Acta Psychologica 170 (2016) 94–102

Contents lists available at ScienceDirect

Acta Psychologica

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

Understanding individual differences in representational abstraction: The role of working memory capacity☆

Loes Stukken *, Bram Van Rensbergen, Wolf Vanpaemel, Gert Storms

University of Leuven, Belgium

ARTICLE INFO

Article history:
Received 5 November 2015
Received in revised form 1 June 2016
Accepted 7 June 2016
Available online 3 July 2016

Keywords:
Category learning
Categorization
Individual differences
Representational abstraction
Working memory

ABSTRACT

Several studies have reported differences in categorization strategies among participants: some learn a category by making abstraction across the category members while others use a memorization strategy. Despite the prevalence of these differences, little attention has been paid to investigating what influences some to use an abstraction strategy and others a memorization strategy. The current study had two goals: in a first experiment we investigated whether these differences were stable across time, using the parallel form method often used in psychometric research, and in a second experiment we investigated whether the individual differences in categorization strategy were related to working memory capacity. We used a modelling strategy, in which we not only focused on full abstraction and memorization strategies, but also on intermediate strategies in which some category members are abstracted and others are not. The first study revealed that the individual abstraction strategy of individual participants in two different experiments, performed at different times, correlate significantly, and second study showed that these individual differences were related to the working memory capacity of the participants.

© 2016 Elsevier B.V. All rights reserved.

It is well known that there are differences in the strategies people rely on when learning categories from labelled exemplars. As early as 1984, Medin, Altom, and Murphy (1984) asked participants to report the strategy they used when learning a category. These self- reported strategies could be classified in five different types, including a prototype and memorization strategy, and the relative frequencies of the types indicated that they were fairly equally used. A similar questionnaire approach was used by Little and McDaniel (2015a), showing that 37% of their participants classified themselves as exemplar-learners, while 51% classified themselves as rule-learners. Transfer performance confirmed the validity of the self-reported strategies. Opting for a different methodology, Craig and Lewandowsky (2012) used the response patterns on transfer items to identify the strategy of individual participants. The two most popular strategies used in their study were a rule-based and an exemplar strategy. The same transfer method was employed by McDaniel, Cahill, Robbins, and Wiener (2014) to discriminate between individuals using a rule-based approach or an exemplar strategy. Using a combined analysis of the response patterns and a modelling perspective, Johansen and Palmeri (2002) reported interindividual differences in addition to intraindividual differences in learning ill-defined categories: by the end of training, participants differed in the strategy they used. They furthermore found evidence that strategies shifted over the course of learning. These results were later confirmed by Raijmakers, Schmittmann, and Visser (2014) using a latent Markov analysis. Others used a modelling perspective to find differences in the extent to which participants used a prototype versus an exemplar strategy. Smith, Murray, and Minda (1997), for example, contrasted an exemplar model and a prototype model using the data of individual participants. Their results indicated that a large subgroup of their participants (40% of the participants across the two experiments) learned the categories by abstracting a prototype while the others used an exemplar strategy (though see, Nosofsky & Zaki, 2002).

Given the strong evidence for the presence of differences in categorization strategy, surprisingly little attention has been paid to understanding these differences. One of the first and most important steps that has to be taken when investigating these differences is to establish whether they are meaningful and represent vast individual differences in categorization strategy. It is therefore worthwhile to first of all demonstrate that these differences are stable over time. A recent study of McDaniel et al. (2014) showed that individuals diverge in the type of strategy they use in a function-learning task and that the preferred strategy generalized to an abstract coherent categorization task. Thus, individual differences in the function-learning task were quite stable across tasks. In Study 1, we further investigated whether individual differences in categorization

☆ Author note: Loes Stukken, Laboratory of Experimental Psychology, University of Leuven, Belgium; Bram Van Rensbergen, Laboratory of Experimental Psychology, University of Leuven, Belgium; Wolf Vanpaemel, Quantitative Psychology and Individual Differences, University of Leuven, Belgium; Gert Storms, Laboratory of Experimental Psychology, University of Leuven, Belgium.

* Corresponding author at: Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102 B3711, 3000 Leuven, Belgium.
E-mail address: loes.stukken@ppw.kuleuven.be (L. Stukken).

http://dx.doi.org/10.1016/j.actpsy.2016.06.002
0001-6918/© 2016 Elsevier B.V. All rights reserved.

strategies can be reliably assessed, using a parallel testing perspective (Lord & Novick, 1968). If the differences in categorization strategy represent vast individual differences, we would expect these differences to be stable over time.

Secondly, the stability of these differences over time would allow us to investigate what makes people different with respect to their categorization strategy. One factor that has been linked to categorization strategy is working memory capacity. Since categorization requires the active processing and manipulation of category information, working memory can be expected to play a role in category learning (see e.g., Craig & Lewandowsky, 2012). Furthermore, Sewell and Lewandowsky (2012) found a relationship between working memory capacity and the extent to which people are able to coordinate different types of partial-knowledge structures, in this case two different types of rules. They showed that participants with higher working memory capacity were better able to coordinate these two types of rules and to adjust their categorization judgment to each of these rules when expected. However, Craig and Lewandowsky (2012) and Little and McDaniel (2015a) found no such relation at all.

In the present study, we further investigate whether the individual differences in categorization strategy were related to working memory capacity. More specifically, following Sewell and Lewandowsky (2012), we expect that the higher one's working memory capacity, the easier it is to actively retain different types of information, and thus the lower the individuals' tendency to make abstraction across the category members. The lower one's working memory capacity, on the other hand, the more need to summarize information and thus the more need to make abstraction across the category members. In sum, we hypothesize that participants with lower working memory capacity will make more abstraction across category members than participants with a higher working memory capacity.

In the next section, we will outline the modelling framework that we used to determine the level of abstraction employed by the participants in their categorization strategies. Then, we discuss two experiments, one in which the stability of individual differences in abstraction was investigated and one in which we related these individual differences to working memory capacity. In the discussion, we relate our findings to previous studies.

1. Assessing representational abstraction

There is a long-standing tradition in the categorization literature to rely on formal models to determine the underlying categorization strategy that participants use. Some models assume a single type of category representation, like the Generalized Context Model (GCM; Nosofsky, 1984) and ALCOVE (Kruschke, 1992), which assume an exemplar strategy; the Multiplicative Prototype Model (Reed, 1972; Nosofsky, 1987; Smith & Minda, 2000), which is based on a prototype representation, or the decision-bound theory (Ashby & Gott, 1988; Ashby & Maddox, 1993) which is built around a rule strategy. Other, so called hybrid models combine several (mostly two) strategies. Examples of hybrid models are RULEX (Nosofsky, Palmeri, & Mckinley, 1994), SUSTAIN (Love, Medin, & Gureckis, 2004), the rational model (Anderson, 1991), COVIS (Ashby, Alfonso-Resse, Turken, & Waldron, 1998) and ATRIUM (Erickson & Kruschke, 1998). Within the RULEX framework it is, for example, assumed that during category learning subjects start by extracting a rule while storing exceptions to this rules separately. The rational model interpolates between an exemplar and prototype representation by assuming that groups of category members are clustered together. Whenever a new stimulus is encountered, it can be added to an existing cluster or it can form a new cluster (see also SUSTAIN, for comparable ideas). In the present study, we rely on a similar hybrid model, the Varying Abstraction Model (VAM; Vanpaemel & Storms, 2008), to determine the degree to which individual participants make abstraction across category members.

The VAM defines category representations that lie on a continuum from minimal (separate exemplars) to maximal abstraction (a single prototype). A category representation is formed by dividing the points in the Multidimensional scaling (MDS) space that make up a category into clusters and by averaging the coordinates of the points that were clustered together. The resulting coordinates define a set of subprototypes that make up the category representation. For example, if a cluster Q_j contains n_j stimuli, the coordinate value μ_{jk} for subprototype j on dimension k is calculated by averaging the coordinate value of all n_j stimuli on dimension k:

$$\mu_{jk} = \frac{1}{n_j} \sum_{x_i \in Q_j} x_{ik},$$

where x_{ik} is the coordinate value of stimulus i on dimension k.

The number of different possible representations depends on the number of category members to be learned. A category with four members, for example, can lead to 15 different representations according to the VAM (see Fig. 1). A category with six members can be represented by 203 different representations.

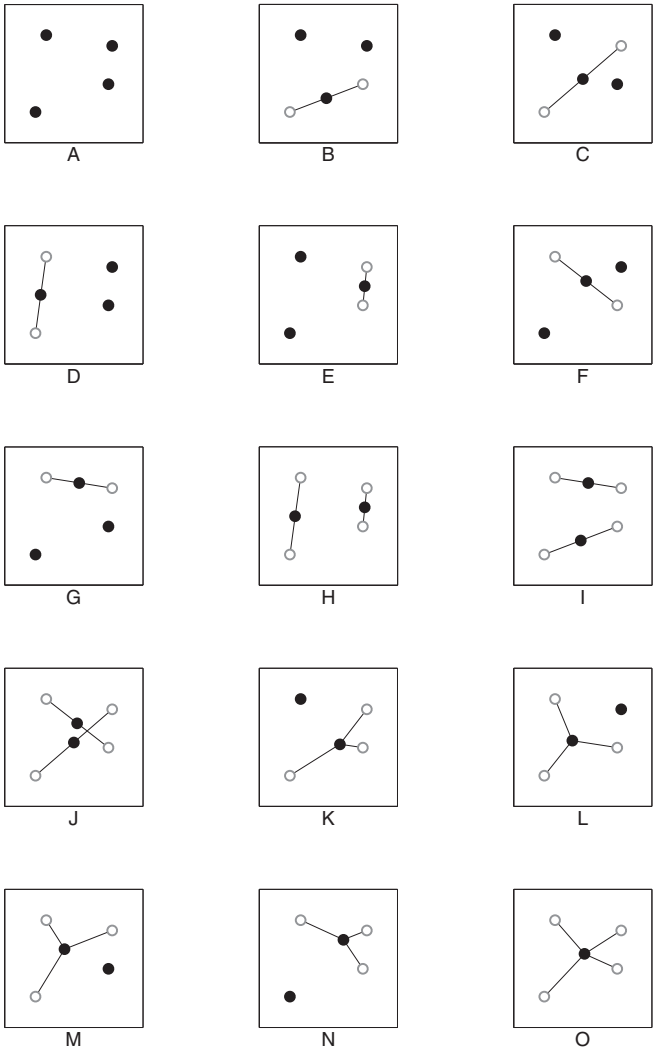

Fig. 1. The 15 possible representations for a category with four members. The subprototypes are represented by the black circles and are connected by lines to the original category members (the white circles). Panel A shows the exemplar representation (no members are merged); Panel O shows the prototype representations (all members are merged together into a single subprototype); Panels B–G show intermediate representations with three subprototypes; and Panels H–N show intermediate representations with two subprototypes. Figure with approval taken from Vanpaemel and Storms (2008).

The VAM uses the same processing assumptions as the GCM, but instead of calculating categorization predictions for only the exemplar representation, it derives predictions for all the representations on the continuum from minimal to maximal abstraction and offers a way to systematically compare these representations. Thus, for every possible representation, the VAM derives predictions by calculating the similarity of the presented stimuli to the category representation in the MDS space. Within the VAM, this means that the similarity of stimulus i to a category A, (S_{iA}) is calculated by summing the similarities between the stimulus i and the subprototypes that constitute category A:

$$S_{iA} = \sum_{\mu_j \epsilon S_A} s\left(x_i, \mu_j\right),$$

in which S_A denotes the set of subprototypes representing A.

The similarities between the stimulus and the subprototypes are computed in the same way as in the GCM: by determining the distances between the stimulus and the subprototypes that make up the category in the multidimensional psychological space, and converting these distances in a measure of similarity. The distance between stimulus i and subprototype j is calculated by:

$$d\left(x_i, \mu_j\right) = \left[\sum_{k=1}^{D} w_k \left|x_{ik} - \mu_{jk}\right|^2\right]^{\frac{1}{2}}$$

where x_{ik} and μ_{jk} are the coordinate values of stimulus i and subprototype j on dimension k and $0 < w_k < 1$ is an attentional weight associated with dimension k. These distances are then converted to obtain a measure of the similarity from each stimulus to every subprototype

$$s\left(x_i, \mu_j\right) = e^{-cd\left(x_i, \mu_j\right)},$$

where c is the sensitivity parameter.

Like the GCM, the VAM then uses Luce's choice rule to derive categorization predictions from these similarities. The probability that stimulus i is categorized in category A is in the VAM, for example, given by:

$$p_{iA} = S_{iA}/(S_{iA} + S_{iB}).$$

Since the VAM shares its processing assumptions with the GCM, it also has the same parameters: the sensitivity parameter (i.e., the rate at which similarity declines with distance) and attentional weights, one per dimension (i.e., the amount of attention given to a particular dimension k).

To keep the number of possible representations within feasible limits, a k-means clustering version of the VAM was introduced by Verbeemen, Vanpaemel, Pattyn, Storms, and Verguts (2007); (see also Stukken, Storms, & Vanpaemel, 2011). In this version, k-means clustering is used to select the most plausible clustering of category members at each level of abstraction. In this way, similar exemplars are assigned to the same cluster while dissimilar members are kept separate. The k-means clustering procedure would, for example, select category representations A (exemplar representation), E (out of the six representations with three subprototypes), N (out of the seven representations with two subprototypes) and O (prototype representation) from the representations shown in Fig. 1. When several clusterings at a given level of abstraction were considered to be equally plausible by the k-means clustering procedure, they were all included in the analysis.

We assess the degree of abstraction used by a participant by determining the number of clusters of the best fitting representation: the more clusters that make up a representation the less abstract the representation is.

2. Study 1: the stability of individual differences

The goal of the first study was to investigate whether individual differences in the tendency to make abstraction across category members assessed using the VAM are stable over time. In order to check the stability of the individual differences we used the parallel form method which is often used in psychometrics to evaluate the reliability of tests assessing characteristics like, for example, personality, intelligence and skills of the individual (Lord & Novick, 1968). Within a parallel form study, participants are presented with two tests, which are similar but not identical, and both tests are administered with days or weeks in between. If several assumptions are met, the correlation between these scores is an approximation to the reliability of the test. In this study, participants performed two different, but similar, category-learning tasks. Using the VAM, we determined for every participant the degree of abstraction in each session. The correlation between both scores is an estimate of the degree to which the individual differences in categorization strategy were stable over time. To be sure that other factors like the category structure or the presentation order of the training stimuli were not influencing the relationship between the categorization scores of the two sessions we decided to test the relationship quite strictly: Where, in psychometrics the two tests are as similar as possible, we made sure in this study that the category structure and the presentation order were different across the two sessions so that they could not influence the relation between the abstraction scores of the two sessions. We wanted to have a clear estimate of the extent to which the categorization strategy of participants is driven by the categorizer self. Given that the category structure and presentation order have, among other factors, an important influence on the categorization strategy of the participants, we don't expect the relationship between the two sessions to be as high as is typically reported in parallel testing in psychometric research.

In addition to categorization strategy, the test-retest method also allows us to check the stability of the parameters of the model. We expected no relation between the values of the w parameter across the two sessions. The value of this parameter represents the amount of attention that is spent on the dimensions of the stimuli, which is dependent upon the type of stimuli and the category structure of the categories that are learned (Nosofsky, 1987). The c parameter on the other hand, represents the degree to which an individual is sensitive to differences between category members (Nosofsky, 1986). Individuals with a high c parameter are more sensitive to differences between category members and are therefore stricter when judging instances as part of the category than individuals with a low c parameter. Although the value of the c parameter is dependent upon a range of factors external to the individual, like stimulus materials and experience with the stimuli (Nosofsky, 1987), it is likely that individuals also vary in the extent to which they are sensitive to differences and that the values of the c parameter are at least partly depending upon a person-specific characteristic of the individual. We therefore expected a relation between the values of the c parameter across sessions.

The raw and processed data are available on osf.io/uynbf/.

2.1. Method

2.1.1. Participants

A total of 97 participants took part in the experiment, with 64 of them reaching the learning criteria[1] in both sessions. The data of these 12 men and 52 women (mean age: 26.08) were used in further analysis.

2.1.2. Stimuli and category structure

The stimuli used in the two different sessions of the experiment were created by Op De Beeck, Wagemans, and Vogels (2001, 2003)

[1] The exact criterion is given below. It was set before data collection.

using seven sinusoidal functions (radial frequency components, RFC's) that were defined by three parameters: frequency, phase, and amplitude. The two different stimulus types (see Figs. 2 and 3) were created based on distinct RFC's. To induce within-group differences in shape, five RFC's were kept constant, while the amplitude of the other two (henceforth the stimuli are two dimensional) were independently varied in small steps. Every combination of the amplitude values of these two RFC's resulted in a specific stimulus (see Figs. 2 and 3 for an overview of the stimuli). Accordingly, stimuli lying close to each other in the grid are quite similar because the differences in amplitude values of the two RFC's are small, while stimuli lying further apart have some distinct differences. Op De Beeck et al. (2003) showed that people roughly represent these stimuli in a psychological space that corresponds to the parametric space in which the stimuli are constructed. The psychological space has the same number of dimensions and the stimuli are ordered in the same way. These stimuli were chosen because they are processed in an integral way (Arguin & Saumier, 2000), which prevented participants from using a unidimensional rule.

Half of the participants were presented with the stimuli from Fig. 2 in the first session and the stimuli from Fig. 3 in the second session. The other half were presented with the stimuli of Fig. 3 in the first session and the stimuli of Fig. 2 in the second session. For each stimulus type, we constructed two categories by dividing the 36 stimuli in Figs. 2 and 3 into two groups (categories A and B) along the vertical axis. For each participant separately, 12 stimuli (6 from each category) were randomly drawn as training stimuli in the experiment. The stimuli that were not selected as training stimuli were used as transfer stimuli in the transfer phase.

2.1.3. Procedure

Participants were invited twice to the laboratory and on each occasion were asked to complete a category learning task. The two sessions were at least two weeks separated in time. The category learning task consisted of two phases: a learning phase and a transfer phase. In the learning phase, participants were trained to distinguish between what was called creatures that belonged to two recently discovered species on, depending on the session, either Mercury or Pluto. Their goal was to learn which creatures belong to category A and which to category B. Both categories were given non-existing names: for half of the participants the two categories were called *Mapapilios* and *Machaons* in the first session and *Lovotomers* and *Lochiers* in the second session and vice versa for the other half. The names were furthermore counterbalanced: For half of the participants category A was called *Mapapilios* for the other half category A was called *Machaons* (the same holds for *Lovotomers*/*Lochiers*). The learning phase was organized in blocks of trials in which all the stimuli were shown once in a random order. After a stimulus was presented, the participants were instructed

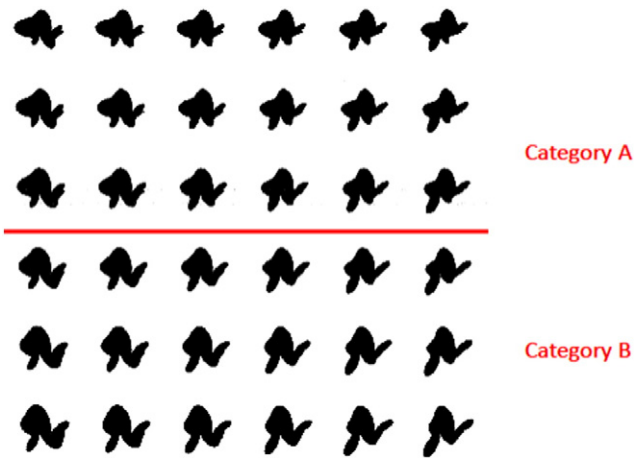

Fig. 2. First type of stimuli.

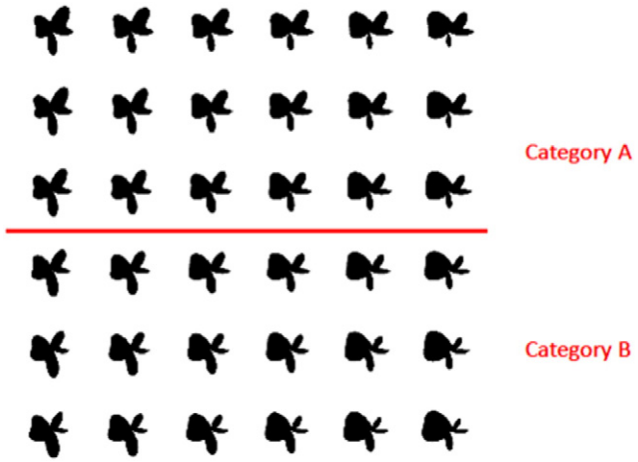

Fig. 3. Second type of stimuli.

to indicate whether the stimulus belonged to the category of the *Mapapilios*/*Lovotomers* or to the category of the *Machaons*/*Lochiers*. Participants received feedback after each response and were encouraged to use this feedback to obtain a higher percentage correct in the next block. At the end of each block, participants were shown the obtained percentage correctly categorized stimuli. The learning stage ended when the participant was able to categorize at least 90% of the stimuli correct in two consecutive blocks. The number of learning blocks were limited to 30, meaning that after 30 learning blocks participants were automatically redirected to the transfer phase. However the data of the participants who did not reach the learning criterion within 30 blocks were not used for further analysis.

In the transfer phase, the training and transfer stimuli were shown one by one and the task for the participants was to judge whether the presented stimulus belongs to the category of the *Mapapilios*/*Lovotomers* or to the category of the *Machaons*/*Lochiers*, based on what they learned in the training phase. This phase consisted of five blocks in which all 36 stimuli of Figs. 2 or 3 were shown in a random order. No corrective feedback was provided in this phase.

2.2. Data analysis and results

2.2.1. Learning phase

From the 98 participants in the experiment, 64 reached the learning criterion in both sessions.[2] Participants needed on average 8.02 blocks (standard deviation: 6.22) to reach the learning criterion in Session 1 and on average 9.08 blocks (standard deviation: 7.06) to reach the learning criterion in Session 2. There was no significant difference between the two types of stimuli in the number of learning blocks needed ($t(63) = -0.84$, $p = 0.40$).

2.2.2. Transfer phase

2.2.2.1. Fits. For each participant individually, the degree of abstraction was assessed by fitting the VAM to the data of the transfer phase using a grid-search routine and maximum likelihood as a loss function. Fig. 4 gives an overview of the distribution of the R-squared and the values of the c and w-parameters for the two sessions across

[2] Ten participants didn't show up for the second session, the data for 3 participants were incorrectly collected (more information on osf), 18 participants did not reach the learning criterion in either Session 1 or 2 and two participants did not reach the learning criterion in both sessions. The criterion was used strictly: even participants that almost reached the 90% correct limit in the 30th block were not used in further analyses. The percentage correct in the final block for participants that did not reach the learning criterion varied between 0.42 and 0.83. The raw data of all these participants are also available on osf.io/uynbf/.

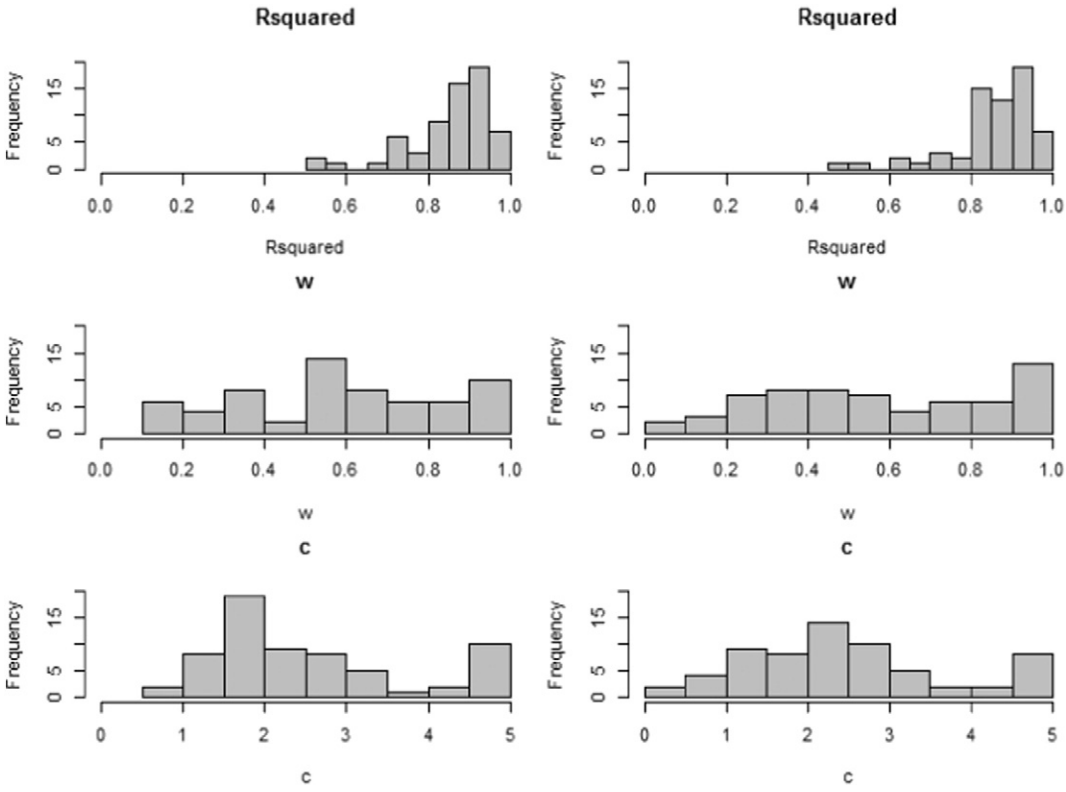

Fig. 4. The distributions of R-squared and the values of the w and c parameter across the best fitting representations of all participants. The distributions for the first session are represented on the left side, the distributions for the second session on the right side.

participants. The average R-squared across participants for Session 1 (0.86) and Session 2 (0.85) indicated that the model fitted the data reasonably well.

2.2.2.2. The stability of individual differences in representational abstraction. The fit of the VAM to the individual data of the transfer phase resulted in a best fitting representation for each participant for both types of stimuli and for each category. The number of clusters in each representation gives an indication of the degree of abstraction in the category representation. Fig. 5 gives an overview of the number of times a specific representation provided the best fit for the stimuli of the upper half of Figs. 2 and 3 (Category A) and the stimuli of the lower half of Figs. 2 and 3 (Category B), across participants. The number of clusters for the stimuli of the upper half of Figs. 2 and 3 is represented on the X-axis, the number of clusters for the stimuli of the lower half of Figs. 2 and 3 on the Y-axis. The colored bars indicate the frequency with which a specific combination of clusters was identified as the optimal representation.

For example, the optimal representation of two participants yielded one cluster for the upper half of Fig. 2 and three clusters for the lower half of Fig. 2. As can be seen in Fig. 5, for both types of stimuli there was considerable variation in the number of clusters in the optimal representation of the participants. The optimal representations of the participants are spread out throughout the continuum of minimal to maximal abstraction. In other words, several participants used a prototype representation in both categories, while others used a

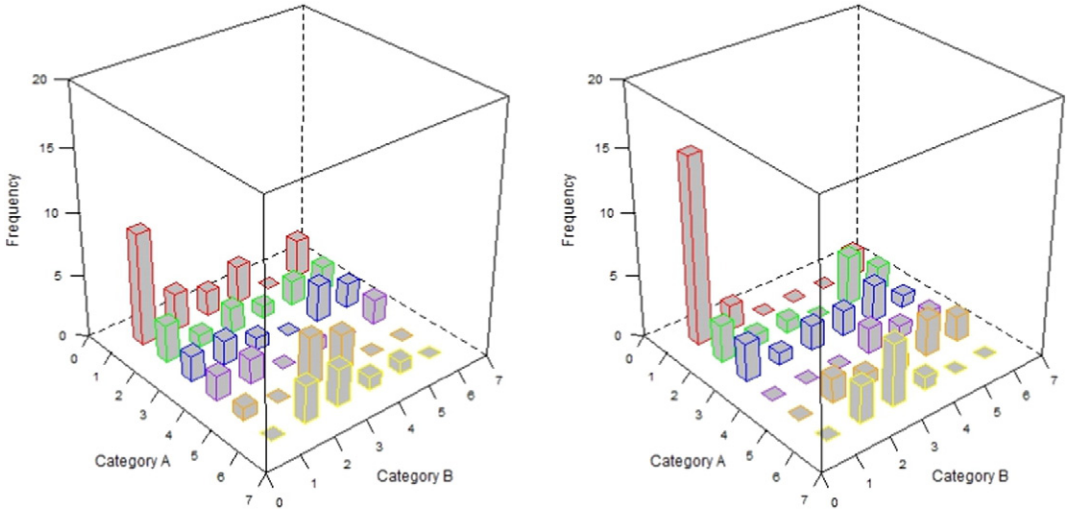

Fig. 5. Best-fitting representations across participants for the stimuli of Fig. 2 (left) and Fig. 3 (right).

representation closer to the minimal abstraction side of the continuum (i.e., an exemplar representation): for example, five clusters in the representation of both categories A and B. The average number of clusters for the stimuli in the upper half of Fig. 2 was 2.94 (standard deviation: 1.79) and the average number of clusters for the stimuli in the lower half of Fig. 2 was 3.05 (standard deviation: 1.75). For Fig. 3 the average number of clusters for the stimuli in the upper half was 3.0 (standard deviation: 1.82) and the average number of clusters for the stimuli in the lower half was 3.02 (standard deviation: 1.85).

To answer the central question that motivated this study, we checked whether the individual differences were stable across time by calculating for each session an abstraction score per participant. This was done by computing the average number of clusters used to represent both categories. The significant correlation, $r = 0.23$, $t(62) = 1.8988$, $p = 0.03$ (one tailed), between the abstraction scores of the two sessions revealed that the individual differences showed stability over time. Participants who used an abstraction strategy in the first session thus were likely to use a similar strategy in the second session, whereas participants who used a memorization strategy in the first session also tended to use a memorization strategy in the second session.

In addition to the correlation between abstraction scores, we also correlated the values of the w and c parameters to check how their values are related across sessions. The correlation between the values of the w parameter across the two sessions was −0.07 and was, as expected, not significantly different from zero ($t(62) = -0.5218$, $p = 0.60$, two tailed). The correlation between the values of the c parameter was 0.25 and, in line with our expectations, was significant ($t(62) = 2.0192$, $p = 0.02$, one tailed).

2.3. Discussion

The first experiment revealed differences between participants in the number of clusters of their optimal representation for both the stimuli of Fig. 2 and Fig. 3. The significant correlation between the abstraction scores of the first session and the abstraction scores of the second session suggests that the individual differences showed stability across time. Although significant, the correlation was quite modest. The relation, however, was tested quite strictly. In both sessions different stimuli, different category structures and a different presentation order were used for each participant. Although we believe that individuals differ in their tendency to use more or less abstraction to represent a category, we also believe that an effective categorizer would be able to adjust her strategy based on situational demands. Hence these situational factors probably temper the relation between the abstraction scores of the two sessions. For example, a participant receiving a category structure in which the training stimuli were situated close to each other and thus were highly similar on the first occasion, is probably influenced by the category structure to use more abstraction. If in the second session, this participant was presented with training stimuli that were situated far apart, the participant might be influenced to use less abstraction.

In addition to the significant correlation between the abstraction scores, we also found a significant correlation between the values of the c parameter across sessions. This correlation suggests that the degree to which individuals are sensitive to differences between stimuli at least partly depends on a stable characteristic of the individual.

Given these findings, the crucial question is what can explain the differences in abstraction strategy. The second experiment sought to answer this question by linking individual differences in abstraction to working memory capacity.

3. Study 2: individual differences and their relation with working memory

The working memory system is used to actively retain and manipulate information in conscious awareness during cognitive tasks (Miyake & Shah, 1999) and can therefore be expected to, at least partly, guide abstraction in category learning. The capacity of the working memory system can be measured by determining the number of items that can be maintained simultaneously in working memory. This number is known to be limited to a few items, although large individual differences in working memory capacity exist (Oberauer, 2005). Sewell and Lewandowsky (2012) found a relationship between working memory capacity and the extent to which people are able to coordinate different types of partial-knowledge structures. Given that the VAM requires the coordination of multiple knowledge structures (in this case subprototypes or clusters), we expect high working memory individuals to lie on the minimal abstraction side of the continuum, using more clusters to represent the category, while the lower working memory capacity individuals are expected to lie on the maximal side of the continuum (using only a small number of clusters to represent the category). In addition, we expect no relation between the values of the c parameter, the w parameter and working memory capacity, since there is no reason to assume that working memory capacity has an influence on the degree to which participants are sensitive to differences between stimuli or the degree to which attention is distributed across the two stimulus dimensions. We think that working memory capacity relates to the ability of participants to coordinate multiple knowledge elements but that it has no influence on the degree to which participants are sensitive to differences or the way participants distribute their attention across stimulus dimensions.

The raw and processed data are available on osf.io/uynbf/.

3.1. Method

3.1.1. Participants

A total of 43 participants completed the experiment. Three of those did not reach the learning criterion (same criterion as in Study 1) and one participant reported being disturbed by loud noises in a nearby room during the working memory task. The data of the remaining 10 men and 29 women (mean age: 20.33) were used in further analyses.

3.1.2. Working memory

To measure working memory capacity of our participants, the GOSPAN task (De Neys, d'Ydewalle, Schaeken, & Vos, 2002) was administered. The GOSPAN is a Dutch and computerized version of the operation span task (OSPAN; La Pointe & Engle, 1990). In the GOSPAN, participants are asked to judge the accuracy of simple math problems (66 operations taken from Engle, Tuholski, Laughlin, & Conway, 1999) while remembering 15 different sets of two-syllable dutch words (varying between two to six words per set). De Neys et al. showed that the GOSPAN scores were reliable and related to the scores on the OSPAN task. Furthermore, it has been shown that the GOSPAN scores, like other working memory scores, are related to performance in conditional reasoning and semantic memory retrieval tasks (De Neys, Schaeken, & d'Ydewalle, 2005; Verschueren, De Neys, Schaeken, & d'Ydewalle, 2002).

3.1.3. Categorization task

The stimuli that were used in this study were the stimuli from Fig. 2. The two categories were formed in exactly the same way as in Study 1. The procedure was exactly the same as in the first study, except that the two categories were called *Mapapilios* and *Machaons*.

3.2. Data analysis and results

3.2.1. Working memory

A participant only received a score for a specific set if all the words in the set were recalled in the order in which they were presented. The score for each set equaled the number of words the set contained. Summing the scores of the sets that were recalled correctly resulted in the working memory score of the participant. The average working memory

capacity score was 39.51 (standard deviation: 11.42) with a minimum of 14 and a maximum of 60.

3.2.2. Learning phase

The 39 participants needed on average 8.2 blocks (standard deviation: 6.17) to reach the learning criterion.[3] The correlation between the number of learning blocks needed and working memory capacity indicated that there was no relation between working memory capacity and speed of learning ($r(37) = -0.18$, $p = 0.138$). Including the three participants that did not reached the learning criteria increased the correlation from −0.18 to −0.22, but this correlation also did not reach significance ($p = 0.08$).

3.2.3. Transfer phase

3.2.3.1. Fits. To assess the degree of abstraction, the VAM was fitted to the data of the transfer phase using a grid-search routine. The data of every participant were fitted individually and the optimal representation was identified by maximum likelihood. Fig. 6 gives an overview of the distribution of R-squared and the values of the c and w parameters across participants. The model fitted the data reasonably well, with an average R-squared across participants of 0.82.

3.2.3.2. The relation between representational abstraction and working memory capacity. Fig. 7 gives an overview of the number of times a specific representation was used for the stimuli of the upper half of Fig. 2 (Category A) and the stimuli of the lower half of Fig. 2 (Category B) across participants. The number of clusters for Category A is shown on the X-axis, the number of clusters for Category B on the Y-axis. The colored bars indicate the frequency with which a specific combination of clusters was identified as the optimal representation. Similarly to the first study, there was considerable variation in the number of clusters in the optimal representations of the participants. Several participants, for example, used a prototype representation in both categories, while others used a representation closer to the minimal abstraction side of the continuum (e.g., five clusters in the representation of both Categories A and B). The average number of clusters for Category A was 2.72 (standard deviation: 1.78) and the average number of clusters for Category B was 2.97 (standard deviation: 1.75).

To check whether the level of abstraction in the participants' category representations was related to their working memory capacity, we calculated an abstraction score for every participant by averaging the number of clusters of the two categories in their optimal representation. The resulting scores were then correlated with the working memory capacity scores of the participants. The positive and significant correlation, $r = 0.32$, $t(37) = 2.0558$, $p = 0.02$ (one-tailed), indicated that, in general, participants with lower working memory capacity used less clusters to represent the category, while participants with higher working memory capacity used more clusters. Thus, the lower the working memory capacity of the participants, the more abstraction used to represent the category.

The correlation between working memory capacity and the c parameter did not reach significance, $r = 0.15$, $t(37) = 0.92$, $p = 0.36$ (two-tailed). This was in line with what we expected: working memory capacity does not influence the degree to which individuals are sensitive to differences between stimuli. Similarly, the correlation between the w parameter and working memory capacity also did not reach significance, $r = -0.15$, $t(37) = -0.94$,$p = 0.35$ (two-tailed), which was again in line with our expectations.

[3] The working memory scores of the three participants that did not reach the learning criterion were close to the average working memory score, respectively, 37,34 and 31. The percentage correct in the final block for participants that did not reach the learning criterion varied between 0.75 and 0.83.

4. General discussion

The starting point of the current study was to investigate whether individuals mutually differ in the extent to which they use abstraction processes to learn a category. We wanted to unravel individual differences in the degree of abstraction in one's category representation. First, we wanted to show that these differences reflect reliable underlying individual tendencies. Second, we wanted to investigate whether these tendencies were related to working memory capacity and thus whether the capacity of an individual's working memory can be linked to the strategy they use to learn categories. Both our experiments revealed clear individual differences in the type of strategy the participants used to learn the categories. Some participants clearly relied on an abstraction strategy to learn categories, where others were somewhere in between and still others preferred a strategy that is predominantly determined by memorization. The significant correlation in the abstraction scores between the two sessions suggests that these differences were stable across time. The significant correlation between the values of the c parameter in the two sessions suggested that the degree to which individuals are sensitive to differences between stimuli is not only related to the characteristics of the stimuli or the experience with the stimuli, but also depends upon an cognitive characteristic of the individual.

We showed in our second experiment that the individual differences in abstraction were related to working memory capacity: participants with lower working memory capacity were more likely to use abstraction processes when learning the categories than participants with high working memory capacity. In the following sections these results and their relation with previous studies will be discussed in more detail.

4.1. Understanding individual differences

The observation that individuals differ in the degree to which they make abstraction across category members and thus that the type of categorization strategy depends at least partly on stable individual differences has some implications for category learning experiments and the artificial category learning literature in general. With a few exceptions (e.g., Medin et al., 1984) the artificial category learning literature was decennia long dominated by its most prominent discussion: whether individuals make some sort of abstraction when learning a category (as formalized in rule-based and prototype models) or whether they use a memorization strategy (as assumed in exemplar models) (Nosofsky, 1986, 1987, 1989; Nosofsky, 1989; Nosofsky & Palmeri, 1997, 1998;). A tacit assumption underlying this discussion was the belief that, regardless of the exact circumstances in which people are categorizing and independent of the characteristics of the categorizer, everyone always uses the same categorization strategy. This is, for example, evident from the common approach in these studies, in which categorization data are aggregated across participants and the categorization model is then fitted to the aggregated data. Our study adds to the growing literature demonstrating that this assumption is questionable. An immediate implication is that instead of trying to unravel which strategy each and every participant uses, it is worthwhile to investigate the characteristics of the individuals that determine their preferred categorization strategies. Furthermore, our results show that individual categorization strategies can even be measured at the level of varying degrees of abstraction, not only in contrasting exemplar and prototype models (Vanpaemel & Storms, 2010).

In our second study, we showed that individual differences, as observed in Study 1, are significantly related to working memory capacity. A goal for future studies is to determine other factors that influence categorization strategies in general, and abstraction processes in specific.

The literature suggests an influence of other variables. McDaniel et al. (2014) investigated (besides working memory capacity) also fluid intelligence as measured by Raven's Advanced Progressive Matrices (RAPM), which gives an indication of the individual's ability to

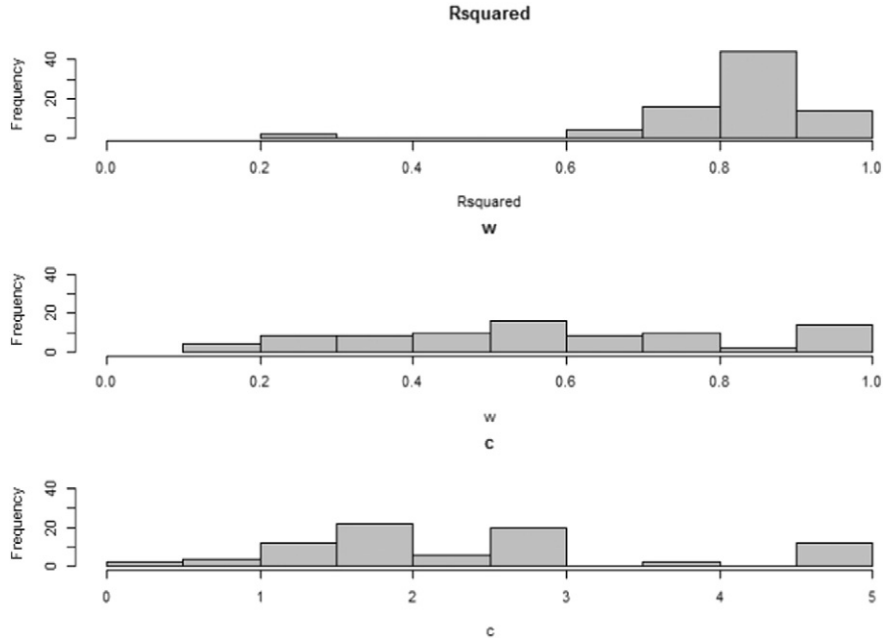

Fig. 6. The distributions of R-squared and the values of the w and c parameter across the best fitting representations of all participants.

extrapolate a rule from a group of stimuli, and found that the higher the RAPM scores, the more likely the individual was a rule learner. Another possible factor that might determine whether participants will use abstraction to learn the category is the extent to which people have a detail-focused cognitive style. People with autism, for example, have the tendency to focus on the details instead of on the whole, they are biased to local features and have difficulties noticing the global pattern. Evidence for this comes from their superior performance in tasks that require local processing (for example, the embedded figures task; Jollifle & Baron-Cohen, 1997) and their decreased performance in tasks that require global processing (for example, a decreased vulnerability to visual illusions; Happé, 1996). One of the consequences of this detail-focused cognitive style is that people with autism might have difficulties making a prototype representation of a category (Klinger & Dawson, 2001). The extent to which a detail-focused cognitive style has an influence on categorization strategy in a normal population remains to be investigated.

Investigating individual differences and the factors that influence these differences is also important given the occurrence of these strategies in educational settings (McDaniel, Frey, Kudelka, & Shields, 2012). Identifying which student is likely to use a given strategy and assessing whether this strategy is effective or not, given a specific course goal, can aid teachers in coaching students towards reaching these goals (Little & McDaniel, 2015b).

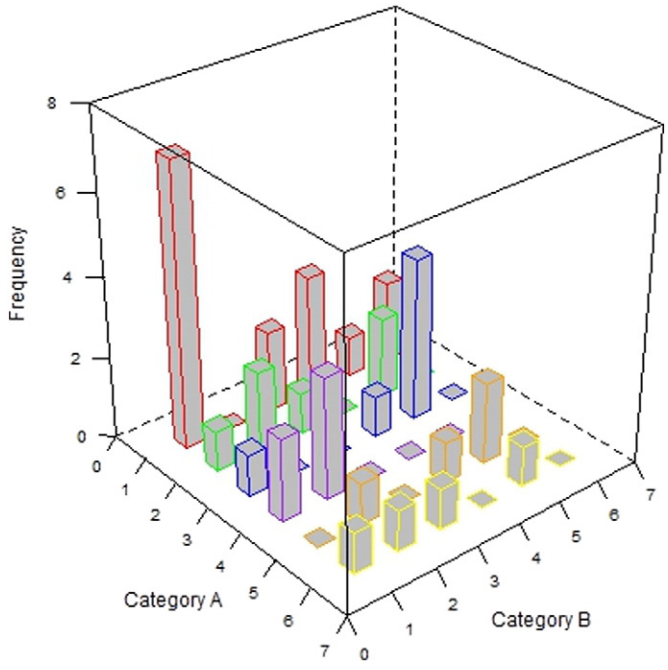

Fig. 7. Best–fitting representations across participants for the second study.

4.2. Working memory

We found evidence for a relationship between working memory capacity and the degree to which participants made abstraction across the category members in our experiment: Higher working memory capacity individuals tended to make less abstraction across category members than individuals with lower working memory capacity. This was in line with our hypothesis: we expected that individuals with higher working memory capacity would be better able to actively coordinate multiple knowledge elements in working memory while performing a categorization task and thus would feel a smaller need to make abstraction across category members than individuals with lower working memory capacity. Our results differ from the results reported by McDaniel et al. (2014) who found the reversed relationship between working memory capacity and abstraction: Their participants with high working memory capacity were more likely to use a rule-based strategy than an exemplar-based memorization strategy. McDaniel et al., however, used a function-learning task in their experiment in which participants had to learn to pair continuous cue values with continuous criterion values following a bilinear function. The way to solve this task is probably very different from the categorization task in our study, as is also evidenced by the fact that the representational shifts in this type of tasks follow the opposite direction (memorization to rules) of the representational shifts (rules/prototype to exemplar) that are typically reported in tasks that resemble the one employed in our experiment (Johansen & Palmeri, 2002; Smith & Minda, 1998; Vanpaemel & Navarro, 2007). In addition, in our type study, making abstraction across the different category members was operationalized by averaging across the category members while in the task of McDaniel et al. making abstraction

corresponded to formulating a rule. Averaging across category members is probably different from formulating a rule in a categorization task: When making abstraction through averaging, only the average (i.e., the summary representation) needs to be retained in working memory. When formulating a rule probably several aspects of the rule might also need to be actively retained in working memory. These differences might also explain the difference in results. Our results also differ from Little and McDaniel (2015a) and Craig and Lewandowsky (2012). We did find a relationship between working memory capacity and strategy use, but, whereas they focused on only the two endpoints of the continuum (full abstraction or no abstraction) we used a more detailed measurement, which – arguably – leads to a more sensitive measure of abstraction. Sewell and Lewandowsky (2012) reported that participants with higher working memory capacity were better at shifting between response strategies and argued they are better able to coordinate multiple elements during category learning. Our results confirm this idea: High working memory capacity participants made less abstraction across category members than participants with lower working memory capacity.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*, 409–429.

Arguin, M., & Saumier, D. (2000). Conjunction and linear non-separability effects in shape encoding. *Vision Research*, *40*, 3099–3115.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33–53.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372–400.

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

Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *Quarterly Journal of Experimental Psychology*, *65*, 439–464.

De Neys, W., d'Ydewalle, G., Schaeken, W., & Vos, G. (2002). A Dutch, computerized, and group administrable adaptation of the operation span test. *Psychologica Belgica*, *42*, 177–190.

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking and Reasoning*, *11*(4), 349–381.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable 16 approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

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

Happé, F. G. E. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions, a research note. *Journal of Child Psychology and Psychiatry*, *37*, 873–877.

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

Jollifle, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures test? *Journal of Child Psychology and Psychiatry*, *38*, 527–534.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*, 111–124.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1118–1133.

Little, J. L., & McDaniel, M. A. (2015a). Individual differences in category learning: Memorization versus rule abstraction. *Memory and Cognition*, *43*, 283–297.

Little, J. L., & McDaniel, M. A. (2015b). Some learners abstract, others memorize examples: Implications for education. *Translational Issues in Psychological Science*, *1*, 158–169.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley Publishing Company.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, *11*, 309–332.

McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General*, *143*, 668–693.

McDaniel, M. A., Frey, R., Kudelka, C., & Shields, S. P. (2012). Individual differences in concept learning tendencies: Spanning the laboratory and the classroom. *Paper presented at the POGIL South Central Regional Meeting, St. Louis, MO*.

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 333–352.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 104–114.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87–108.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, *45*, 279–290.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Nosofsky, R. M., & Palmeri, T. J. (1998). A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psychonomic Bulletin & Review*, *5*(3), 345–369.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 924–940.

Nosofsky, R. M., Palmeri, T. J., & Mckinley, S. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53–79.

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*, 368–387.

Op De Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent lowdimensional configurations of parameterized shapes. *Nature Neuroscience*, *4*, 1244–1252.

Op De Beeck, H., Wagemans, J., & Vogels, R. (2003). The effect of category learning on the representation of shape: Dimensions can be biased but not differentiated. *Journal of Experimental Psychology: General*, *132*, 491–511.

Raijmakers, M. E. J., Schmittmann, V. D., & Visser, I. (2014). Costs and benefits of automatization in category learning of ill-defined rules. *Cognitive Psychology*, *69*, 1–24.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.

Sewell, D. K., & Lewandowsky, S. (2012). Attention and working memory capacity: Insights from blocking, highlighting and knowledge restructuring. *Journal of Experimental Psychology: General*, *141*, 444–469.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1411–1436.

Smith, D. J., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 3–27.

Smith, J. D., Murray, M. J., & Minda, J. P. (1997). Straight talk about linear separability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 659–680.

Stukken, L., Storms, G., & Vanpaemel, W. (2011). Explaining categorization response times with varying abstraction. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2842–2847). Austin, TX: Cognitive Science Society.

Vanpaemel, W., & Navarro, D. (2007). Representational shifts during category learning. In D. McNamara, & J. Trafton (Eds.), *Proceedings of the annual conference of the cognitive science society* (pp. 1599–1604). Austin, TX: Lawrence Erlbaum Associates.

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.

Verbeemen, T., Vanpaemel, W., Pattyn, S., Storms, G., & Verguts, T. (2007). Beyond exemplars and prototypes as memory representations of natural concepts: A clustering approach. *Journal of Memory and Language*, *56*, 537–554.

Verschueren, N., De Neys, W., Schaeken, W., & d'Ydewalle, G. (2002). Working memory capacity and the nature of generated counterexamples. *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (pp. 914–919). Mahwah, NJ: Erlbaum.