Preprint; please don't quote

The Influence of Presentation Order on Category Transfer

Fabien Mathy
Université Nice Sophia Antipolis–France

Jacob Feldman
Rutgers University – New Brunswick, USA

This study of supervised categorization shows how different kinds of category representations are influenced by the order in which training examples are presented. We used the well-studied 5-4 category structure of Medin & Schaffer (1978), which allows transfer of category learning to new stimuli to be discriminated as a function of rule-based or similarity-based category knowledge. In the rule-based training condition (thought to facilitate the learning of abstract logical rules and hypothesized to produce rule-based classification), items were grouped by sub-categories and randomized within each subcategory. In the similarity-based training condition (thought to facilitate associative learning and hypothesized to produce exemplar classification), transitions between items within the same category were determined by their featural similarity and sub-categories were ignored. We found that transfer patterns depended on whether the presentation order was *similarity-based*, or *rule-based*, with the participants particularly capitalizing on the rule-based order.

A few studies have shown that aspects of stimulus presentation can influence category formation (Carvalho & Goldstone, 2011, 2014; Clapper & Bower, 1994; Elio & Anderson, 1981, 1984; Gagné, 1950; Goldstone, 1996; Kornell & Bjork, 2008; Kang & Pashler, 2012; Mathy & Feldman, 2009; Medin & Bettger, 1994). One way to test this is to evaluate the *transfer* of learning to novel instances, in order to see whether different orders can, in fact, alter the way training examples are generalized (Elio & Anderson, 1981, 1984). In this paper we pursue this strategy, using the popular 5-4 category structure of Medin and Schaffer (1978) for which generalization patterns have been scrutinized in the past (Johansen & Palmeri, 2002) to test whether rules or similarity can alter the representation of categories.

Rule-based vs. similarity-based processing is often induced exclusively by the characteristics of the categorization tasks (Ashby & Ell, 2001), but sometimes both types of processes can be observed within a single task (e.g., Allen & Brooks, 1991; Shanks & Darby, 1998), for instance depending on the presence of a concurrent load during training (Wills, Graham, Koh, McLaren, & Rolland, 2011). The present study adopts a different strategy and seeks to manipulate these two types of processing as factors within a single task. Although the rule-based and similarity-based distinction is subject to theoretical difficulties, there is some advantage of having these two explanatory constructs under different experimental conditions (see the special issue of *Cognition* in Vol. 65, e.g., Hahn & Chater, 1998; Sloman & Rips, 1998; E. E. Smith, Patalano, & Jonides, 1998).

Mathy and Feldman (2009) for instance tested three types of presentation orders to study category discovery (not category generalization, as in the present study). The rule-based order, in which stimuli obeying different rules are separated to facilitate a rule abstraction process, led to the most effective learning overall. Two other types of presentation order were found less beneficial to category learning, although as discussed below they may have subtle benefits for other kinds of learning. The similarity-based order (Elio & Anderson, 1981), in which the stimuli are ordered by proximity in the stimulus space, facilitates exemplar memorization. (The temporal contiguity of the most similar stimuli could reinforce the local associations of exemplars, but it could also induce overly specific rules). The dissimilarity-based order, in which the stimuli are ordered so as to maximize distance between successive items, tends to disrupt almost any type of learning mechanism (Gagné, 1950, found in this latter case no difference with a random presentation). The present study employs the first two types of orders, because both can easily be matched to a specific learning mechanism (respectively, rule-based and exemplar-based) that we sought to track during the generalization of category learning to new stimuli.

We hypothesized that participants in a rule-based training condition would show generalization patterns consistent with rule-based retrieval and that participants in a similarity-based training condition would show generalization patterns consistent with exemplar retrieval. We provide strong evidence that participants capitalized differently on the two types of stimulus order from the analysis of the individual generalization patterns. To anticipate further, the generalization patterns seem more typical of a rule-based classification following a rule-based training, and the connection is weaker for the similarity-based condition.

This research was supported by the Agence Nationale pour la Recherche (ANR) Grant # ANR-09-JCJC-0131-01 awarded to Fabien Mathy. We are grateful to Azizedine Elmahdi and Nicolas Heller for their assistance in data analysis during their engineer internship in the lab. Correspondence concerning this article should be addressed to BCL Lab, CNRS, UMR 7320, Nice, Université Nice Sophia Antipolis, France, or by e-mail to fabien.mathy@unice.fr.

Method

The present experiment used the 5-4 categories (see Fig. 1 and Table 1). The 5-4 category set was first studied by Medin and Schaffer (1978), as well as in many subsequent studies reanalyzed by J. D. Smith and Minda (2000), and in other subsequent studies (Cohen & Nosofsky, 2003; Johansen & Kruschke, 2005; Johansen & Palmeri, 2002; Lamberts, 2000; Lafond, Lacouture, & Mineau, 2007; Minda & Smith, 2002; Rehder & Hoffman, 2005; Zaki, Nosofsky, Stanton, & Cohen, 2003). This structure makes it possible to study the way in which 7 unclassified items are categorized during a transfer phase, after the learning of 5 + 4 = 9 items. The 5-4 ill-defined category structure has been evaluated many times without diminishing its current importance. For instance, Raijmakers, Schmittmann, and Visser (2014) have recently suggested that prolonged learning of this category can be modeled using latent Markov models, with the aim of detecting the underlying strategies used by participants and the transitions between those strategies. Our objective here was different in that our manipulation was designed to induce a particular strategy (rule-based or exemplar) at the beginning of the learning process, and to find out if different generalization patterns could be observed depending on whether training was similarity- or rule-based right after a short training process (i.e., an amount of training sufficient to acquire the correct categories). We based our analysis on the list of the rule-based and exemplar-based patterns of generalization that was provided by Johansen and Palmeri (2002) for the 5-4 categories (but see Raijmakers et al., 2014 for a more recent and more complete exploration of the possible fluctuation of the strategies used by participants during transfer, including guessing).

Two types of presentation order was used (rule-based or similarity-based) for the training phase, and this manipulation was a between-participants factor. In both conditions, a single fully-blocked (i.e., separating the two categories into two distinct sub-blocks; see Clapper & Bower, 2002) ordering of the stimuli was generated for each participant and then used for all training blocks[1]. Blocking was used to reinforce the effect of ordering the stimuli within category. After the category was learned (up to a chosen criterion explained below), the participants were administered a transfer phase in which old and new stimuli were shown randomly.

Participants. The participants were 44 freshman or sophomore students from the University of Franche-Comté, who received course credits in exchange for their participation.

Choice of categories studied. Each participant was administered a single 5-4 category set. The 5-4 is shown in the bottom hypercube of Fig. 1. In the hypercube, made of $2^4 = 16$ stimuli, the 5 examples of category A are indicated by black circles and the 4 examples of category B are indicated by empty vertices. The hypercube, also known as a Hasse diagram, is extremely useful for quickly looking at the category structure, which does not easily appear in the corresponding truth table such as the one provided in Table 1.

Stimuli. Stimulus objects varied across four Boolean dimensions (Color, Shape, Size, and Filling texture). Each dimension was instantiated by the same physical dimension for all participants. This choice was made to homogenize attention to the dimensions across participants to limit the distribution of the generalization patterns within presentation orders. A color dimension differentiated the objects at the top of the hypercube from those at the bottom (green *vs*. red respectively); a shape dimension differentiated the objects at the front from those at the back (square *vs*. circle); a size dimension distinguished the objects in the left cube from those in the right cube (small vs. large); and finally, the left and right objects within the cubes were hatched *vs*. plain. Overall, the combination of these four separable dimensions (Garner, 1974) formed 16 single unified objects (e.g., a small hatched red square, a large plain green circle, etc.).

Clusters. The 5-4 notation refers to the presence of one 5-member category and another 4-member category (Medin & Schaffer, 1978; J. D. Smith & Minda, 2000). For this concept, the objects of each category were numbered by virtually indexing each of the training stimuli using *A* (first category), *B* (second category), and *T* (transfer items). This notation refers to previous research (e.g., Smith & Minda, 2000, p. 4).

To identify the items in a rule-based fashion, we rely on the simplest *one-dimensional* ***Color*** *rule plus exception* which groups the objects of category A into two mutually exclusive sub-categories or clusters[2]. Category A was organized into two clusters: the largest cluster represented the rule for the Red feature and the smallest cluster represented the exception for the Green feature, that is, Cluster 1 = (A_5,

[1] Although many studies have shown that induction benefits from interleaving the categories, (e.g., Kornell & Bjork, 2008), blocked presentations are sometimes more beneficial (Goldstone, 1996), for instance with low-similarity category structures such as the one used in the present study (Carvalho & Goldstone, 2014). In this case, blocking tends to promote the identification of the shared features (a process that seems neutral regarding the rule-based vs. similarity-based processes). See also Clapper and Bower (1994, 2002) for a positive effect of blocking. Another strategy for presenting the stimuli (such as presenting first the most different stimuli close together in temporal sequences to highlight the discriminative features of the categories) would argue for the use of interleaving instead.

[2] There are two dimensions in these stimuli that admit a rule-plus-exception strategy (i.e. Size and Color), but Color was chosen to facilitate the task.

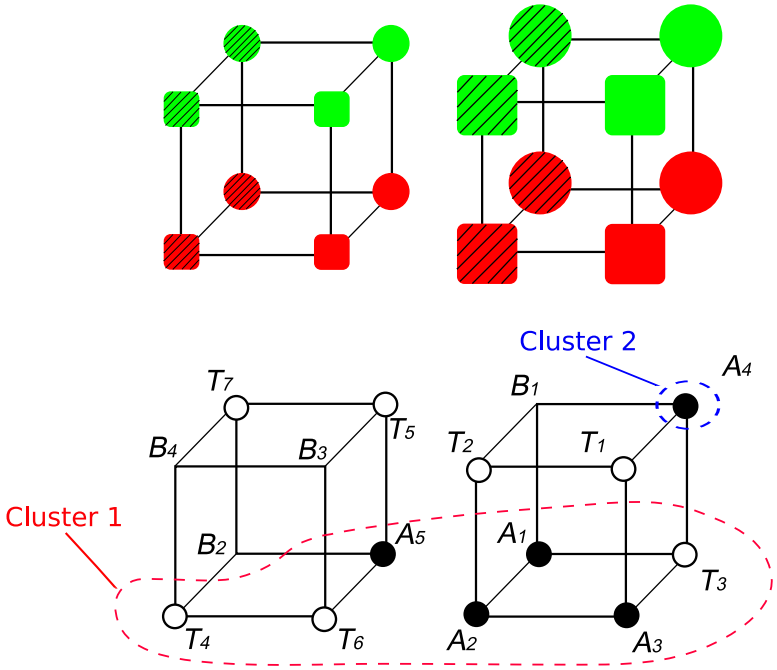


Figure 1. Four-dimensional Hasse diagram, training sample, and the 5-4 categories. *Note.* The figure illustrates two hypercubes, one per row. The first hypercube exemplifies how the dimensions were implemented in the experiment. The second hypercube shows the structure of the 5-4 categories. The examples of category A are indicated by black circles, whereas the examples of category B are represented by empty vertices. The white circles represent the transfer items. The 5-4 notation refers to the presence of 5 examples in category A *vs*. 4 examples in category B (Medin & Schaffer, 1978; Smith & Minda, 2000). For this concept, following Smith & Minda, 2000 (p. 4), the objects were numbered by indexing each of the examples using letters (*A*, for the examples belonging to category *A*; *B*, for the second category; and *T* for the examples that were presented during the transfer phase, which the participants could freely classify as *A*'s or *B*'s. The numbers following the letters *A*, *B*, and *T* are arbitrary numbers that distinguish the individual items. Note that T_1,..., T_7 in Johansen & Palmeri (2002), and in the present study, are respectively indexed T_{10},..., T_{16} in Smith & Minda (2000). Among the examples of category A, there are two clusters. The red set indicates the objects of category A belonging to the biggest cluster (T_4, T_6, and T_3 are virtually included in the cluster under the hypothesis that the participant in the rule-based order is induced to form a "All red except the small hatched circle" rule). The blue set indicates the objects of category A from the second cluster. Category-B clusters (described in the text) are not represented in order to lighten the presentation. The clusters are circled by discontinuous curves to acknowledge the fuzziness of their boundaries. Indeed, the representation of the clusters in the participant's mind is unclear because seven examples (the transfer items) are not associated with clear categories during the training phase.

A_1, A_2, A_3), and Cluster 2 = A_4. The two combined clusters thus represented the rule "All red, except the small hatched circle" and the "large plain green circle" exception. These two clusters are circled by discontinuous red and blue curves in Fig. 1. Similarly, the objects of the second category were B_3, B_4, and B_1 (Cluster 1) and B_2 (Cluster 2), representing the rule "All green, except the large plain circle" and the "small hatched red circle" exception[3].

Ordering of stimuli. For both conditions (Rule-based and Similarity-based), the stimuli were fully separated/blocked according to category (Category *A* and Category *B*). The presentation of Category *A* always preceded the presentation of Category *B* in two successive sub-blocks.

[3] Note that the choice of clusters is hypothetical. We refer here to the idea that these clusters can result from the participants' conceptualization of the induced simplest one-dimensional rule plus exception during the training phase, and can be applied to the transfer phase. However, other participants' conceptualizations are possible. For instance, a more complicated alternative representation for Cluster 1 could be the set A_5, T_6, A_1, A_2, T_3, and A_3 (i.e., all red, except small hatched). This rule is simpler if we refer to its length but it is more elaborated in that it supposes that the participant makes the abstraction that B_2 is not the only exception in play. The only effective constraint on the ordering of stimuli was the separation between the two clusters within-categories, in order to help the abstraction of the one-dimensional rule plus exception allowing one

Table 1
Category structure of the 5-4 categories and observed generalizations patterns by order type.

A		B		Transfer		Observed Generalization patterns + Frequency (bottom)												
A1	1110	B1	1100	*T1	1001	A	A	A	A	A	A	A	B	B	B	B	B	B
A2	1010	B2	0110	*T2	1000	A	A	A	A	B	B	B	A	A	B	B	B	B
A3	1011	B3	0001	T3	1111													
A4	1101	B4	0000	*T4	0010	A	B	B	B	A	B	B	A	B	A	B	B	B
A5	0111			*T5	0101	A	A	B	B	B	A	A	B	A	B	A	A	B
				*T6	0011	A	A	A	B	A	A	B	B	A	A	A	B	B
				T7	0100													
						Frequency												
				Order:	*Sim.*	4	2	1	2	1	1	1	0	0	7	0	1	1
				Order:	*Rule*	2	0	0	1	0	1	0	1	1	14	1	1	0

Note. * highlights the critical transfer stimuli. A and B stand for 'Category A' and 'Category B' respectively. The feature vectors such as 1110 in the second column represents the Size-Shape-Color-Filling dimension values respectively. The values 1111 represent respectively the features Large-Red-Circle-Plain. The values 0000 represent respectively the features Small-Green-Square-Hatched. Among the 32 possible generalization patterns, we observed only 13 different ones in our experiment. The patterns are reported in columns in the right part of the table.

In the rule-based order, the first objects of the first category (i.e., category A) were randomly drawn from Cluster 1 until all had been presented. This was followed by the object in Cluster 2. Thus, in the rule-based order, all the members of the biggest cluster were presented first in a random order, but separated from the exceptional member, in order to encourage participants to abstract the simplest rule. The presentation within Cluster 1 was random in accordance with a rule-abstraction process that is supposed to impede stimulus singularity. The ordering procedure was similar for the second category (i.e., category B): the objects of Cluster 1 (randomly ordered) preceded the object of Cluster 2.

In the similarity-based order, the first object of category A was chosen at random, and subsequent objects of the same category were chosen randomly from those with maximal similarity to the previous object, and so forth until the set of examples was exhausted. Ties were resolved randomly. The same algorithm was applied to the second category (i.e., category B). From an exemplar point of view (Elio & Anderson, 1981, 1984), this type of order strictly followed the similarity structure between the stimuli with the objective to reinforce exemplar memorization.

Dissimilarity between two stimuli i and j was computed using the city-block distance

$$d_{ij} = \sum_{a=1}^{a=4} |x_{ia} - x_{ja}| \qquad (1)$$

where x_{ia} is the value of stimulus i along dimension a. Similarity was simply computed using $s_{ij} = 4 - d_{ij}$. The most important aspect of this procedure was that the ordering did not necessarily respect the cluster boundaries targeted in the rule-based order, as similarity steps could cross in and out of clusters (for instance, A_5, A_4, A_1, etc.). Another important aspect of the similarity-based order is that some orders could induce hypotheses more specific than those targeted in the rule-based order: for instance, "large red squares" after the presentation of A_3 and A_2; "large red squares OR large red hatched", or "red large, except plain circle", after the presentation of A_3, A_2, and A_1; "large red squares OR large circles" after the presentation of A_3, A_2, A_1, and A_4, etc.

One sample of a rule-based presentation for Category A followed by Category B could be: A_1, A_5, A_3, A_2), and A_4, followed by B_3, B_4, B_1, and B_2. One sample of a similarity-based presentation for Category A followed by Category B could be: A_2, A_1, A_5, A_4), and A_3, followed by B_1, B_2, B_3, and B_4.

Justification for the types of presentation order chosen. These two types of presentation order match two extreme ways of learning: a complex inductive process based on abstraction and an elementary process with underlying associative mechanisms (Sloman, 1996).

to reach the learning criterion faster according to our hypothesis.

1. The rule-based condition uses a set of clusters which are presented to the participants in an order depending on their magnitude (since in many domains, exceptions are learned last), with no distinction within clusters (since the abstraction process supposedly cancels out any effects of the non-diagnostic features on learning). Because the objects are supposed to involve common abstract properties within clusters, they are drawn randomly. The goal is to make participants form a logical rule. Overall, the rule-based order is thought to help the learner separate different clusters in order to abstract a simple logic describing the stimuli. The identification of the clusters is also facilitated by the randomization of the steps within clusters. The randomization of the steps avoid the learner to be mislead by a too specific hypothesis.

2. The similarity-based condition follows a more simple hypothetical associative process that uses the temporal contiguity of the stimuli to reinforce the memory traces of two stimuli locally and by extension, by a chaining process, the entire similarity structure. Because exemplar models assume that participants classify objects according to their relative summed similarity to exemplars of the two categories, it is hypothesized that exemplar memorization can capitalize on a reinforced similarity structure.

To compare the two conditions, imagine the learner has already encountered a small and green object (e.g., B_3), which gave the learner some information about the different dimension values. In this context, the presentation of a large plain red square A_3 followed by a large hatched red circle A_1 directly rules out the overly specific "large red square" rule, in favor of the more general "large and red" rule. On the contrary, a large plain red square A_3 followed by a large hatched red square A_2 (an order typically favored by the similarity-based order) tends to temporarily mislead the learner about the possible "large red square" rule, until the large hatched red circle A_1 is presented. A similarity-based order operates otherwise to favor an exemplar-based learning mechanism: when a large plain red square A_3 is followed by a large hatched red square A_2 , the two exemplars are not only memorized as single entities but the association of the two can be thought to subsequently reinforce the entire similarity structure (see Sloman, 1996, who develop a explanation of why associative learning use similarity and temporal relations rather than symbolic structures).

Procedure. Participants were individually tested. The participants sat approximately 60-cm from the computer screen and were given a tutorial before the task began. Each participant was then asked to learn a single 5-4 category set and was administered a series of training blocks until the learning criterion (detailed later) was met.

Presentation order (rule-based or similarity-based) was a between-subject manipulation during the training phase. The participants were randomly assigned to one of these two conditions. Because there were several possible variations within each presentation order type (depending for instance on which stimulus was first presented within blocks), one presentation order was randomly chosen for a given participant before the experiment started and the chosen presentation order then applied across the training blocks until the learning criterion was met.

Independently of the two types of orders, the training phase was arranged to be as simple as possible. Each participant was administered a series of identical blocks in which, within each block, the five examples of the first category of the to-be-learned 5-4 category structure preceded the four examples of the second category (i.e., learning was massed; see Birnbaum, Kornell, Bjork, & Bjork, 2013; Carvalho & Goldstone, 2014; Kornell & Bjork, 2008; Kornell, Castel, Eich, & Bjork, 2010; Kang & Pashler, 2012; Wahlheim, Dunlosky, & Jacoby, 2011). Although not specifically tested on the 5-4 category structure, Mathy and Feldman (2011) previously showed that massing (relative to interleaving) the stimuli by category (such that the stimuli of each category appeared consecutively) facilitated both a rule-based and a similarity-based binary-valued artificial classification learning task such as the one studied here. The repetition of identical orders across blocks within participants was not the procedure used by Mathy and Feldman (2009) either. However, by removing the variations between blocks during training, our goal was to have the participants form a consistent representation of the category structure.

During a training block, the stimuli were displayed one at a time on the top half of the computer screen for one second. The A and B categories corresponded respectively to the up and down keys, and to two category pictures on the right hand side of the screen. In the category frame, a school bag was located at the top, and a trash can[4] at the bottom (to match the response keys). When the stimulus was presented, the correct category label was displayed below the stimulus (i.e., 'schoolbag' or 'trash can') for one second while the corresponding category picture was displayed (for instance, the school bag was shown for one second, while the trash can was hidden for a positive category). This instruction was followed by a confirmation phase during which the participant had to press the response key corresponding to the category picture that had just been shown to them. After the key was pressed, feedback indicating a correct or incorrect classification was given at the bottom of the screen for two seconds. This feedback was useful despite presenting the correct answer prior to the participant's response because participants

[4] The school bag and the trash can have been used in many other similar experiments on artificial categorization in our lab, and have proved useful for children, especially when giving the instructions. We chose to use the same experimental setup across populations to facilitate the comparison of the results.

in categorization tasks might systematically inverse the response keys at one point or another of the experiment. Our pretests showed that in this condition, the participants could not fail to correctly give any of the instructed responses. Therefore, the participants were expected to get 100% correct feedback during the training blocks. This confirmation phase was used to make sure that the participants were following the learning phase actively and that they did not miss any of the instructed categories.

Because a 100% correct feedback did not guarantee learning during the training blocks (e.g., the participants could repeat the instructed responses without paying attention to the stimuli), each training block was followed by a categorization block to test the participants along the training process. In the categorization blocks, the nine stimuli were randomly permuted to test the participants. Category learning was also supervised (i.e., with feedback) in the categorization blocks. To summarize, the main difference between the training blocks and the categorization blocks is that order was manipulated in the training blocks (which was observational in nature and which did not allow testing the participants) whereas order was random during the categorization blocks to test the participants. Although the categorization blocks could generate noise in the manipulated order, the participants indicated in our pretests that the learning process was clearly facilitated by the recurring training blocks.

To further encourage learning during the categorization blocks, the participants scored one point on a progress bar for each correct response. The number of empty boxes in the progress bar was $4 \times (5 + 4)$. One empty box was filled whenever a correct response was given, but the progress bar was reset if an incorrect response was given. This criterion was identical to the one used by Shepard, Hovland, and Jenkins (1961) in their first experiment and by Mathy and Feldman (2009). Consequently, the participants had to correctly classify stimuli in four consecutive categorization blocks of nine stimuli to complete the training phase of the experiment. This setting required them to correctly classify all the stimuli, including those considered as exceptions. This intentionally limited the number of strategies that could provide partial solutions such as being able to classify stimuli on the basis of a limited number of features with less than 100% accuracy. Although this criterion seems difficult to reach (in Johansen & Palmeri, 2002, only 75% of their participants reach a criterion of 80% accuracy in the final blocks), our procedure combining the observation of the category labels and the fully-blocked procedure was thought to outperform the usual trial-and-error learning procedure. Our participants were not trained to categorize the stimuli during 32 training blocks like in some previous studies (Johansen & Palmeri, 2002; Raijmakers et al., 2014), because we were interested in observing the participants' representations after they had acquired the correct category. Because the training blocks and the categorization blocks were interspersed, the progress bar was hidden during the training blocks. The number of points that were accumulated on the progress bar was restored whenever a categorization phase began.

Once the participants reached the learning criterion (at this point, the progress bar was equal to 4×9), we conducted a transfer phase during which both the training and transfer stimuli were presented (each once in a block). The transfer phase was composed of 5 blocks of 16 stimuli. The order of all stimuli was randomized within-blocks for the transfer phase.

Results

Learning Phase. One participant (in the similarity-based order) who was not able to meet the learning criterion in the time allowed by the experiment did not complete the transfer phase and was removed from the analysis. To make sure that the similarity-based order differed from the rule-based order, inter-item similarity was computed for each type of order and for all contiguous pairs of examples within the manipulated blocks before being averaged by participants. The mean inter-item similarity per block was significantly lower in the rule-based condition (2.05, $sd = .19$) than in the similarity-based condtion (2.2, $sd = .13$), $t(41) = 2.9$, $p = .006$. Learning was faster in the rule-based order than in the similarity-based order (6.4, $sd = 2.6$, vs. 9.0, $sd = 5.7$, number of training blocks to criterion on average), and this difference was significant ($t(41) = -2.39, p = .051$, but $Z_{Wilcoxon} = -2.11, p = .035$; the non-parametric test was preferred because both distributions significantly deviated from a normal distribution using the Shapiro-Wilk test, due to a positive skewness in both groups). However, no correlation between Inter-item similarity per block and Number of training blocks was found significant in either of the order conditions.

Transfer Phase. The following analysis of the transfer phase follows that of Johansen and Palmeri (2002) (from p. 495), which we take here for granted. Johansen and Palmeri developed a precise analysis of the patterns of categorization during the transfer phase, with some of the patterns reflecting rule-based category representations. These patterns only vary for a subset of stimuli labeled the critical stimuli. As in Johansen and Palmeri (2002), our analysis does not focus on the categorization probabilities for the nine examples encountered by the participants in the training phases, as these stimuli were globally categorized in the transfer phase as they were learned in the training phase. Figure 2 shows the average categorization probabilities corresponding to the critical transfer stimuli (those that are diagnostic of a rule-based or an exemplar-based generalization pattern: T_1, T_2, T_4, T_5, T_6). A generalization pattern for a given participant is defined by how these five critical transfer stimuli were categorized across the transfer blocks. A participant

who would classify the five critical transfer stimuli in the A category (on average across the five transfer blocks) would exhibit a AAAAA pattern. Another participant who would classify only T_1 in the A category on average would exhibit a ABBBB pattern. As indicated by Johansen and Palmeri on p. 490, participants who would apply a rule along the Size and Color dimensions would exhibit the AABBB and BBABA patterns respectively. Using more complex computation (which we cannot detailed in the present short paper), Johansen and Palmeri identified ABBBA as a pattern typical of an exemplar-based categorization.

The categorization probability $p(A)$ was the observed proportion of a stimulus categorized as A during the transfer phase across the five blocks (this calculation was made for each participant, before being averaged across participants in Figure 2). For instance, a stimulus categorized five times out of five as belonging to category A simply corresponds to $p(A) = 1$. Participants were classified as using a particular transfer pattern based on whichever response was given more often out of the five transfer responses. A within-subjects Stimulus Type × between-subjects Presentation Order ANOVA on the proportion of A responses restricted to the five critical stimuli showed a significant interaction, $F(4, 164) = 3.2, p = .014, \eta_p^2 = .07$. Figure 2 shows that following the rule-based presentation order, the average pattern (across participants) is *BBABA* (a typical rule-based generalization pattern), as opposed to *ABABA* for the similarity-based presentation order (a non-typical pattern, although similar to *ABBBA*, a typical exemplar-based pattern). We therefore found a prominent rule-based generalization pattern (*BBABA*) following the rule-based presentation order, which corresponds to a one-dimensional rule (plus exceptions) based on the Color dimension, i.e. "all red objects except B_2 versus all green objects except A_4").

Because averaging the individual patterns can be misleading, Table 1 reports the 13 different patterns that we observed in our experiment when we focused on the distribution of the generalization patterns at the individual participant level ($N = 43$). The patterns are reported in columns in the right part of the table. The respective distributions of the number of patterns observed for the Similarity and Rule conditions are reported below the patterns. Because it makes sense that an 'AAAAA' pattern is less different from an 'AABAA' pattern than from a 'BBBBB' pattern, we considered that the 13 patterns represented an ordinal variable, since each of the binary pattern can be transform into a decimal. The two-sample Kolmogorov-Smirnov test was computed using the Matlab kstest2 function, which is based on the exact theoretical distribution function of D. We found a $p = .028$ when the directional parameter was set to 'larger' (to test the alternative hypothesis that the cumulative distribution function for Sim is larger), showing that the frequency distributions for Sim and Rule are different. A simple Gamma test for the Frequency × Order Type crosstab was also significant ($p = .017$).

The distribution of the generalization patterns shows in particular that the *BBABA* pattern was less common in the similarity-based presentation order than in the rule-based order (7 vs. 14 participants respectively). Note that the *AABBB* pattern (another prominent rule-based pattern, based on a one-dimensional rule using Size instead of Color as the main dimension) is represented twice by participants who were given a similarity-based order and once by a participant in the rule-based presentation order condition. Note that in the present study, 93% of the participants who applied a one-dimensional rule still classified the exceptional items of Category A and B in their correct category during transfer, regardless of order condition. We conclude that the participants tended to use Color to separate the categories. A total of 24 participants eventually categorized the transfer objects in a way that suggests that they applied a rule-based strategy[5]. Overall, our result clearly indicates a distortion in the generalization patterns according to presentation order, and this distortion is mainly visible in the frequency associated with the *BBABA* pattern and the result of the Kolmogorov-Smirnov test.

A simpler test to see whether there was statistically more prominent rule-based generalizations in the rule-based training fell short of significance. For instance, a total of 15 participants in the rule-based condition categorized the transfer objects in a way that reflected one of the two prominent *BBABA* or *AABBB* rule-based patterns, against 7 participants for all other patterns (expressed henceforth as 15 vs. 7). Conversely, the distribution was 9 vs. 12 in the similarity-based order. This allowed us to compute a χ^2 for this 2×2 crosstab, which fell short of significance ($p = .09$, without Yates' correction; $p = .07$ using the Fisher exact test). A 16 vs. 6 distribution (instead of the 15 vs. 7 observed) in the rule-based condition would have reached significance. Note that the 16 vs. 6 simulation is quite informative since two of the 7 participants only showed a single different response to the *BBABA* pattern in the rule-based condition (i.e., *BABAA* and *BBBAA*).

Another potential issue of the preceding analyses is that a participant who responded for instance ABABA on all five transfer trials (i.e., five A responses for the first critical transfer item, five B responses for the second critical transfer item etc.) was treated the same as a participant who responded A on four trials and B on one trial for the first critical transfer item, and so on. The latter participant would still be classified as an ABABA responder since a particular transfer pattern was based on whichever response was given more often

[5] Twenty-four (7+14+2+1 = 24) is quite a high value in relation to the fairly erratic distribution of patterns that has previously been observed (see Johansen & Palmeri, 2002, p. 491).

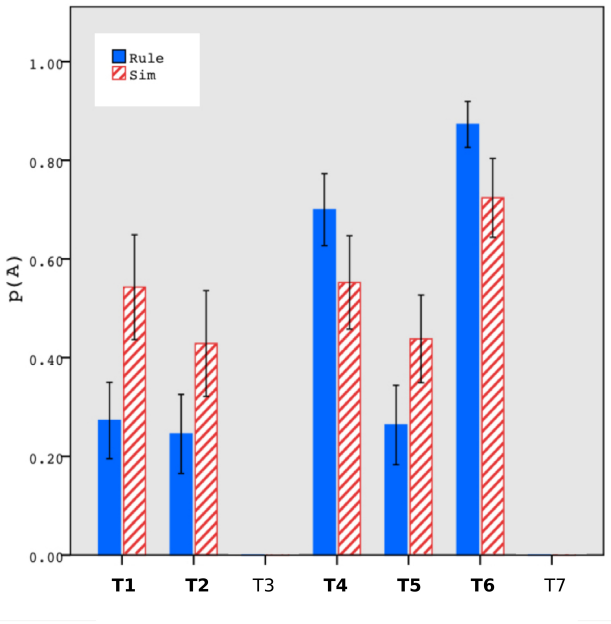


Figure 2. Average categorization probabilities of the critical transfer items (T_1, T_2, T_4, T_5, T_6) in the 5-4 category structure during the transfer phase (amounting to 5 blocks). Note. $p(A)$ is the observed proportion that each of the stimuli labeled under the abscissa was categorized as A during the transfer phase; $p(A)$ was first computed for each of the participants before being averaged across participants. The proportions are broken down by presentation order conditions (rule- vs. similarity-based). The graph does not include the categorization probabilities for the nine examples that the participants encountered in the training phases. Error bars show +/- one s.e.

out of the 5 transfer responses. Latent class analysis would be appropriate for determining whether any strategies were dominant in either conditions, but this method unfortunately requires either large data sets consisting of long time series of transfer responses from each subject, or else many more subjects (Raijmakers et al., 2014). In order to see whether performance in the similarity-order condition was more random than in the rule-based order, we considered the *proportion* of A responses to transfer trials for each of the transfer stimuli, which the previous analysis had ignored. One possibility is that the participants were simply worse at applying one of the rules to the transfer items in the similarity-order condition. We therefore classified the participants as 100% consistent when they gave at least four similar responses to each transfer item ; they were rated as 80% consistent when only one item was classified inconsistently (i.e., two or three times in a given given category), 60% consistent when two items were classified inconsistently, and so on. The mean proportion of this consistency measure across participants was slightly lower in the rule-based condition (.85, sd = .20) than in the similarity-based condition (.94, sd = .11), but this difference did not reach significance ($t(41) = -1.9, p = .06$). A second test was made to substitute the preceding consistency criterion into a more severe one by rating a transfer response as consistent only when the participant gave either five A responses or five B responses for a critical item. In that case, the mean proportion reduced to .65 (sd = .31) in the rule-based condition and .70 (sd = .25) in the similarity-based condtion, and this smaller difference did not reach significance either ($t(41) = -.57, p = .57$). This result contradicts the idea that the participants in the similarity-based condition were behaving more randomly when categorizing the new items.

General Discussion

The current study corroborates our earlier finding that a rule-based presentation order aids learning (Mathy & Feldman, 2009), this time with the 5-4 categories. More importantly, we hypothesized that participants in a rule-based training condition would show generalization patterns consistent with rule-based retrieval, and that participants in a similarity-based training condition would show generalization patterns consistent with exemplar retrieval. Our results show that participants particularly capitalized on the rule-based order. When our participants in the rule-based condition were asked to transfer their learning to unlearned instances, they mostly showed generalization patterns consistent with rule-based retrieval; and, likewise, participants in the similarity-based condition showed fewer generalization patterns consistent with rule-based retrieval, although

they did not generally exhibit generalization patterns consistent with exemplar-based retrieval. The distribution of the generalization pattern seems more random in the similarity-based condition (except for the larger group which adopted one rule-based generalization). Two thirds of these participants may have represented the categories in a heterogeneous way, almost as if the presentation had been random. However, our consistency measure of the transfer responses in the similarity-based condition did not provide support for the hypothesis that the participants behaved more randomly when categorizing the new items. One remaining interpretation is that the similarity-based condition only generated more noise during the training process. These analyses lead to our conclusion that the participants particularly capitalize on the rule-based order during learning because they can in fact abstract rules that are both consistent and simple, while participants in the similarity-based order struggle abstracting the simplest rules and memorize the stimuli without relying on a typical exemplar-based categorization process either.

These findings are important because they suggest that the manipulation of order within categories can define a context which may affect not only the speed of learning but also the nature of learning, in much the same way as co-presented items can have powerful effects on same-category comparisons (Andrews, Livingston, & Kurtz, 2011; Hammer, Diesendruck, Weinshall, & Hochstein, 2009). The different representations that were observed during the transfer phase suggest a potential connection with the idea that a mere manipulation of order inside a category can determine the meaning that is inferred from a set of objects, like a disease from a checklist of symptoms for instance (Kwan, Wojcik, Miron-Shatz, Votruba, & Olivola, 2012).

Limitations

One limitation of the present study is that the total amount of category learning was not equated in the two experimental conditions. Because we used a learning criterion, the order conditions were confounded with the number of blocks of training; on average participants in the rule-based condition reached the learning condition in fewer blocks. In effect, using a fixed learning criterion meant that rule-based participants received on average fewer training trials—an inevitable result of the superior learning entailed by the rule-based order. Note though that alternative procedures would have created even more serious design problems. For example, equating the number of blocks in each condition would have led either to extremely poor learning in some conditions (if the number were low), and thus transfer performance reflecting more noise than learning); or (if the number were high) ceiling effects and mostly exemplar-based memorization. Effectively, early phases of learning tend to be more rule-based, and later phases more exemplar-based (Johansen & Palmeri, 2002), potentially introducing a further bias if learning is prolonged. Of course, our results might in part reflect the specific learning criterion we chose, and only future experiments can reveal whether the pattern of results might sublty alter with alternative criteria. However, the essential point is that because two presentation orders lead to such different rates of learning, it seems necessary that when comparing them we equate the amount of actual learning rather than the mere exposure to training stimuli. A similar limitation had been noted by (Casale, Roeder, & Ashby, 2012), which also questioned whether the specific transfer observed for their rule-based condition might have been due to the faster learning of their rule-based task. The authors also rejected the difficulty account arguing for instance that transfer should occur in any task in which there is learning, and concluded that rule-based category learning is functionally distinct from other categorization tasks during transfer.

To confirm our interpretation of the results, we reanalyzed our results by matching the participants of both groups by the number of blocks that each of them used to reach the learning criterion. The participants who could not be matched given their performance during training were left out of the new analysis. (For example, because a single participant reached the learning criterion using 12 blocks, in the rule-based order, this participant was omitted from the new analysis). As a result, presentation order and amount of learning were equated a posteriori. We then reanalyzed the data, without the stimulus-exposure confound. The revised results were virtually identical to the original, with for instance a $r = .99$ pearson correlation between the mean probabilities shown in Figure 2 and the corresponding means in the new analysis. We conclude that the differing patterns of transfer between the two presentation orders cannot be attributed to the different amounts of stimulus exposure in the two conditions.

A second limitation is due to a choice we made in the design of the experiment that may hamper its generalizability. For instance, only one kind of clustering (based on Color) was used to determine the clusters of the rule-based condition, and a similar experiment could have been conducted using the Size and the Shape dimensions as well. Still, because two-thirds of our participants followed the *one-dimensional Color rule plus exception* that was induced in the Rule condition, we felt that tripling our sample to use for instance all of the Shape/Color/Size canonical dimensions to confirm this trend would have been too costly in the short-term.

Conclusion

Our main results tends to agree with previous theories implying that different learning processes mediate category representations (Ashby & Ell, 2001; Sloman, 1996) as we found that different types of presentation orders change the generalization patterns of a single category structure. Our previous work (Mathy & Feldman, 2009) already showed

that rule-based orders and similarity-based orders inflect the speed of category learning. The present result further documents that different category representations can be induced by such orders and that the distortion of these representations are extended to several stimuli among the categories induced. It is not clear whether the similarity-based training leads to exemplar-based generalizations because there is no clear evidence for similarity-based responding in our result, but the patterns of generalization that the learners show change in comparison to the rule-based condition[6]. The rule-based condition however more clearly produced rule-based classification during transfer and a clear peak for one of the typical generalization patterns.

A reasonable question is whether our participants are truly using rule-based representations within a putative dual process of reasoning. Our results are also compatible with other approaches which tend to see the two systems of reasoning as the two endpoints of a continuum (Briscoe & Feldman, 2011). For instance, selective attention in an exemplar-based processes can produce patterns of responses similar to single-dimension rule use (indeed, this is how Johansen & Palmeri, 2002, model rules; see also Kruschke, 1992), and as such, responding on the basis of a single dimension does not in itself indicate rule use (likewise, participants can also apply multidimensional rules which can account for exemplar effects; see Nosofsky, Palmeri, & McKinley, 1994). However, we believe that our procedure involved no pressure to increase the likelihood of single-dimension responding as predicted for instance by the Extended General Context Model (Lamberts, 1998) or Combination Theory (Wills, Inkster, & Milton, 2015).

The rule-based effect is particularly interesting in light of the fact that the 5-4 categories are known to predominantly elicit memorization of individual objects (Blair & Homa, 2003). Thus it cannot be easily argued that the category structure itself generally favor a rule-based learning mechanism, although the learning process was not prolonged and might have mostly induced a rule-based process. Overall, these results might be crucial for testing rule- and exemplar-based models of human categorization behavior (Rouder & Ratcliff, 2004) and sequential hypothesis-testing models (Markant & Gureckis, 2014), as well as better identifying rule learners *vs*. exemplar learners (McDaniel, Cahill, Robbins, & Wiener, 2014), especially for testing models that are known to implement an incremental architecture (Love, Medin, & Gureckis, 2004; Sakamoto, Jones, & Love, 2008; Stewart, Brown, & Chater, 2002) to account for such sequence effects.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, *120*, 3-19.

Andrews, J. K., Livingston, K. R., & Kurtz, K. J. (2011). Category learning in the context of co-presented items. *Cognitive processing*, *12*, 161–175.

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, *5*, 204-210.

Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & cognition*, *30*, 666-677.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: the roles of discrimination and retrieval. *Memory & cognition*, *41*, 392-402.

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.

Briscoe, E., & Feldman, J. (2011). Conceptual complexity and the bias/variance tradeoff. *Cognition*, *118*, 2–16.

Carvalho, P. F., & Goldstone, R. L. (2011). Stimulus similarity relations modulate benefits of blocking and interleaving during category learning. In *Abstracts of the 52nd annual meeting of the psychonomic society.* Seattle, WA..

Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & cognition*, *42*, 481–495.

Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory & cognition*, *40*, 434–449.

Clapper, J. P., & Bower, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 443-460.

Clapper, J. P., & Bower, G. H. (2002). Adaptative categorization in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 908-923.

Cohen, A. L., & Nosofsky, R. M. (2003). An extension of the exemplar-based random-walk model to separable-dimension stimuli. *Journal of Mathematical Psychology*, *47*, 150-165.

Edmunds, C. E. R., Milton, F., & Wills, A. J. (2015). Feedback can be superior to observational training for both rule-based and information-integration category structures. *Quarterly journal of experimental psychology (2006)*, *68*, 1203-1222.

Elio, R., & Anderson, J. R. (1981). Effects of category generalizations and instance similarity on schema abstraction. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 397-417.

[6] One possibility is that showing our participants the correct category label during training tended to discourage associative learning (Ashby, Maddox, & Bohil, 2002). However, a subsequent study has shown that, under more controlled conditions, the presence of feedback impacts rule-based and similarity-based processes to a similar degree (Edmunds, Milton, & Wills, 2015).

Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition*, *12*, 20–30.

Gagné, R. M. (1950). The effect of sequence of presentation of similar items on the learning of paired associates. *Journal of Experimental Psychology*, *40*, 61-73.

Garner, W. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory & Cognition*, *24*(608-628).

Hahn, U., & Chater, N. (1998). Similarity and rules: distinct? exhaustive? empirically distinguishable? *Cognition*, *65*, 197-230.

Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: what makes the difference? *Cognition*, *112*, 105–119.

Johansen, M. K., & Kruschke, J. K. (2005). Category representation for classification and feature inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1433-1458.

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

Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*, 97-103.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: is spacing the "enemy of induction"? *Psychological Science*, *19*, 585-592.

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*, 498-503.

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

Kwan, V. S. Y., Wojcik, S. P., Miron-Shatz, T., Votruba, A. M., & Olivola, C. Y. (2012). Effects of symptom presentation order on perceived disease risk. *Psychological science*, *23*, 381–385.

Lafond, D., Lacouture, Y., & Mineau, G. (2007). Complexity minimization in rule-based category learning: Revising the catalog of boolean concepts and evidence for non-minimal rules. *Journal of Mathematical Psychology*, *51*, 57-74.

Lamberts, K. (1998). The time course of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 695-711.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227-260.

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

Markant, D., & Gureckis, T. (2014). Is it better to select or to receive? learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*, 94–122.

Mathy, F., & Feldman, J. (2009). A rule-based presentation order facilitates category learning. *Psychonomic Bulletin & Review*, *16*, 1050-1057.

Mathy, F., & Feldman, J. (2011). Presentation order effects on category learning and category generalization. In *52nd Annual Meeting of the Psychonomic Society* (Vol. Nov 3-6, p. 58). Seattle, WA..

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.

Medin, D. L., & Bettger, J. G. (1994). Presentation order and recognition of categorically related examples. *Psychonomic Bulletin & Review*, *1*, 250-254.

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

Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 275-292.

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

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

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.

Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of experimental psychology. General*, *133*, 63–82.

Sakamoto, Y., Jones, M., & Love, B. C. (2008). Putting the psychology back into psychological models: Mechanistic versus rational approaches. *Memory & Cognition*, *36*(1057-1065).

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 405-415.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*, 13, whole No. 517.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.

Sloman, S. A., & Rips, L. J. (1998). Similarity as an explanatory construct. *Cognition*, *65*, 87–101.

Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, *65*, 167-196.

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.

Stewart, N., Brown, G. D. A., & Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 3-11.

Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: an investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, *39*, 750-763.

Wills, A. J., Graham, S., Koh, Z., McLaren, I. P. L., & Rolland, M. D. (2011). Effects of concurrent load on feature- and rule-based generalization in human contingency learning. *Journal of experimental psychology. Animal behavior processes*, *37*, 308–316.

Wills, A. J., Inkster, A. B., & Milton, F. (2015). Combination or dif-

ferentiation? two theories of processing order in classification. *Cognitive Psychology*, *80*, 1–33.

Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2003). Prototype and exemplar accounts of category learning and attentional allocation: a reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1160-1173.