Acta Psychologica 138 (2011) 11–18

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Representation at different levels in a conceptual hierarchy ☆

Wouter Voorspoels *, Gert Storms, Wolf Vanpaemel

Department of Psychology, K.U. Leuven, Belgium

ARTICLE INFO

Article history:
Received 28 December 2010
Received in revised form 5 April 2011
Accepted 19 April 2011
Available online 28 May 2011

PsycINFO classification:
2300
2340
2343

Keywords:
Categorization
Concept representation
Typicality
Computational models

ABSTRACT

The present study examines the influence of hierarchical level on category representation. Three computational models of representation – an exemplar model, a prototype model and an ideal representation model – were evaluated in their ability to account for the typicality gradient of categories at two hierarchical levels in the conceptual domain of clothes. The domain contains 20 subordinate categories (e.g., trousers, stockings and underwear) and an encompassing superordinate category (CLOTHES). The models were evaluated both in terms of their ability to fit the empirical data and their generalizability through marginal likelihood. The hierarchical level was found to clearly influence the type of representation: For concepts at the subordinate level, exemplar representations were supported. At the superordinate level, however, an ideal representation was overwhelmingly preferred over exemplar and prototype representations. This finding contributes to the increasingly dominant view that the human conceptual apparatus adopts both exemplar representations and more abstract representations, contradicting unitary approaches to categorization.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Categorization is a fundamental human cognitive function. Grouping the environment in different categories opens up a range of possibilities, such as labeling objects, communicating about and interacting with them, inferring properties concerning them and so on. A short reflection on the difficulty of treating every experience as unique and new, not having access to the vast amounts of information and knowledge that are available in concepts, reveals that categorizing is basic to most of our cognitive activity and, by extension, our daily life. Not surprisingly, it has received a great deal of attention in cognitive science, resulting in elaborate research in both artificial categorization contexts and natural language concepts, and the development of different theories on how categories are mentally represented.

At the core of the debate between different views and theories concerning the nature of category representation lies the issue of abstraction. On the one hand, the exemplar view states that a category is represented, by the set of its previously observed members, the stored exemplars (e.g., Brooks, 1978; Medin & Schaffer, 1978). According to this view, no abstraction across these encountered members takes place. On the other hand, the abstractionist view states that a category is represented by a summary representation, i.e., general category information abstracted from the members of a category. The summary representation can, for example, consist of a rule (e.g., Bourne, 1970), or the average across members (a prototype, e.g., Posner & Keele, 1968), or an ideal (e.g., Barsalou, 1985; Chaplin, Oliver, & Goldberg, 1988).

In early categorization literature, it was often assumed that there exists a single category learning and representation system, and much effort was put in providing evidence for one of the fundamentally different views on this single categorization system. The general finding emerging from this line of research was that categories are represented by previously encountered members of the category rather than by a summary representation (e.g., Busemeyer, Dewey, & Medin, 1984; Medin & Schaffer, 1978; Nosofsky & Stanton, 2005; Rehder & Hoffman, 2005; Stanton, Nosofsky, & Zaki, 2002; Storms, De Boeck, & Ruts, 2001; Voorspoels, Vanpaemel, & Storms, 2008; for reviews see Nosofsky, 1992, and Vanpaemel & Storms, 2010).

More recently however, the research focus has been shifting from efforts at deciding between the different views on category representation, to exploring potential factors that lead to the formation of different types of representation. Categorization is nowadays considered to cover a diverse load, defying captivity in a single and unitary framework (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & O'Brien, 2005; Kruschke, 2005; Minda & Smith, 2010; Poldrack & Foerde, 2008). This translates into the view that human conceptual structure is sufficiently flexible to adopt highly abstracted representations as well as rely on exemplar representations (e.g.,

☆ The Research in this article is part of research project G.0281.06 sponsored by the Belgian National Science Foundation-Flanders, given to the second author.

* Corresponding author at: Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium. Tel.: +32 16 325986.
E-mail address: wouter.voorspoels@psy.kuleuven.be (W. Voorspoels).
URL: http://ppw.kuleuven.be/concat (W. Voorspoels).

0001-6918/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2011.04.007

Love, Medin & Gureckis, 2004; Minda & Smith, 2001; Vanpaemel & Storms, 2008). In other words, while there is indeed strong evidence that in many circumstances exemplar retrieval drives categorization, sometimes summary representations are more appropriate (e.g., Ashby & Maddox, 2005; Ashby & Valentin, 2005; Johansen & Palmeri, 2003; Minda & Smith, 2001).

The dominant research question has moved from investigating which representational mode is the correct one, to understanding under which conditions a certain representation is more likely. For example, the characteristics of a category, such as size and coherence, have been shown to influence the type of representation of a category. Categories that contain only a few (encountered) members, or that contain very diverse stimuli, seem to elicit exemplar retrieval categorization processes, while larger categories or categories with a coherent family resemblance structure, are more likely to foster a central tendency representation (e.g., Blair & Homa, 2003; Homa, Sterling, & Trepel, 1981; Minda & Smith, 2001; Reed, 1978).

The present study focuses on another factor that potentially promotes or hinders abstraction: The hierarchical level of a category. Most everyday categories in the real world – such as furniture, fruit or animals – are characterized by a complex hierarchical structure. The human conceptual apparatus allows multiple categorizations of nearly all objects through a class inclusion hierarchy. For example, my neighbor's poodle, Jack, obviously is a dog. Moreover, he is undoubtedly a mammal and an animal, and unmistakably a living thing. Jack can be classified in all mentioned categories and each categorization implies membership of the higher level categories. Such a category hierarchy structure is intrinsic to the majority of conceptual domains, and has been proven an important factor in category-based tasks such as classification, feature verification and inference (e.g., Coley, Medin, & Atran, 1997; Collins & Quillian, 1969; Murphy & Lassaline, 1997; Murphy & Wisniewski, 1989; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Given the apparent salience of the hierarchical structure in everyday conceptual domains, it is reasonable to evaluate whether exemplar and abstractionist representations are differentially appropriate for different levels in an established natural language conceptual hierarchy.

1.1. Overview

The primary aim of the present study is to contribute to the growing body of evidence that the human conceptual system is sufficiently flexible to employ different types of representation. In particular, we examine whether the level in a conceptual hierarchy is a factor that might contribute to the formation of abstract representations. To this end, we will introduce a natural language concept hierarchy that contains lower level categories and an encompassing higher level. At both levels in the hierarchy we will examine whether an abstract category representation or an exemplar representation is more likely, through quantitative comparisons of their account of the typicality gradient within the categories.

We started from a rich, hierarchically structured dataset of clothes, containing members such as 'leather pants', 'bowl hat', 'plastic boots', etc. The members were generated for one of 20 subordinate categories of the concept CLOTHES (such as trousers, head wear, shoes). One group of participants rated the typicality of the members for their respective subordinate category (how typical are 'leather pants' for the category trousers?), and a different group of participants judged the typicality of the members for the superordinate concept CLOTHES (how typical are 'leather pants' for the category CLOTHES). To investigate whether the type of representation varies as a function of hierarchical level, we evaluated three models: the generalized context model (GCM; Nosofsky, 1986), and two abstractionist models, a standard multiplicative prototype model (MPM; e.g., Smith & Minda, 2000) and an ideal dimension model (IDM; Barsalou, 1985; Lynch, Coley, & Medin, 2000; Voorspoels et al., in press). The models were compared in their ability to account for the typicality gradient of the subordinate categories and of the superordinate category.

We will start by giving an overview of the computational models and the model evaluation strategy. Then we will describe in more detail the data set and the results of the analyses.

2. Models and model evaluation strategy

We tested three models that implement different views to concept representation. Formally, all three models are based on an underlying geometric representation, in which stimuli are points in an M-dimensional space and the distance between two points is inversely related to the similarity of the corresponding stimuli. Details on the formal implementation of the models can be found in Appendix A. Here we will in short discuss the main ideas that underlie the three models under consideration: the prototype view, the ideal view and the exemplar view.

The prototype view (e.g., Hampton, 1993; Coleman & Kay, 1981; Minda & Smith, 2010; Posner & Keele, 1968; Reed, 1972; Younger & Cohen, 1983) states that a category is represented by an abstract summary, referred to as a prototype. In this view, the concept vehicle is a summary representation of what vehicles are generally like, abstracted from specific instances of vehicles, containing information such as 'moves people or cargo from point A to point B'. Car is a typical member of the category vehicle because it is highly similar to the abstract vehicle prototype. We implement this approach through the MPM, a standard multiplicative prototype model (e.g., Nosofsky, 1992; Smith & Minda, 2000).

The ideal view of concept representation (Barsalou, 1985; Chaplin et al., 1988; Davis & Love, 2010; Feldman, 2000; Lynch et al., 2000) assumes that a category is represented by an ideal, constituted by the ideal dimension: a relevant characteristic, or a combination of ideal features, that sets an extreme, potentially unrealistic norm. In terms of typicality, the more a category member possesses of whatever it is the ideal dimension of the category consists of, the more typical the member is of that category. For example, the lower a food product is on calories, the more it is typical of the category 'things you eat when on a diet', and the more valuable something is, the more typical it is of 'things you rescue from a burning house' (Barsalou, 1985). For more common categories, it may be more difficult to articulate the specific combination of features that might make up the ideal (but see Burnett, Medin, Ross, & Blok, 2005; Lynch et al., 2000). To put it somewhat trivially: according to the IDM, a car is typical for the category of vehicles if it has a lot of the combination of features that make up "vehicle-ness". We implement the ideal approach through the recently proposed ideal dimension model (IDM; Voorspoels et al., in press).

The exemplar view (Brooks, 1978; Hintzman, 1986; Medin & Schaffer, 1978) proposes that a category is represented by previously encountered instances of the category. Typicality is conceptualized as the summed similarity of a category member to all stored members of the category. In this view, the concept vehicle consists of representations of previously encountered instances of vehicles, such as train, plane and metro. Car is a typical vehicle because it is highly similar to many stored instances of vehicle. In this study, the exemplar view is implemented by the generalized context model (GCM; Nosofsky, 1986).

Our model analyses will take two different, but complementary perspectives. We first investigate whether the models can adequately account for the typicality gradient. We will do this by evaluating the models on their optimal fit. Given that the models provide a reasonable account of the data, a natural follow-up question concerns which of these models capture the observed typicality ratings best. The model comparison question cannot be solved by merely looking at goodness-of-fit, since inherent complexity differences between the models will bias the conclusion. Rather, model comparison should

rely on a measure of generalizability (Myung, 2000; Pitt, Kim & Myung, 2003), which balances the inherent complexity in the models with their ability to fit the empirical data. In this study, we relied on the marginal likelihood as a measure of generalizability (Kass & Raftery, 1995). The marginal likelihood of the model is the weighted average likelihood of the model across the parameter space (Kass & Raftery, 1995; Myung & Pitt, 1997). More detailed information on the model evaluation procedures can be found in Appendix B.

3. Method

We focused on the semantic superordinate category CLOTHES. Typicality ratings were collected for two hierarchical levels. Further, because all models rely on a geometric representation of the stimuli, a pairwise similarity measure was collected to derive such a representation.

3.1. Stimuli

The stimuli were a set of 192 verbal descriptors of members of clothes generated for 20 subordinate categories taken from De Wilde, Vanoverberghe, Storms, and De Boeck (2003): accessories, baby clothes, beach clothes, bedroom clothes, coats, evening wear, casual clothing, head wear, rain clothes, shoes, skirts, sport wear, socks, sweaters, traditional clothing, trousers, underwear, uniforms, winter clothes and work clothes.[1] For each of the 20 subordinate categories, we withheld members that were generated by at least 2 participants. This selection resulted in categories containing between 9 and 18 members.

3.2. Typicality ratings

3.2.1. Participants

Participants were 180 undergraduate students that participated for course credit.

3.2.2. Procedure

Participants were asked to judge how typical the members of a category (presented as verbal stimuli) were for the category on a 10-point Likert-scale ranging from 1 (not typical at all) to 10 (very typical). The task was presented in a pen-and-paper format. Participants were given a short instruction, followed by a list of members of the category, each member accompanied by the 10-point scale. For each category, three random orders of the stimuli were used.

A first group of 160 participants rated the typicality for the subordinate categories. Participants were asked to judge how typical the members of a particular category were for that category (e.g., typicality of 'leather pants' for trousers). Each participant was asked to judge the typicality of the members of three categories, randomly selected from the set of 20. This resulted in 24 judgments for the members of each subordinate category.

A separate group of 20 participants judged the typicality of all 192 members for the superordinate category CLOTHES (e.g., typicality of 'leather pants' for CLOTHES).

3.3. Pairwise similarity measure

3.3.1. Participants

Twenty five undergraduate students participated for course credit.

3.3.2. Procedure

Given the large size of the stimulus set, we relied on a card sorting task to derive the necessary pairwise similarity measure for the members of CLOTHES (e.g., Ameel & Storms, 2006; Van der Kloot & Van Herk, 1991). Faced with the 192 word cards that were randomly spread out on a table, participants were asked to sort the cards into piles based on similarity, the only restriction being that there had to be more than one pile and less than 192. No instructions were given regarding which features, dimensions or principles should be applied in sorting the cards into piles. The initial sort served as a starting point for two additional sorts. First, the participants were asked to join their initial piles together into larger piles, if they thought it was possible. Then, again starting from the initial sort, participants were asked to further subdivide the initial piles, if deemed possible.

4. Preliminary analyses

4.1. Typicality ratings

The reliability of the typicality judgments was evaluated by the split-half correlation for the typicality judgments within each of the 21 categories (20 subordinate and the superordinate category), corrected with Spearman–Brown. Estimated reliabilities ranged from 0.83 to 0.96 across the categories. Typicality judgments were averaged across participants, resulting in a typicality average per stimulus, and a measure of data precision (the standard deviation across participants). The typicality scores on average ranged from 5 to 9.2, with an average standard deviation (averaged across categories) of 1.31.

4.2. Similarity scaling

The sorting task resulted in 75 (25 participants times 3 sorts) exemplar-by-exemplar matrices containing zeros and ones, referring to whether two exemplars were in the same pile in a particular sort or not. The matrices were then summed, producing a 192×192 matrix containing measures of similarity for each stimulus pair (i.e., the frequency of two stimuli sharing a pile, summed across the 75 sorts). The estimated split-half reliability of the sort task data was 0.97. The summed similarity matrix was used as input for a SAS MDS-analysis. In the interest of generality, solutions in 2 to 8 dimensions were considered. The appropriateness of a spatial stimulus representation was evaluated by calculating the percentage of triplets in the similarity matrix that violates the triangle inequality and a badness-of-fit measure (stress) for the MDS-solutions, which are both presented in Table 1.

As can be seen in Table 1, the number of triplets violating the triangle inequality is very small, suggesting that a spatial configuration is appropriate. The stress values show a clear monotonically decreasing pattern, suggesting that the MDS-algorithm did not get trapped in local minima. Following Kruskal (1964), solutions with stress-values exceeding 0.10 are not considered, leaving representations from Dimensionality 5 to 8 for further analyses.

5. Model-based analyses

We are interested in the representation of concepts at different levels in a hierarchy. The presented stimulus set provides us with a conceptual domain with concepts at two hierarchical levels: subordinate concepts (e.g., trousers, head gear) and a superordinate concept (CLOTHES). The three representational models give a clear and different account of the typicality gradient: the MPM relates

Table 1
Percentage of Triangle Inequality Violations (TIV) and stress-values as a function of the dimensionality of the MDS-solution.

TIV	2D	3D	4D	5D	6D	7D	8D
0.27%	0.24	0.17	0.12	0.09	0.07	0.06	0.05

[1] The study was done using Dutch category labels. Some of these labels may not cover the exact extension of the best English translations.

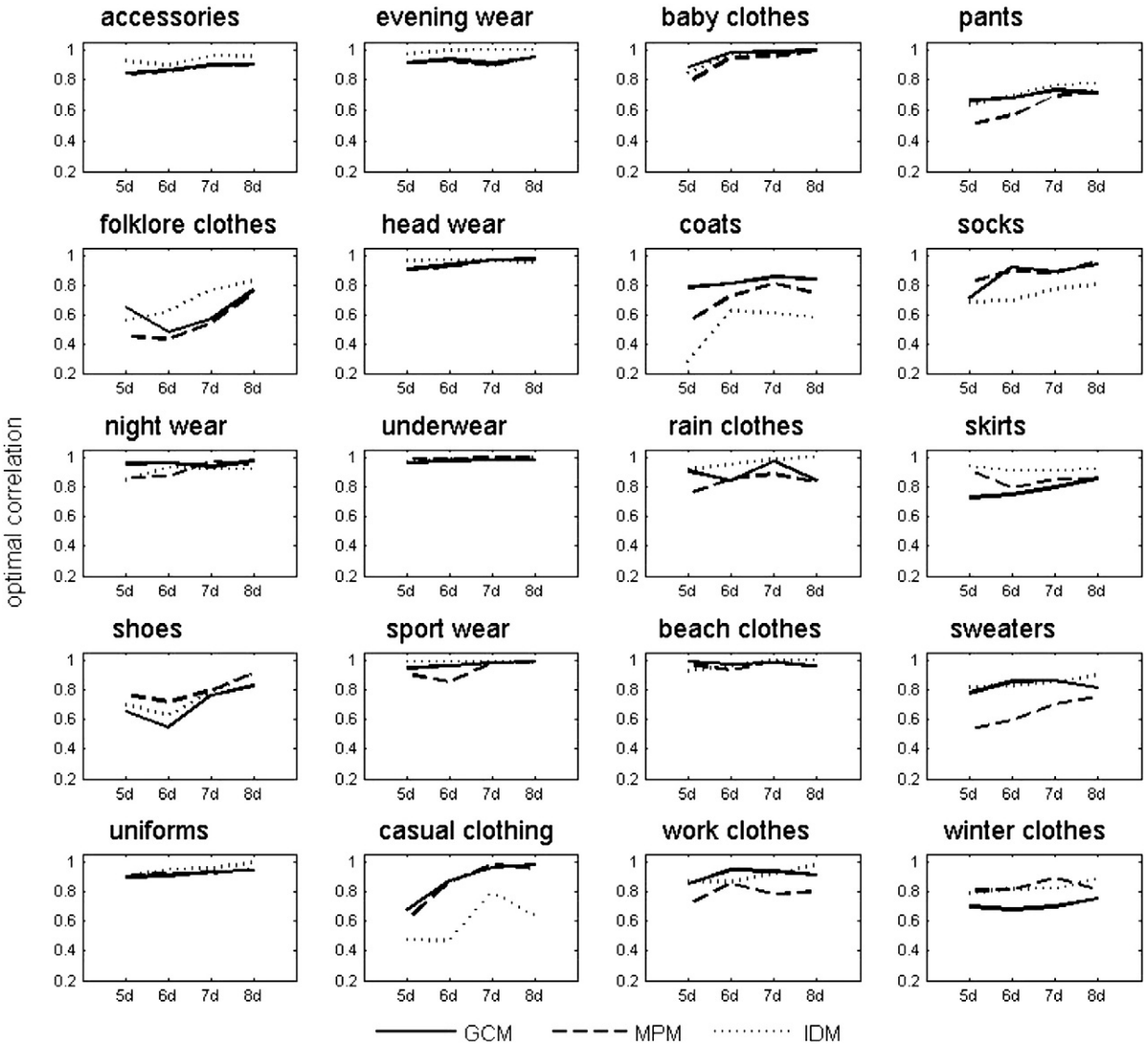


Fig. 1. Optimal correlations between observed and model-based typicality ratings for each of the 20 subordinate concepts presented for the three models as a function of dimensionality of the underlying representation.

typicality of a category member to the similarity towards the average of the category, the IDM reduces typicality to the value on an underlying ideal dimension, and the GCM accounts for typicality in terms of the summed similarity towards all other members of the category. Based on the empirical typicality gradient of the categories, and the constructed representation, we evaluate the models on two hierarchical levels: the subordinate level (e.g., the typicality of 'leather pants' for trousers) and the superordinate level (the typicality of 'leather pants' for CLOTHES). As explained earlier, the models will be evaluated on both goodness-of-fit and generalizability. The results of the model fitting and model comparison analyses are presented separately for the subordinate level categories, and the superordinate level category.

5.1. Subordinate level

Results of the model fit analyses are presented in Fig. 1. For each of the 20 subordinate categories, we calculated the optimal correlation between observed typicality ratings and model-based scores derived from the three models. The optimal correlations are presented as a function of the dimensionality of the underlying representation.

Some exceptions aside (the IDM in casual clothing and coats), it is clear from Fig. 1 that the three models provide a more than reasonable account of the observed typicality ratings in all 20 subordinate categories. Averaged across categories, the GCM, the MPM and the IDM arrive at an average optimal correlation of respectively 0.89, 0.90 and 0.88 for the eight-dimensional stimulus representation. Clearly all three models are able to provide a sufficient account of the typicality gradient of categories at this hierarchical level, making them viable accounts of the data. To decide which model captures the regularities in the data best, we need to rely on generalizability.

Results of the analyses concerning generalizability of the models are shown in Fig. 2. Presented are the model weights for each model in each category. The model weight reflects the relative evidence in favor of a particular model, given in a set of models (Burnham & Anderson, 2002; Lee, 2004; Wagenmakers & Farrel, 2004). The evidence for a particular model is quantified in the marginal likelihood of the model.[2]

[2] We relied on standard non-informative parameters prior to calculate the marginal likelihood of the models. For more information see the model descriptions in Appendix A.

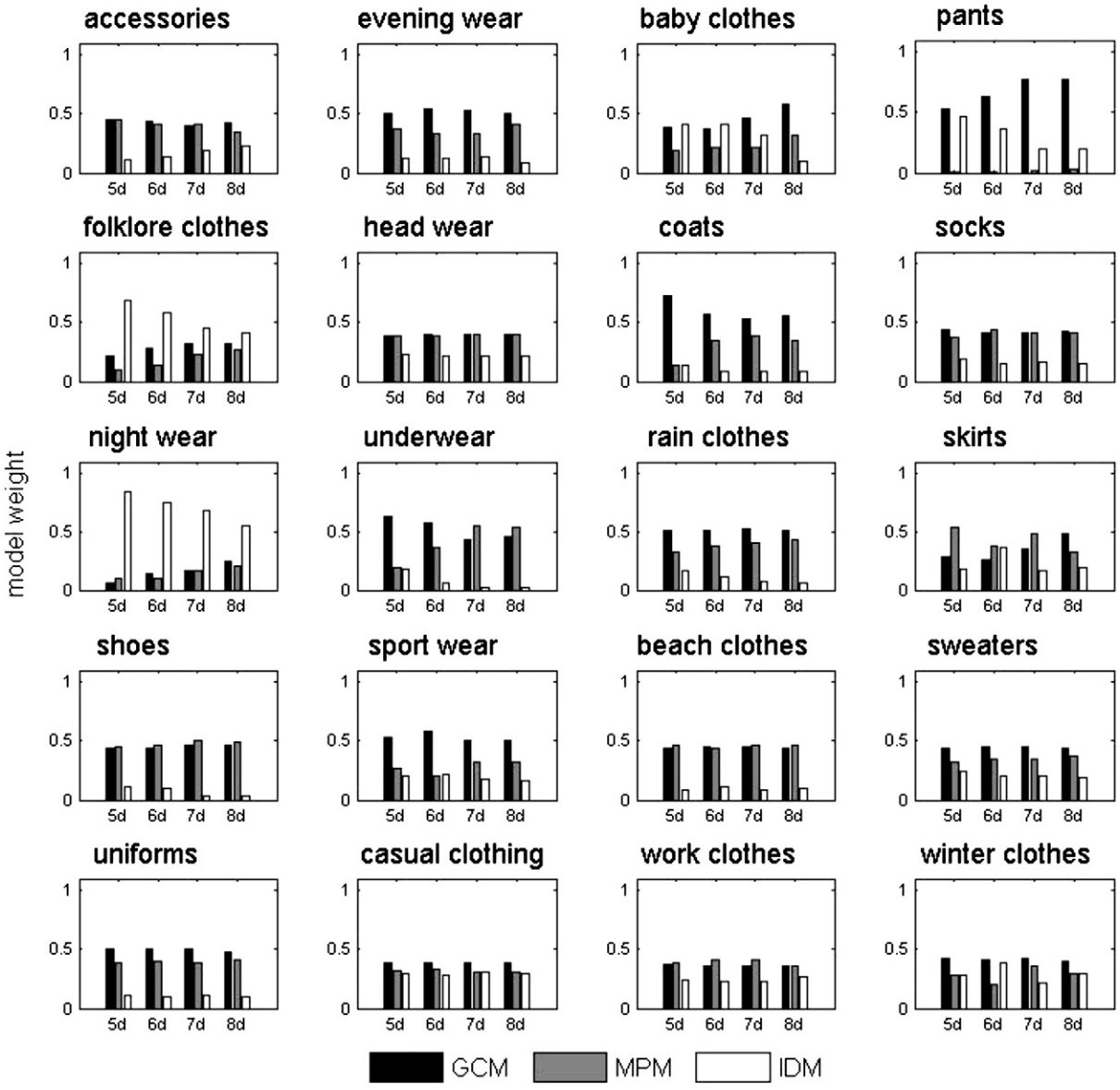


Fig. 2. Relative marginal likelihoods for the three models as a function of Dimensionality.

For 6 out of 20 categories (accessories, head wear, socks, shoes, beach clothes and work clothes), there is no clear dominant model, indicating that the data did not provide sufficient evidence to convincingly conclude for a particular model over the others. In three categories (baby clothes, underwear, and skirts), there is a clear dominant model for each dimensionality, but the results are inconsistent across dimensionalities, so drawing conclusions is hard. For the remaining 11 categories, the conclusions are clear: The GCM is the preferable model consistently across dimensionalities for 9 of these 11 categories. The IDM is the preferable model in only two categories (folklore clothes and night wear). Taken together, most support is found for exemplar representations, having the upper hand in nearly all consistent categories, and being competitive in nearly all other categories, for which differentiation between models is difficult.

5.2. Superordinate level

To investigate whether representation is affected by the hierarchical level, we evaluated the same models for the higher level concept CLOTHES, again in terms of optimal fit and generalizability. Results are summarized in Fig. 3.

In contrast to the somewhat muddled picture at the subordinate level, the results displayed in Fig. 3 are abundantly clear. The IDM is overwhelmingly supported, both in terms of optimal correlation and generalizability. Both the GCM and the MPM have problems accounting for the typicality gradient of CLOTHES, arriving at optimal correlations of 0.61 and 0.57 in Dimensionality 8. The IDM provides an adequate account of the observed typicality, reflected in an optimal correlation of around 0.80, relatively independent of the underlying dimensionality. The marginal likelihoods, in which model fit and model flexibility are balanced, provide overwhelming evidence in favor of the IDM. The model weights of both the GCM and the MPM approximate zero.

Taken together, these results clearly support the conclusion that the typicality gradient of the higher level concept CLOTHES as constituted by low level members (leather pants, bowl hats, etc.) is determined by a summary representation as proposed by the IDM, rather than by an exemplar representation or a prototype that is the average of all subordinate members.

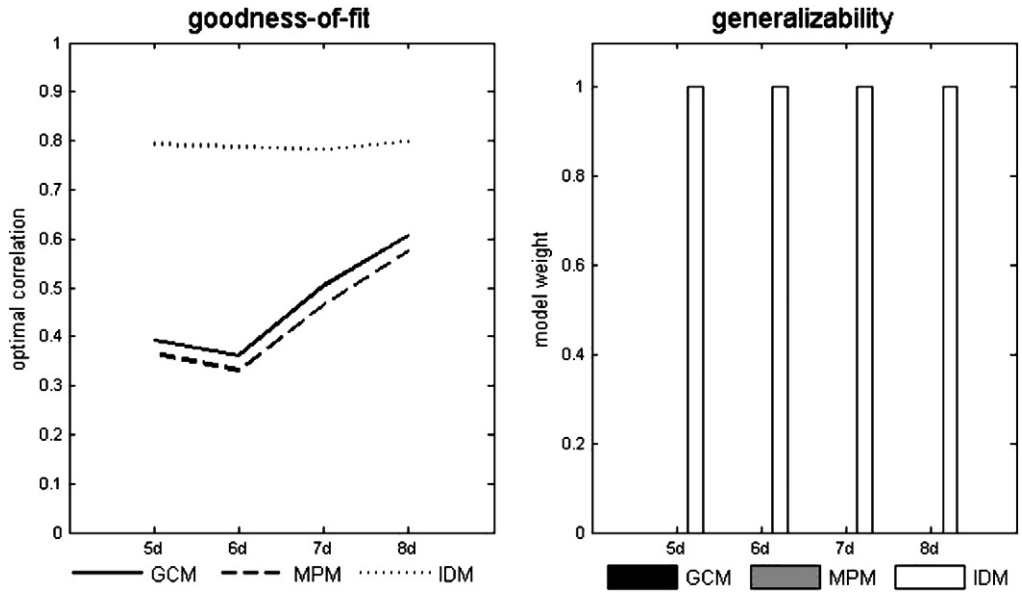


Fig. 3. Optimal correlations and relative marginal likelihoods for the superordinate concept, as a function of Dimensionality.

6. General discussion

Despite the salience and importance of conceptual hierarchy in natural language categories (Rosch, 1999), it is an open question whether a single formal model of representation can account for representation of categories at different hierarchical levels (Kruschke, 2005; Medin, Lynch, & Solomon, 2000). Most thorough model comparisons are performed in the artificial category learning paradigm, and few studies in this line of research have implemented a vertical dimension in the category structure. Exceptions are Palmeri (1999) and Verheyen, Ameel, Rogers, and Storms (2008), who found that an exemplar approach can capture some aspects of the learning of hierarchical categories. In natural language categories, on the other hand, much of the research concerning the vertical dimension in conceptual structures concerns the existence and correlates of a privileged level of categorization (the basic level, e.g., Coley et al., 1997; Lassaline, Wisniewski, & Medin, 1992; Rosch et al., 1976), and the application of computational models in this field is less widespread.

In the present study, we set out to test whether a single type of representation can sufficiently capture the typicality gradient of concepts at different hierarchical levels in natural language categories. To this end we evaluated three major views on representation – the exemplar view, the prototype view and the ideal view – in their account of the typicality gradient in natural language concepts at a lower hierarchical level (i.e., trousers) and at a higher hierarchical level (CLOTHES). We found support for exemplar representations at the subordinate level (i.e., trousers, sweaters, head wear), confirming results of earlier studies in natural language concepts and artificial category learning (Nosofsky, 1992; Storms et al., 2001; Voorspoels et al., 2008). Interestingly, at the superordinate level in the hierarchy (CLOTHES), we found overwhelming evidence for a summary representation, in the form of an ideal representation.

The finding that the human conceptual system does not only seem to accommodate exemplar representations, but also abstracted summary representations, fits nicely with the increasingly dominant view that categorization is not a unitary phenomenon. Indeed, recent empirical research has demonstrated the need for different types of representation of varying abstraction, and different systems of categorization to account for human behavior (e.g., Ashby & Valentin, 2005; Blair & Homa, 2003; Love & Gureckis, 2007; Minda & Smith, 2001; Poldrack & Foerde, 2008; Storms et al., 2001; Vanpaemel & Storms, 2010). On a more theoretical side, computational frameworks have been developed that explicitly implement the possibility of different degrees of abstraction in category representation (e.g., SUSTAIN, Love, Medin, & Gureckis, 2004; the Varying Abstraction Model, Vanpaemel & Storms, 2008), or take a hybrid approach (e.g., COVIS, Erickson & Kruschke, 1998). These models explicitly acknowledge the flexibility of our cognitive system, by allowing different representations and degrees of abstraction, a view that has provided a promising extension of the categorization literature.

While the present study is, to our knowledge, the first to evaluate different computational models of representation on different levels in a natural language conceptual hierarchy, our findings can be understood in the light of earlier findings in artificial category learning. First, higher level categories differ from lower level categories in their size, a factor that has been shown influential in the debate on abstraction in artificial category learning. Categories that contain many members favor representations containing category general information, i.e., summary representations (Homa et al., 1981; Minda & Smith, 2001), which is obviously more cognitively economical than an exemplar representation. The IDM proposes a summary representation in the form of an ideal, and therefore may be the more appropriate model at a level where economy is more demanding.

The question then rises why the average prototype, a more popular conception of a summary representation, is not competitive in the higher level categories. Possibly, the within category similarity of higher level categories is too low. It is difficult to think of a sensible average of a set of items that are as diverse as leather pants, soccer shoe, bowl hat and rain jacket. Rosch et al. (1976) indeed proposed that the basic level is the highest level at which a more or less unitary image can be formed of a category. Rising in a hierarchy makes it increasingly difficult to form such an image, and consequently an average prototype might become more unlikely.

This argument is closely related to a second potential factor that influences category representation: Higher level categories are less coherent than subordinate categories. Objects in higher level categories have less in common than members in lower level categories, and what they have in common – such as their function or other abstract properties – is qualitatively different from the commonalities at a lower level, where perceptual features are often shared (Barsalou, 1991; Murphy & Wisniewski, 1989; Rosch et al.,

1976). Whereas subordinate categories reflect a structure of co-varying (perceptual) properties in the world (Rosch & Mervis, 1975), the representation of superordinate categories may depend less on the inherent structure of the world, allowing more influence of cultural utility (Malt, 1995) — for example through cultural goals or implicit norms. As Malt (1995) notes, superordinate categories indeed reflect more cross-cultural diversity, which points to a greater dependence on culture-specific rather than environmental factors.

For lack of an imposing feature structure in the world ensuring category coherence, an ideal representation may constitute the relevant dimension, derived from cultural utility, that makes the higher level category coherent, despite the obvious (perceptual) differences between the members. In this sense, higher level categories perhaps resemble goal-derived categories, for which ideal representations have been proven successful (Barsalou, 1985). Only after knowledge of what a goal derived category is about ('things you rescue form a burning house'), a seemingly random set of objects (car keys, pet dog, love letters, baby) makes sense as a category, and the dimension that demarcates and holds the category together is obvious ('valuableness'). Perhaps for higher level concepts too, category coherence is reduced to a single relevant dimension that sets a norm for category inclusion and typicality. As such, the IDM resembles a rule-based approach, where a rule can be conceptualized as one dimensional similarity, in this case towards an ideal or a norm (Pothos, 2005).

A noteworthy characteristic of the present study is that it focused on the everyday artifact category, CLOTHES. Future work will have to consider whether the results generalize to non-artifact, natural categories, such as ANIMALS or PLANTS. One can wonder, for example, whether an ideal would be relevant in the context of high level natural categories, given that cultural utility may be less applicable for these categories. Another extension of the present findings would be to the domain of artificially learned categories. As already noted, most research in this domain has focused on subordinate level categories that are not embedded in a hierarchy.

7. Conclusion

We presented evidence that, in a conceptual hierarchy, exemplar representations are more likely for concepts at lower levels and that ideal representations become more likely when rising in hierarchy. Possibly, this finding can be related to the category characteristics of both lower and higher level concepts. Lower level categories count fewer members, which reflect a higher within category coherence, through the sharing of many common, perceptual features. Higher level categories on the other hand count many more members, and have a lower, and possibly qualitatively different, within category similarity structure, allowing more influence of cultural utility in the representation. The present study adds to the growing support for the view that multiple systems of categorization might be necessary to explain human categorization behavior.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.actpsy.2011.04.007.

References

Ameel, E., & Storms, G. (2006). From prototypes to caricatures: Geometrical models for concept typicality. *Journal of Memory and Language*, *55*, 402–421.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Science*, *2*, 83–89.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442–481.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 49–178.

Ashby, F. G., & Valentin, V. V. (2005). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 547–572). New York: Elsevier.

Barsalou, L. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *11*, 629–654.

Barsalou, L. W. (1991). Deriving categories to achieve goals. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, vol. 27*. (pp. 1–64). San Diego: Academic press.

Blair, M., & Homa, D. (2003). As easy to memorize as they are to classify: The 5-4 categories and the category advantage. *Memory & Cognition*, *31*, 1293–1301.

Bourne, L. E. (1970). Knowing and using concepts. *Psychological Review*, *77*, 546–556.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum.

Burnett, R. C., Medin, D. L., Ross, N. O., & Blok, S. (2005). Ideal is typical. *Canadian Journal of Experimental Psychology*, *59*, 3–10.

Burnham, K., & Anderson, D. (2002). *Model selection and multi-model inference: A practical, information-theoretic approach.* New York: Springer.

Busemeyer, J. R., Dewey, G. I., & Medin, D. L. (1984). Evaluation of exemplar-based categorization and the abstraction of categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 638–648.

Chaplin, W., Oliver, J., & Goldberg, L. (1988). Conceptions of states and traits: Dimensional attributes with ideals as prototypes. *Journal of Personality and Social Psychology*, *54*, 541–557.

Coleman, L., & Kay, P. (1981). Prototype semantics: The English verb lie. *Language*, *57*, 1–13.

Coley, J. D., Medin, D. L., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*, 73–112.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240–247.

Davis, T., & Love, C. B. (2010). Memory for category information is idealized through contrast with competing options. *Psychological Science*, *21*, 234–242.

De Wilde, E., Vanoverberghe, V., Storms, G., & De Boeck, P. (2003). The instantiation principle re-evaluated. *Memory*, *11*, 533–548.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology. General*, *127*, 107–140.

Feldman, J. (2000). Bias toward regular form in mental shape spaces. *Journal of Experimental Psychology. Human Perception and Performance*, *26*, 1–14.

Hampton, J. A. (1993). Prototype models of concept representation. In I. Van Mechelen, J. A. Hampton, R. S. Michalski, & P. Theunis (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 67–95). London: Academic Press.

Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 418–439.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, *45*, 482–553.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Kruschke, J. K. (2005). Category learning. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognitive psychology* (pp. 1–57). London: Sage.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1–27.

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). Basic levels in artificial and natural categories: Are all basic levels created equal? In B. Burns (Ed.), *Percepts, concepts, and categories: The representation and processing of information*. Amsterdam: North-Holland.

Lee, M. D. (2004). A Bayesian analysis of retention functions. *Journal of Mathematical Psychology*, *48*, 310–321.

Love, B. C., & Gureckis, T. M. (2007). Models in search of a brain. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 90–108.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *111*, 309–332.

Lynch, E. B., Coley, J. D., & Medin, D. L. (2000). Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices. *Memory & Cognition*, *28*, 41–50.

Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology*, *29*, 85–148.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology*, *51*, 121–147.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*, 775–799.

Minda, J. P., & Smith, J. D. (2010). Prototype models of categorization: Basic formulation, predictions, and limitations. In E. Pothos & A. Wills (Eds.), *Formal approaches in categorization*. Cambridge, UK: Cambridge University Press.

Murphy, G. L., & Lassaline, M. E. (1997). Hierarchical structure in concepts and the basic level of categorization. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 93–131). Hove, East Sussex, UK: Psychology Press.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*, 572–586.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, *4*, 79–95.

Myung, J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psycholology*, *44*, 190–204.
Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology. General*, *115*, 39–57.
Nosofsky, R. M. (1992). Exemplars, prototypes, and similarity rules. In A. F. Healy & S. M. Kosslyn (Eds.), *Essays in honor of William K. Estes, vol. 1: From learning theory to connectionist theory; vol. 2: From learning processes to cognitive processes* (pp. 149–167). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
Nosofsky, R. M., & Stanton, R. D. (2005). Speeded classification in a probabilistic category structure: Contrasting exemplar retrieval, decision-boundary, and prototype models. *Journal of Experimental Psychology. Human Perception and Performance*, *32*, 314–334.
Palmeri, T. (1999). Learning categories at different hierarchical levels: A comparison of category learning models. *Psychonomic Bulletin & Review*, *6*, 495–503.
Pitt, M. A., Kim, W., & Myung, J. (2003). Flexibility versus generalizability in model selection. *Psychonomic Bulletin & Review*, *10*, 29–44.
Poldrack, R. A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience and Biobehavioral Reviews*, *32*, 197–205.
Posner, M., & Keele, S. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.
Pothos, E. M. (2005). The rules versus similarity distinction. *The Behavioral and Brain Sciences*, *28*, 1–49.
Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.
Reed, S. K. (1978). Category versus item learning: Implications for categorization models. *Memory & Cognition*, *6*, 612–621.
Rehder, B., & Hoffman, A. B. (2005). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*, 811–829.
Rosch, E. (1999). Principles of categorization. In E. Margolis & S. Laurence (Eds.), *Concepts: Core readings (Chapter 8)*. Cambridge, MA: MIT Press.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *64*, 217–228.
Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.
Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*, 3–27.
Stanton, R. D., Nosofsky, R. M., & Zaki, S. R. (2002). Comparisons between exemplar similarity and mixed prototype models using linearly separable category structure. *Memory & Cognition*, *30*, 934–944.
Storms, G., De Boeck, P., & Ruts, W. (2001). Categorization of unknown stimuli in well-known natural concepts: A case study. *Psychonomic Bulletin & Review*, *8*, 377–384.
Van der Kloot, W. A., & Van Herk, H. (1991). Multidimensional scaling of sorting data: A comparison of three procedures. *Multivariate Behavioral Research*, *26*, 563–581.
Vanpaemel, W., & Storms, G. (2008). In search of abstraction: The varying abstraction model of categorization. *Psychonomic Bulletin & Review*, *15*, 732–749.
Vanpaemel, W., & Storms, G. (2010). Abstraction and model evaluation in category learning. *Behavior Research Methods*, *42*, 421–437.
Verheyen, S., Ameel, E., Rogers, T. T., & Storms, G. (2008). Learning categories at different levels of abstraction. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 751–756). Austin, TX: Cognitive Science Society.
Voorspoels, W., Vanpaemel, W., & Storms, G. (2008). Exemplars and prototypes in natural language concepts: A typicality based evaluation. *Psychonomic Bulletin & Review*, *15*(3), 630–637.
Voorspoels, W., Vanpaemel, W., & Storms, G., (in press). A formal ideal-based account of typicality.
Wagenmakers, E. J., & Farrel, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*, 192–196.
Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Child Development*, 858–867.