-cluster results account for only 1.6% of the predictions. Similarly, for *ecological*, in-cluster results sum 97.7% of the predictions. Finally, for *symbolic*, an individual performance of 52.8% increases to 77.5% when considering it along with *Bayesian* and *connectionism* predictions. Interestingly, *Bayesian* is confused 13.1% of the times with *ecological*, and *connectionism* is confused as much with with *symbolic* as it is with *enactive*.

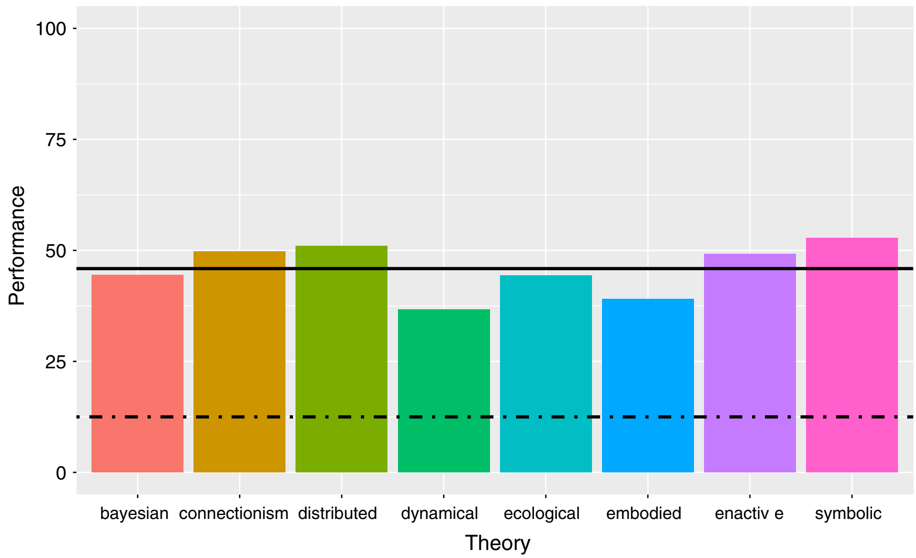

Fig. 7 Mean performance of the GLM of each theory using new data set, using 5 dimensions. Solid line shows the average performance. Dot-dashed line shows chance of 12.5% performance. Results aggregate 10, 000 iterations

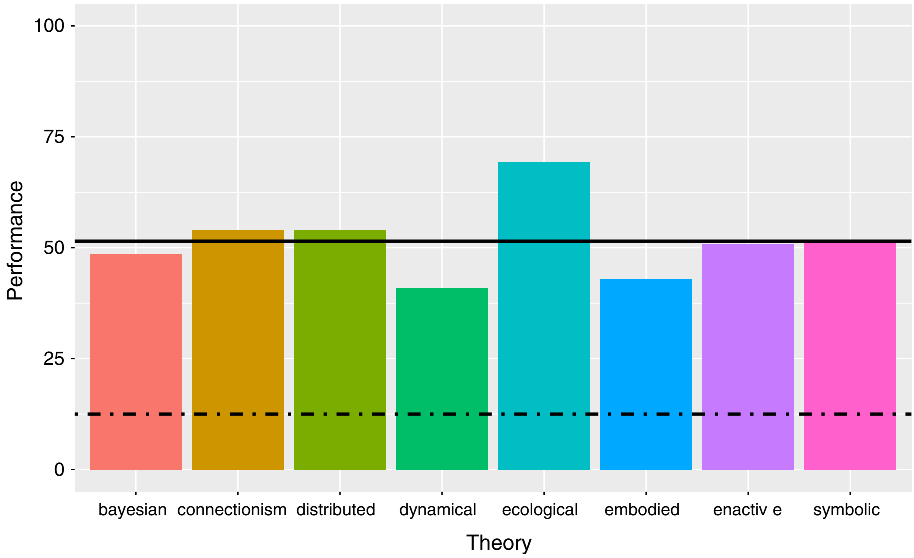

Fig. 8 Mean performance of the GLM of each theory using new data set, using 20 dimensions. Solid line shows the average performance. Dot-dashed line shows chance of 12.5% performance. Results aggregate 10, 000 iterations

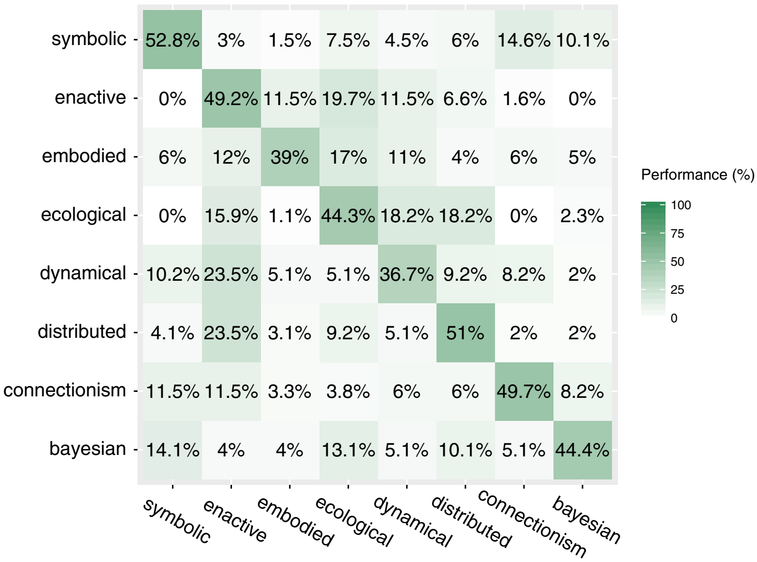

Fig. 9 Confusion matrix for $D = 5$ using replication data

Discussion

These results support the notion that different theories can be successfully represented by using latent semantic analysis (LSA) on word use in abstracts. Firstly, we showed that the theories have different individual properties. There are statistically significant differences between them in how terminologically coherent they are, and how similar they are to other theories.

Secondly, the similarity relations between theories reflect the common characterization of the field of cognitive science as "rooted" in two distinct traditions (Gomila and Calvo 2008; Spivey 2008; Chemero and Silberstein 2008; Wilson 2002; Contreras Kallens 2016; Wheeler 2014; Menary 2010b). These characterizations have emphasized the existence of two approaches to cognitive science: "classical", on the one hand, and "alternative" or "embodied," both of them composed of different subsidiary theories that have varying relationships to one another. Although there are competing characterizations of what exactly is the difference between these two approaches, the former emphasizes cognition as computation over internal representations, and the latter focuses on ways in which the body, the environment, and social interaction, among others, can explain cognitive capacities. The different sub-theories of each have different perspectives on the same kind of approach to explaining cognition. This is especially visible on the "non-classical" approaches, as multiple fairly recent volumes published under the header of "embodied cognition" actually include several of them (Shapiro 2014; Calvo and Gomila 2008; Robbins and Aydede 2009) as do the various journal special issues exploring the topics (Menary 2010a; Almeida e Costa 2005; Ziemke 2002).

Finally, we showed that patterns of word use uncovered by LSA and modeled using a GLM can successfully capture the pattern that is characteristic of each theory. Prediction of which theory each paper espouses with these models is considerably higher than chance both for seen and unseen data, and it can reach considerably high performance.

Examination of the meaning of dimensions

One of the limitations of LSA is that the dimensions of the semantic space that the methodology creates can lack an intuitive interpretation without further processing (Berry et al. 1995; Hu et al. 2007). The meaning of each dimension is highly abstract (Olmos et al. 2014), which makes identifying the meaning of the categories found by this methodology an under-explored area of research (Evangelopoulos 2013). However, a number of procedures that attempt to do just that have been proposed recently.[4] One is based on performing rotations on the matrices resulting from the SVD (Sidorova et al. 2008; Evangelopoulos 2013), and another is based on projecting these matrices onto a new space (Hu et al. 2007; Olmos et al. 2014, 2016). We applied the first one to our model. This decision was based on the relative ease with which we obtained encouraging results.[5]

The methodology we used was presented by Sidorova et al. (2008) and Evangelopoulos, Evangelopoulos et al. (2012). Inspired by Factor analysis, they propose performing a varimax rotation on the term loadings resulting from the SVD, and then projecting the document loadings onto the new rotated space. Both loadings matrices use only the first D dimensions resulting from the SVD. The meaning of each dimension is then extracted by looking at the highest scoring terms on it. In our application of the methodology, we should aim to extract the highest scoring terms of the dimensions that best characterize each theory.

[4] We thank the anonymous reviewers for pointing us towards this research.

[5] We also explored the second methodology, as it holds much promise. However, our study is based on a relatively small set of words, and so choosing a new "word base" for doing this transformation proved to be more difficult, and tended to produce less stable results.

To apply this methodology to our dataset and categories, we need to first choose which dimensions would be associated with each theory. Thanks to the steps involved in our methodology to predict the theory of papers, we had already devised a way to rank dimensions in relation to our eight theories in a given matrix of document loadings. Following our previous method, we ranked each dimension in the new semantic space according to how well it predicted that a document belongs to a theory when compared to each other theory using the GLM models. The coefficient of each of the dimensions selected for the theories was used to determine if the terms are positively or negatively related to a given theory, and how strong that relationship is. For the former, we used the sign of the coefficient: if the coefficient resulting from the GLM was positive, then the terms positively predict that theory, and vice versa. For the latter, we used the value of the coefficient multiplied by the score of that term in that dimension. If the score of the term in that dimension is negative, it was changed to 0. After determining the weight for each dimension and each theory, the scores for each term in each of the dimensions were summed. This was the final weight of the term for that theory.

As an example, in the procedure applied to a varimax rotation of $D = 50$, the term *affordances* appears on dimension 22, which is the second dimension that most positively predicts *ecological*. The coefficient of dimension 22 for *ecological* is 8.516. The score of *affordances* on dimension 22 is 0.145. Thus, the weight for *affordances* on dimension 22 is 1.226. After summing the weights of *affordances* in all 50 dimensions, its final weight is 18.061, the fourth highest weighted term for *ecological*, with a median weight of 1.957 and a mean of 2.469.

For visualizing the resulting terms of the dimensions, we decided to use word clouds generated through the R package *wordcloud* (Fellows 2014). This package shows terms in a size dependent on their frequency; to obtain this, we rounded the final weights of each term to obtain an integer. For proper visualization, the word clouds were limited to 50 dimensions per theory. Furthermore, we generated two word clouds per theory: one including the terms with the most weights in the dimensions that positively predict the theory, and one with those in the dimensions that negatively predict the theory.[6]

Figures 10 and 11 show the word clouds for *ecological* and *Bayesian* respectively, for a varimax rotation of $D = 50$. These figures align very well with intuitions regarding the contents of these theories: *perception*, *affordances*, *Gibson* and *environment* feature prominently in the *ecological* cloud, while *inference*, *rational*, *probabilistic* and *prior* appear in the *Bayesian* one. There are still problematic members of the set, however, like *language* and *strategy* in *ecological*.

Because of how sensitive the procedure is to the choice of parameters, specifically the number of dimensions chosen to be rotated, the intuitiveness of the results for each theory can vary depending on it. Figures 12 and 13 show the word clouds for *embodied* at two different parameter settings. Although $D = 50$ features some important words for the theory like *action*, *perception*, *sensorimotor* and *motor*, $D = 80$ features more centrally a wider array of relevant concepts for the theory in addition to them, like *emotion*, *language*, and *experience*. Meanwhile, Fig. 14 shows the word cloud at $D = 80$ for *ecological*. In it, one of the most central concepts for the theory—the concept of *affordances*—is displaced by less intuitive terms like *social*, *state* and *temporal*.

Finally, the words that most negatively predict the theories can also be intuitive. Figure 15 shows the negative cloud for *symbolic* at $D = 50$, which features some intuitive

[6] Due to space constraints, we can only show a few of these word clouds. However, a more thorough exploration of the parameters can be found in the aforementioned GitHub repository.

Fig. 10 Word cloud for *ecological*. Varimax solution of $D = 50$

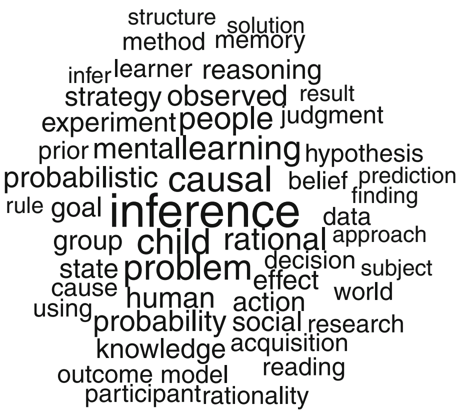

Fig. 11 Word cloud for *Bayesian*. Varimax solution of $D = 50$

Fig. 12 Word cloud for *embodied*. Varimax solution of $D = 50$

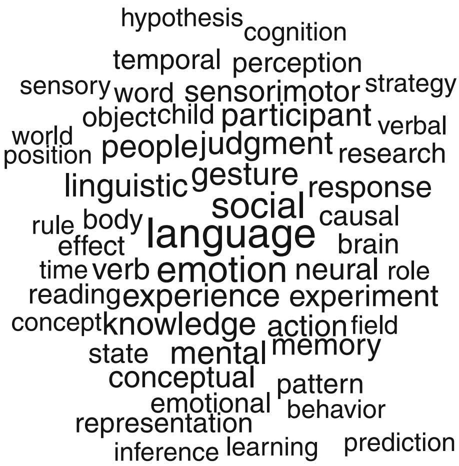

Fig. 13 Word cloud for *embodied*. Varimax solution of $D = 80$

Fig. 14 Word cloud for *ecological*. Varimax solution of $D = 80$

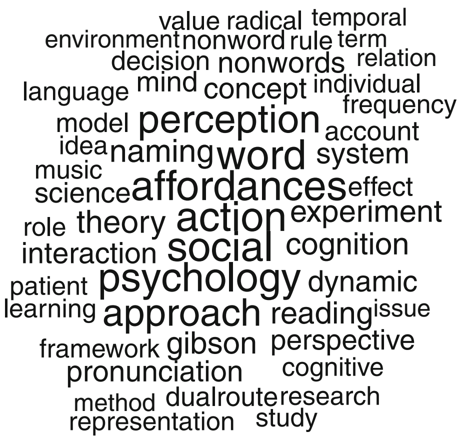

Fig. 15 Negative word cloud for *symbolic*. Varimax solution of $D = 50$

terms such as *affordances*, *action*, *social*, *dynamic* and *interaction*. However, there are also seemingly unrelated terms, considering the history of this approach to cognitive science, such as *cognition*, *word*, *experiment* and *representation*.

It is therefore possible to extract intuitive and meaningful results from the semantic model we have used of each of the theories. However, the methodology we used is not free from important considerations. In particular, it is very sensitive to parameters and choices. Thus, the results in this section serve merely as tentative and qualitative explorations. Future analysis may leverage these and other methods to uncover robust word-theory relationships. We have shown a few of these here, suggesting the goal has some promise.

General discussion

In the preceding sections, we aimed to test if a quantitative method could be used to capture the theoretical make up of a field. We have found that LSA on a sample of text can indeed capture interesting properties of this theoretical landscape: Different theories of a field could be related to different patterns of word use. Interestingly, the insights shared long ago by Vygotsky and Kuhn about the role of language and theory development may ring true under a modern quantitative lens. In particular, the LSA model reveals semantic dimensions that characterize these fields differently. This could be an explanation of the perceived theoretical stalemate of the discipline, as semantic variance has traditionally been seen as one of the possible sources of theory incommensurability (Sankey 1997).

Consider, for example, the results on self-similarity. The most self-similar frameworks in our dataset are *enactive* and *ecological*, suggesting that these may be very tightly knit abstracts terminologically (and perhaps, though with great caution, we might say *conceptually*). This may reflect a kind of conceptual boundary that demarcates these two theories, insulating them from cross-talk with others, such as more traditional information-processing accounts. Take another example from our qualitative investigation of how these theories cluster (e.g., Fig. 5). There, we find that *dynamical* is very confusable with some of the members of its cluster. *Enactive* and *embodied* are confused over 15% of the time with it. A similar point can be made about *Bayesian* approaches based on the dendrograms. Although it clusters with *connectionism* and *symbolic*, it is the closest one to the other cluster; moreover, as can be seen in "Appendix 3", it is the only theory that changes cluster in lower dimensionality. This, and the other cross-cluster confusions, suggest a possible way of bridging gaps in the future. By finding that there is terminological proximity between theories A and B, and between B and C, we may find that working towards rapprochement among all three is through theory B. In our results, this theory B would be *Bayesian*, as it may have ingredients inviting connections among this cluster. Indeed, Bayesian models are often computations over individuated "representations" (hypotheses), a process that is probabilistic (akin to more dynamic, gradient theories), all performed through experiential input. Its status as a bridge may be surprising to some, but our analysis suggests it.

These efforts could be aided by a more robust exploration of the meanings in the semantic space. We uncovered some interesting and intuitive patterns that suggest that the bridges between theories could be looked for in the specific terminological patterns of each theory. Moreover, this could also reveal the underlying structure of each cluster, and aid in the current efforts of characterizing cognitive science and its sharp theoretical divide. A peripheral benefit of this characterization might be practical in nature. Classic application of LSA and related semantic space models is on document retrieval problems (Dumais

1991). Document indexing systems could be equipped with "theoretically normalized" versions of the indexing data. Such normalization may be based on human judges, or by automated techniques based on generalizations of what we show here. This might allow researchers to connect related ideas or threads of research that share conceptual foundations, facilitating new connections among researchers and literatures.

Our results involve a variety of such observations. But it is merely speculative that this terminological coherence and connectivity would reflect the psychology of scientific practitioners. This cause or correlation—between LSA results and theoretical and scientific practice—would be difficult to test, but is certainly an important question in regards to the external validity of these LSA analyses. However, although exploratory, our efforts in quantitative exploration of the features of different theories and their prediction based on those features shows that there is a conceptual landscape in which theoretical issues could be situated. Moreover, even though they are still tentative, our exploration of the meaning of those dimensions shows promise that this landscape not only exists, but could be precisely mapped in the future.

Our study has a number of limitations that narrow its scope and power. These are important to recognize, as they may motivate future applications of these quantitative techniques to theoretical questions.

Firstly, our data collection method is dependent on queries that can seem arbitrary. This is in addition to the fact that our analysis greatly depends on the predefined taxonomy that we decided to use, which the search queries reflect. Moreover, the scope of WOK also limits our results to a constrained database of papers, limiting how representative our results can be of the whole discipline. Nevertheless, our sample consists of 1000 papers sampled from these key terms. There is considerable room to expand these papers, but this may also introduce noise—our selective sample allowed us to have some better confidence that the papers reflected activity in cognitive science, and not inadvertently some other distant field in which the terms may also be used (e.g., the term "symbolic" also occurs in a wide variety of fields in the humanities). However, we think that our methodology is flexible enough to support subsequent improvements of the categories being tested based on both internal and external measures of performance.

The data retrieval process can also be expanded and improved upon. We intend to explore different, more rigorous methods of gathering papers to complement the one applied in the present study. Alternative methods could point to gathering fundamental papers from co-citation analysis of the discipline, or off-loading the decision to members of the community through a survey. With this, our approach could also incorporate into the semantic analysis the influence of social aspects of the organization of scientific communities that remained in the background in this study. It could also open a window to use these tools in a discipline with which we are not previously acquainted, reducing even more the chances of "seeding" the theoretical clusters.

Secondly, and related to the first limitation, the work presented here uses only measures of performance internal to our corpus—in this case, how well the words of the abstracts predicted the labels with which we obtained them. Although the results show that there are indeed different categories within the corpus, corresponding approximately with our assigned labels, the high performance could be reflecting categories whose interpretations differ from our expectation. Thus, a future improvement upon this work would include external human judges with which to compare the performance of the model.[7]

[7] We thank an anonymous reviewer for this suggestion.

Thirdly, a purely semantic way for conceiving of theories downplays the social aspects that have traditionally been used to study theoretical landscapes, such as citation or social network analysis (Bergmann and Dale 2016). Our database is limited exclusively to abstracts, which, apart from the obvious differences of length, have a different communicative goal than full articles.

Finally, and most relevantly, our application of methodologies for extracting the meaning out of the semantic model remains incomplete. Because of the heavy parametrization needed to achieve better results, this fell outside of the scope of our current analysis. Fine-tuning the methodology we used, and exploring the results of other promising approaches, is an obvious extension of the current work.

We could also look to topic modeling to complement these analysis (Griffiths and Steyvers 2004). Previously, in contrast with our assumptions about the underlying semantic structure of the discipline, cognitive science has been represented as being comprised of a set of topics, which consist of probability distributions of words. Each document, then, can be seen as having been probabilistically generated following the topics it contains. Using these assumptions, Bergmann and Dale (2016) explored the topic makeup of a specific issue within cognitive science, the evolution of language. They found three clearly defined clusters of 20 topics in the submissions to the *EvoLang* conference. Using a similar strategy, Priva and Austerweil (2015) modeled the topic makeup of the articles published in the journal *Cognition* over its existence. They found that topic models can successfully track significant changes over time of the predominant issues and problems in cognitive science, such as the rise of moral cognition, through changes in the proportion of the contribution of each topic each year.

Time dynamics could be incorporated into the analysis of the discipline. In line with some of the work by Gentner (2010), one could explore the proportion of documents coming from each of the disciplines that make up cognitive science's core, using the journal *Cognitive Science* and the *Cognitive Science* conference over time. Gentner found a progressive hegemonization of the discipline by psychology, with shrinking contributions from, for example, computer science. Our tools could be used, then, to explore the changing relations of the different theories that we identified in the preceding work. There is a recent precedent for the use of LSA in exploring the time dynamics of word: *Word Maturity* (Kireyev and Landauer 2011; Jorge-Botana et al. 2018). This methodology models the change of meaning of a complete word space by comparing a reference representation of a corpus with intermediate stages, measuring the maturity a given word at a given time step. Although it has been used mainly for studying language development in infants by using an adult corpus as reference, the methodology could be applied to a scientific discipline by using a recent year as reference to measure the trajectories that the meanings of specific key words have followed.

Despite these limitations, our study does show that theoretical outlook has a correlate with observable and quantifiable aspects of the papers of a discipline. This does not, of course, offer a clear way forward for integration or resolution in the debates of cognitive science. It does, however, provide an initial demonstration that this problem can be supplemented with quantitative analysis. It may be possible to build a quantitative map of the landscape of problems and solutions in contention within the discipline. Such a quantitative map may serve as a useful tool for confirming profound theoretical disagreement, and perhaps even finding unexpected rapprochement.

Acknowledgements We want to thank professors Paul Smaldino and Jeff Yoshimi for their feedback on this paper. Thanks to Martin Irani for his help with coding and feedback on the preliminary results.

Appendix 1: Search filters

For the search procedure, quotes were used on the keywords that generated the most confusion on the results due to how common the words used are. Only references classified as *articles*, *proceedings paper*, *reviews*, *book chapters* and *editorial material* were downloaded.

The following subdiscipline categories provided by WoS were used (in alphabetical order):

Behavioral Sciences; Computer Science—Artificial Intelligence; Computer Science—Cybernetics; Computer Science—Information Systems; Computer Science—Interdisciplinary Applications Computer Science—Theory Methods; History Philosophy of Science; Language—Linguistics; Linguistics; Neurosciences; Philosophy; Psychology; Psychology—Applied; Psychology—Biological; Psychology—Developmental; Psychology—Educational; Psychology—Experimental; Psychology—Mathematical; Psychology—Multidisciplinary; Psychology—Social; Robotics; Social Sciences—Interdisciplinary

Appendix 2: Entropy calculation

The formula with which each cell was weighted in our analysis, from (Martin and Berry 2007, p. 38). Each cell was weighted locally with the logarithm of frequency plus 1, $\log\left(f_{ij}+1\right)$. Then, that value was multiplied by the entropy of each one of the terms:

$$1+\sum_{j}\frac{P_{ij}\times\log_{2}P_{ij}}{\log_{2}n}$$

where P_{ij} is the number of times the term i appears in document j, divided by the number of times the term appears in all of the documents, and n is the total number of documents in the dataset. This formula assumes that terms that appear in fewer documents are more informative than terms that appear in more documents. Thus, the values of the former are relatively increased, while the values of the latter are relatively diminished.

Appendix 3: Other dendrograms

In Figs. 16 and 17, we present the dendrograms that obtain by using the values lower ($D=3$) and ($D=5$) than the range of the stable dendrogram presented in Fig. 2. Figure 18 shows the same D as the one used to produce Fig. 2 ($D=10$), but including a randomization of the theory labels attached to each paper (Figs. 19, 20, 21, 22, 23).

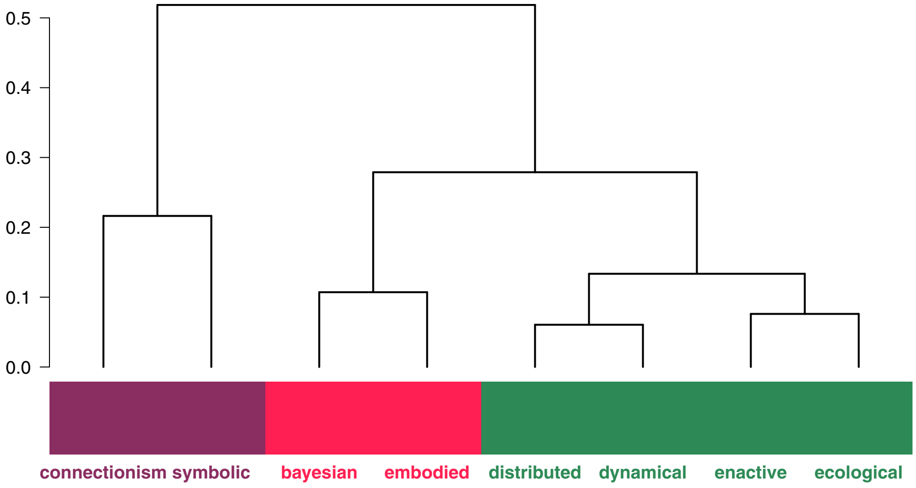

Fig. 16 $D = 3$

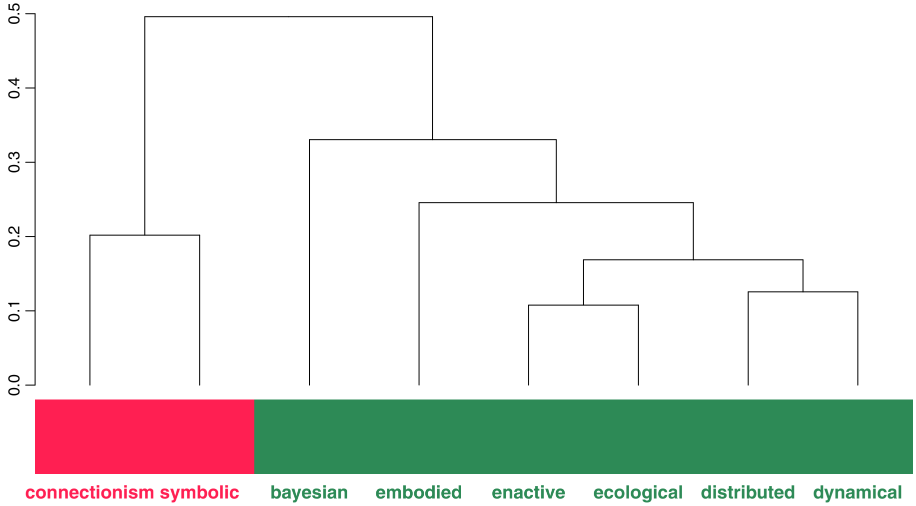

Fig. 17 $D = 5$

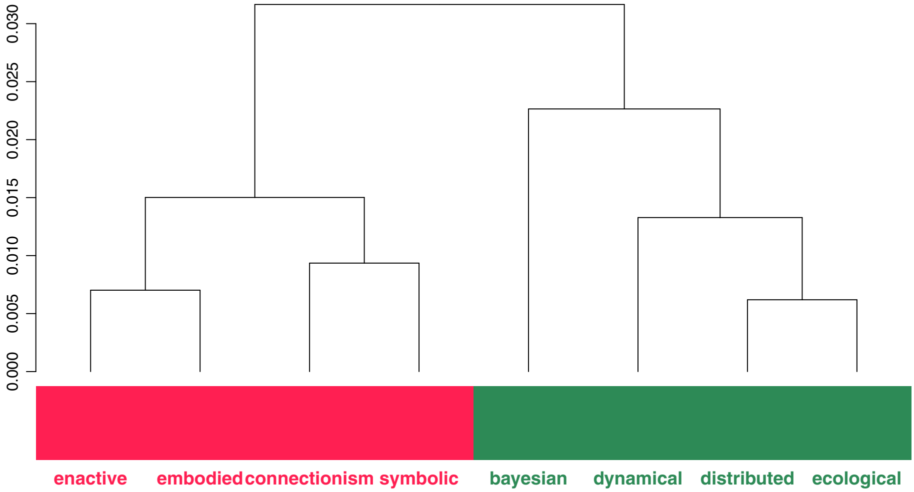

Fig. 18 $D = 10$. Randomized theories. Note the low distance values between the tree branches in comparison to Fig. 2, on the y axis

Appendix 4: Selection of number of dimensions to evaluate

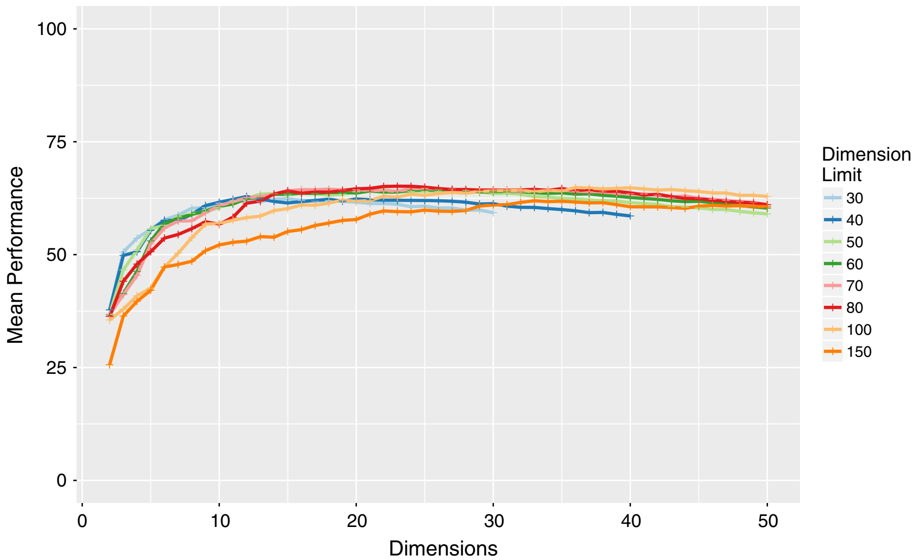

Fig. 19 Mean performance of the prediction by changing the number of dimensions allowed to be evaluated when selecting the best predictors for each theory. Peak performance is achieved by limiting it to 80 dimensions (red line). However, performance is robust, so this parameter can be changed without much decrease in performance. Results aggregate over 1000 iterations. (Color figure online)

Appendix 5: Prediction performance across values of *D*

Fig. 20 Mean performance of the GLM of each theory. D is the number of dimensions used for the model. Results aggregate over 10, 000 iterations

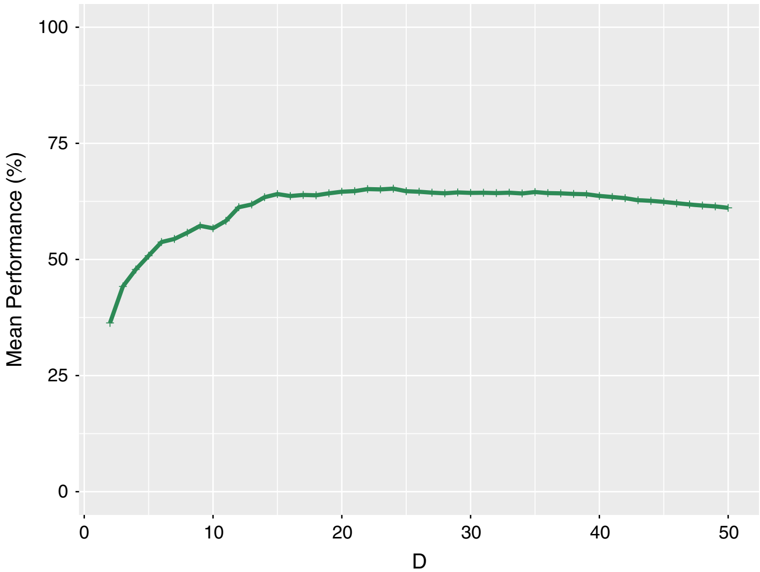

Fig. 21 Mean performance of the 8 models across values of *D*. Aggregated over 10, 000 iterations

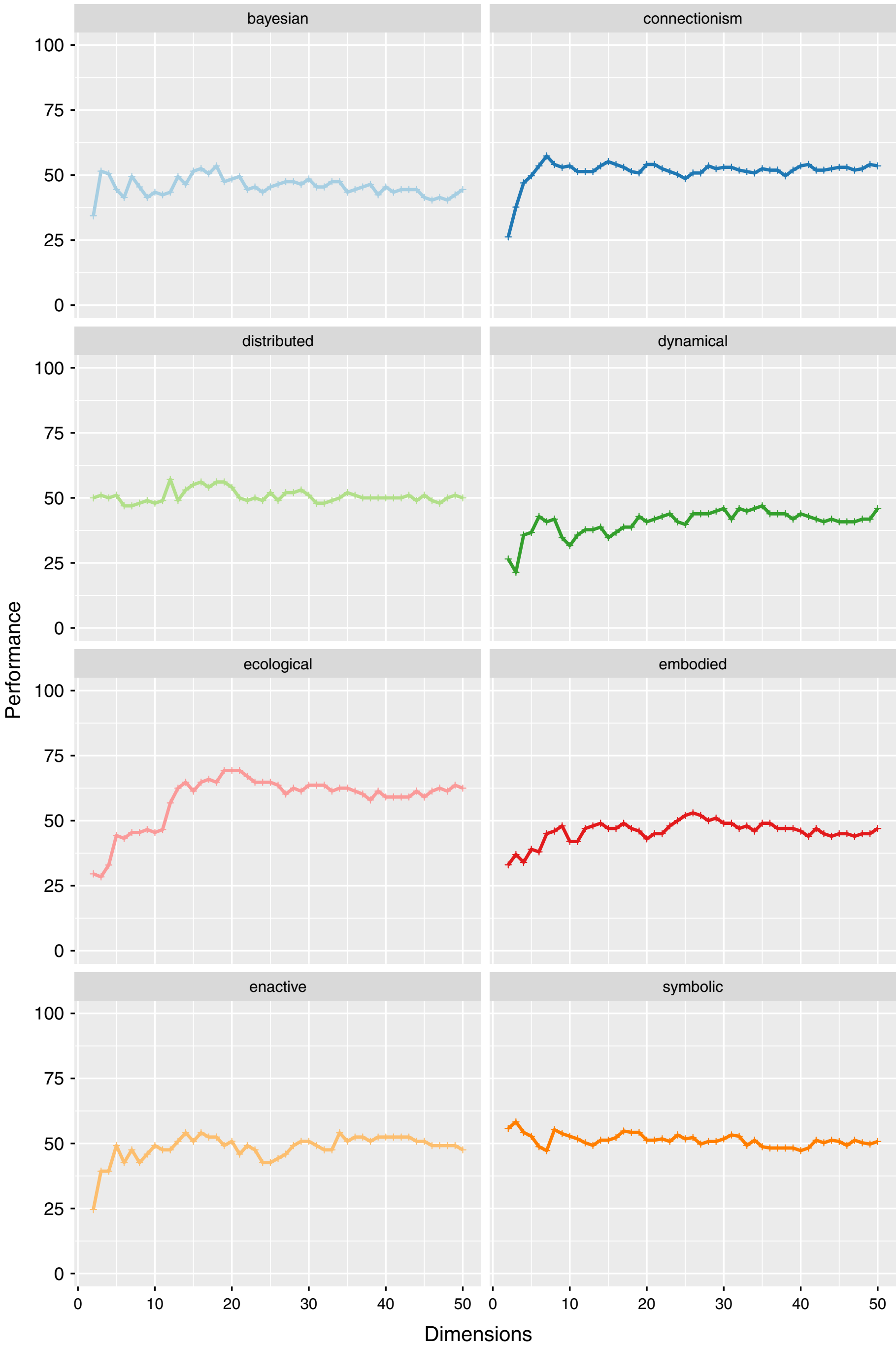

Fig. 22 Mean performance of the GLM of each theory using new data set. *D* is the number of dimensions used for the model. Results aggregate 10, 000 iterations

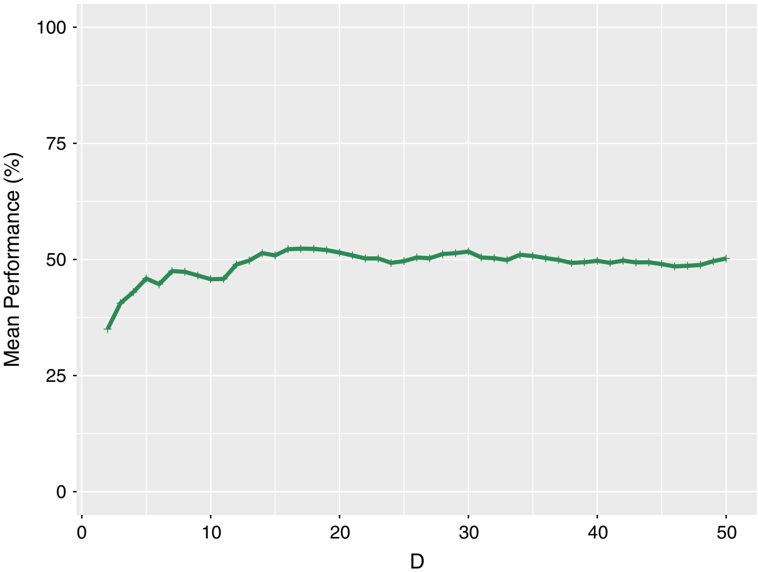

Fig. 23 Mean performance of the 8 models across values of D using new data set. Results aggregate 10, 000 iterations

Appendix 6: Prediction performance in randomization condition

Figure 24 shows the prediction confusion matrix for $D = 20$ the eight different theories with randomization of labels. The maximum value shown is 16.9% of confusion (*embodied* - *ecological*) and the minimum value is 8% (*distributed* - *enactive*). Figure 25 shows the mean performance for the prediction (y-axis) for each theory (x-axis) when theory labels are randomized.

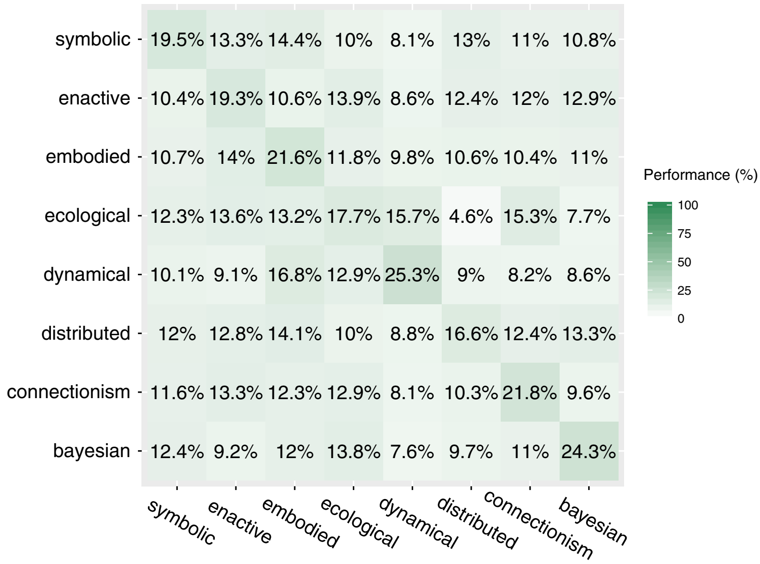

Fig. 24 Confusion matrix with randomized theories, $D = 20$

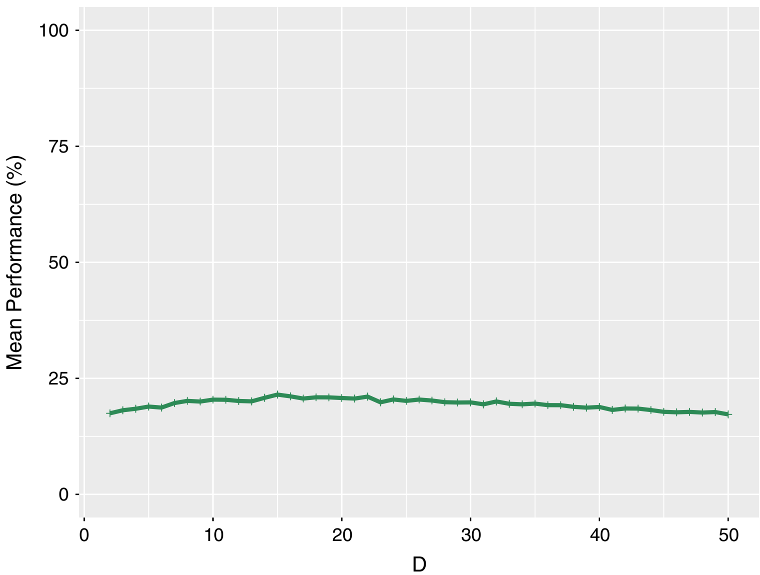

Fig. 25 Mean performance by *D*. Randomized

References

Adams, F., & Aizawa, K. (2010). Defending the bounds of cognition. In R. Menary (Ed.), *The extended mind* (pp. 67–80). Cambridge, MA: The MIT Press.

Alhazmi, F., Beaton, D., & Abdi, H. (2017). The latent semantic space and corresponding brain regions of the functional neuroimaging literature. *bioRxiv*. https://doi.org/10.1101/157826.

Almeida e Costa, F. (Ed.). (2005). Embodied and situated cognition [Special Issue]. *Artificial Life, 11*(2).

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 91–130.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*(4), 637–660.

Bechtel, W., & Abrahamsen, A. (2006). *Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective. Contemporary debates in cognitive science*. Oxford, UK: Basil Blackwell.

Bechtel, W., & Graham, G. (1998). *A companion to cognitive science*. Oxford, UK: Blackwell.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, *72*(1–2), 173–215.

Bergmann, T., & Dale, R. (2016). A scientometric analysis of evolang: Intersections and authorships. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (EVOLANGX11)*. http://evolang.org/neworleans/papers/182.html. Retrieved 22 June 2018.

Berry, M. W., Dumais, S. T., & O'Brien, G. W. (1995). Using linear algebra for intelligent information retrieval. *SIAM Review*, *37*(4), 573–595.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media Inc.

Blatt, E. (2009). Differentiating, describing, and visualizing scientific space: A novel approach to the analysis of published scientific abstracts. *Scientometrics*, *80*(2), 385–406.

Calvo, P., & Gomila, T. (2008). *Handbook of cognitive science: An embodied approach*. New York, NY: Elsevier.

Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Chemero, A., & Silberstein, M. (2008). After the philosophy of mind: Replacing scholasticism with science. *Philosophy of Science*, *75*(1), 1–27.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, MA: The MIT Press.

Contreras Kallens, P. A. (2016). *La máquina del fantasma: unidad de análisis en la ciencia cognitiva* (Unpublished master's thesis). Universidad de Chile, Av. Capitán Ignacio Carrera Pinto 1025, Nuñoa, Santiago, Chile.

Cowley, S. J. (2011). *Distributed language* (Vol. 34). Amsterdam, Netherlands: John Benjamins Publishing.

Dale, R. (2008). The possibility of a pluralist cognitive science. *Journal of Experimental and Theoretical Artificial Intelligence*, *20*(3), 155–179.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391.

de Oliveira, G. S., & Chemero, A. (2015). Against smallism and localism. *Studies in Logic, Grammar and Rhetoric*, *41*(1), 9–23.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*(4), 429–452.

Dumais, S. T. (1991). Improving the retrieval of information from external sources. *Behavior Research Methods, Instruments, & Computers*, *23*(2), 229–236.

Edelman, S. (2008). On the nature of minds, or: Truth and consequences. *Journal of Experimental & Theoretical Artificial Intelligence*, *20*(3), 181–196.

Evangelopoulos, N., Zhang, X., & Prybutok, V. R. (2012). Latent semantic analysis: Five methodological recommendations. *European Journal of Information Systems*, *21*(1), 70–86.

Evangelopoulos, N. E. (2013). Latent semantic analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(6), 683–692.

Feldman, J. A., & Ballard, D. H. (1982). Connectionist models and their properties. *Cognitive Science*, *6*(3), 205–254.

Fellows, I. (2014). wordcloud: Word clouds. Retrieved from https://CRAN.R-project.org/package=wordcloud (R package version 2.5).

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Garfield, E., et al. (1970). Citation indexing for studying science. *Nature*, *227*(5259), 669–671.

Gentner, D. (2010). Psychology in cognitive science: 1978–2038. *Topics in Cognitive Science*, *2*(3), 328–344.

Gibbs, R. W, Jr. (2005). *Embodiment and Cognitive Science*. Cambridge, UK: Cambridge University Press.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. Oxford, UK: Oxford University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception: Classic edition (2014)*. Hove, UK: Psychology Press.
Gomila, T., & Calvo, P. (2008). Directions for an embodied cognitive science: Toward an integrated approach. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 1–25). San Diego, CA: Elsevier.
Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364.
Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, *101*(suppl 1), 5228–5235.
Haugeland, J. (1978). The nature and plausibility of cognitivism. *Behavioral and Brain Sciences*, *1*(2), 215–226.
Haugeland, J. (1981). Semantic engines: An introduction to mind design. In J. Haugeland (Ed.), *Mind design* (pp. 34–50). Cambridge, MA: MIT Press.
Hohwy, J. (2013). *The predictive mind*. Oxford, UK: Oxford University Press.
Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human–computer interaction research. *ACM Transactions on Computer–Human Interaction (TOCHI)*, *7*(2), 174–196.
Hu, X., Cai, Z., Wiemer-Hastings, P., Graesser, A. C., & McNamara, D. S. (2007). Strengths, limitations, and extensions of LSA. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 401–426). New York, NY: Routledge.
Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.
Jorge-Botana, G., Olmos, R., & Luzón, J. M. (2018). Word maturity indices with latent semantic analysis: why, when, and where is Procrustes rotation applied? *WIREs Cogn Sci*, *9*, e1457.
Kireyev, K., & Landauer, T. K. (2011). Word maturity: Computational modeling of word knowledge. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies - Volume 1* (pp. 299–308). Stroudsburg, PA: Association for Computational Linguistics.
Kugler, P. N., & Turvey, M. T. (2015). *Information, natural law, and the self-assembly of rhythmic movement*. Abingdon, UK: Routledge.
Kuhn, T. S. (2000a). Commensurability, comparability, communicability. In J. Conant & J. Haugeland (Eds.), *The road since Structure: Philosophical essays 1970–1993, with an autobiographical interview* (pp. 33–57). Chicago IL: The University of Chicago Press.
Kuhn, T. S. (2000b). What are scientific revolutions? In J. Conant & J. Haugeland (Eds.), *The road since Structure: Philosophical essays 1970–1993, with an autobiographical interview* (pp. 13–32). Chicago, IL: The University of Chicago Press.
Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211.
Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284.
Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: The dynamic tree cut package for R. *Bioinformatics*, *24*(5), 719–720.
Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, *43*(1), 5–25.
Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*(2), 273–302.
Martin, D. I., & Berry, M. W. (2007). Mathematical foundations behind latent semantic analysis. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 35–56). New York, NY: Routledge.
Marx, W., & Bornmann, L. (2010). How accurately does thomas kuhn's model of paradigm change describe the transition from the static view of the universe to the big bang theory in cosmology? *Scientometrics*, *84*(2), 441–464.
Marx, W., & Bornmann, L. (2013). The emergence of plate tectonics and the kuhnian model of paradigm shift: A bibliometric case study based on the anna karenina principle. *Scientometrics*, *94*(2), 595–614.
McCauley, R. N., & Bechtel, W. (2001). Explanatory pluralism and heuristic identity theory. *Theory & Psychology*, *11*(6), 736–760.
McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., et al. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*(8), 348–356.
Menary, R. (Ed.) (2010a). 4e cognition: Embodied, embedded, enacted, extended [Special Issue]. *Phenomenology and the Cognitive Sciences*, *9*(4), 459–463.

Menary, R. (2010b). Introduction to the special issue on 4e cognition. *Phenomenology and the Cognitive Sciences*, *9*(4), 459–463.
Michaels, C. F., & Carello, C. (1981). *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall.
Moravcsik, M. J., & Murugesan, P. (1979). Citation patterns in scientific revolutions. *Scientometrics*, *1*(2), 161–169.
Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In *Visual Information Processing* (pp. 283–308).
Oaksford, M., & Chater, N. (2009). Précis of bayesian rationality: The probabilistic approach to human reasoning. *Behavioral and Brain Sciences*, *32*(1), 69–84.
Olmos, R., Jorge-Botana, G., León, J. A., & Escudero, I. (2014). Transforming selected concepts into dimensions in latent semantic analysis. *Discourse Processes*, *51*(5–6), 494–510.
Olmos, R., Jorge-Botana, G., Luzón, J. M., Martín-Cordero, J. I., & León, J. A. (2016). Transforming lsa space dimensions into a rubric for an automatic assessment and feedback system. *Information Processing & Management*, *52*(3), 359–373.
Paxton, A. E. (2015). *Coordination: Theoretical, methodological, and experimental perspectives*. Ph.D. thesis, University of California, Merced.
Priva, U. C., & Austerweil, J. L. (2015). Analyzing the history of cognition using topic models. *Cognition*, *135*, 4–9.
R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Richardson, M. J., Shockley, K., Fajen, B. R., Riley, M. A., & Turvey, M. T. (2008). Ecological psychology: Six principles for an embodied–embedded approach to behavior. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 159–187). San Diego, CA: Elsevier.
Robbins, P., Aydede, M., et al. (2009). *The Cambridge handbook of situated cognition*. Cambridge, UK: Cambridge University Press.
Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. Cambridge, MA: The MIT Press.
Rumelhart, D. E., McClelland, J. L., Group, P. R., et al. (1987). *Parallel distributed processing* (Vol. 1). Cambridge, MA: The MIT Press.
Salton, G., Wong, A., & Yang, C.-S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, *18*(11), 613–620.
Sankey, H. (1997). Incommensurability: The current state of play. *Theoria: An International Journal for Theory, History and Foundations of Science, 12*(3), 425–445.
Shapiro, L. (2010). *Embodied cognition*. Abingdon, UK: Routledge.
Shapiro, L. (2014). *The Routledge handbook of embodied cognition*. Abingdon, UK: Routledge.
Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Ramakrishnan, T. (2008). Uncovering the intellectual core of the information systems discipline. *MIS Quarterly*, *32*(3), 467–482.
Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, *11*(1), 1–23.
Spivey, M. (2008). *The continuity of mind*. Oxford, UK: Oxford University Press.
Stepp, N., Chemero, A., & Turvey, M. T. (2011). Philosophy for the rest of cognitive science. *Topics in Cognitive Science*, *3*(2), 425–437.
Stewart, J., Stewart, J. R., Gapenne, O., & Di Paolo, E. A. (2010). *Enaction: Toward a new paradigm for cognitive science*. Cambridge, MA: The MIT Press.
Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318.
Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279–1285.
Thelen, E., & Smith, L. B. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press.
Van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy*, *92*(7), 345–381.
Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: The MIT Press.
Von Eckardt, B. (1995). *What is cognitive science?* Cambridge, MA: The MIT Press.
Vygotsky, L. S. (1997). *The collected works of LS Vygotsky: Problems of the theory and history of psychology* (Vol. 3). Berlin, Germany: Springer.
Wheeler, M. (2014). Revolution, reform, or business as usual? The future prospects for embodied cognition. In L. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 374–383). Abingdon, UK: Routledge.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.
Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences-Cognition*. Cambridge, UK: Cambridge University Press.
Yoshimi, J. (2012). Active internalism and open dynamical systems. *Philosophical Psychology*, *25*(1), 1–24.
Ziemke, T. (Ed.). (2002). Situated and embodied cognition [Special Issue]. *Cognitive systems research 3*(3).
Ziemke, T. (2003). What's that thing called embodiment. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual meeting of the cognitive science society* (pp. 1305–1310). Mahwah, NJ: Lawrence Erlbaum.
Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, *18*(5), 229–234.