Title
Sequential similarity and comparison effects in category learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Carvalho, Paulo
Goldstone, Robert

Publication Date
2011

Peer reviewed

Sequential similarity and comparison effects in category learning

Paulo F. Carvalho (pcarvalh@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University 1101 East Tenth Street
Bloomington, IN 47405 USA

Robert L. Goldstone (rgoldsto@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University 1101 East Tenth Street
Bloomington, IN 47405 USA

Abstract

Order effects in category learning have been previously demonstrated. Specifically, alternation between exemplars of two categories has been shown to improve category learning and discrimination, compared to presenting exemplars of each category in separate blocks. However, the mechanisms underlying order effects are still not completely known. Remaining issues pertain to the relevance of within and between category similarities, and the role of comparing sequentially presented objects. We present two experiments: in Experiment 1 within- and between-category similarity are manipulated simultaneously with presentation schedule. In Experiment 2, alternation between categories is compared to two blocked conditions: one in which very similar stimuli are presented successively, and another in which they are dissimilar. Our results show a clear overall advantage of low similarity in categorization performance, but no effect of presentation schedule. Also, alternation between categories is shown to result in better performance than the blocked condition with more dissimilar stimuli. **Keywords:** category learning; order effects; spacing; perceptual learning.

Introduction

The ability to discriminate between two similar stimuli has been shown to be influenced by the way stimuli are presented. Take, for example, two categories A and B. Exemplars from these two categories can be presented interleaved, i.e., A B A B A B. Another way to present the same exemplars is to show all exemplars of category A before starting presentation of the exemplars from category B, i.e., A A A B B B. This exemplar ordering manipulation can have a great impact on the ability to discriminate between A and B. Indeed, results show that performance improves considerably with interleaved exposure to the stimuli (e.g., Dwyer, Hodder, & Honey, 2004; Lavis & Mitchell, 2006).

This effect has been replicated several times, using different tasks, stimuli, and sensory modalities both in humans and other species. Improved discrimination for interleaving presentations also seems to be highly transferable to new situations that make use of the same kind of knowledge (Mitchell, Kadib, Nash, Lavis, & Hall, 2008) and to result in better deductive extraction of the relevant categories, such as painter identities (Kornell & Bjork, 2008). In light of this evidence, alternation of exemplars has been described as highly beneficial for discrimination learning, and can be potentially implemented in several informal and formal learning contexts (Rohrer & Pashler, 2010; Taylor & Rohrer, 2010).

However, a full description of how the interleaving advantage occurs has not yet been provided. Some authors have argued for the long-known mnemonic powers of spaced rather than massed practice (Kornell, 2009; Kornell & Bjork, 2008; Vlach, Sandhofer, & Kornell, 2008) and have made a connection between interleaving and spacing. When different presentations of A are interleaved with B presentations, then the A presentations will be more spaced. However, evidence inconsistent with this proposal has been presented. Mitchell, Nash, and Hall (2008), for example, presented three highly similar stimuli, A, B and C, to participants. In one condition A and B were interleaved and C was presented spaced. In the spaced condition, two instances of C were separated by an interval during which no information was presented. The duration of this interval was equivalent to the presentation of another stimulus. This manipulation still resulted in better discrimination for stimuli presented interleaved (for a similar demonstration using paitings see Kang & Pashler, in press). In a follow-up experiment, the authors added a new stimulus, D. This stimulus was very different from A, B and C. Interleaving A and B resulted in better performance than interleaving C and D.

Mitchell and collaborators (Mitchell, Kadib, et al., 2008; Mitchell, Nash, et al., 2008) suggested that the advantage of interleaving might still result from different memory encodings promoted by each presentation schedule. Their proposal is that when highly similar stimuli are alternated, then the features that they share will be presented several times in a row, and thus attention shifts towards the differences, enriching the memory trace for the differences between the two categories. When stimuli are blocked, however, all of the features of each category, both those unique to each category and those that are shared across categories, are presented several times in a row, resulting in a relative decrease in attention to unique features and a poorer memory trace for the purposes of discrimination.

Also central to Mitchell and collaborators' proposal (Mitchell, Kadib, et al., 2008; Mitchell, Nash, et al., 2008) is the process of comparison that results from shifting attention to discriminating features. The importance of comparison in the sequence of object presentations was also proposed by Eleanor Gibson (1969) in her influential theory of perceptual learning. Gibson proposed that the opportunity to compare the stimuli was crucial in facilitating a process of differentiation that would result in better discrimination. Unfortunately, the mechanism through which this differentiation process might occur was never fully articulated. Nonetheless, recent work supports Gibson's proposal by demonstrating that simultaneous presentation of two highly similar stimuli results in better discrimination than interleaving them (Mundy, Honey, & Dwyer, 2007, 2009).

Thus, the advantage of interleaving seems to stem in part from the similarity between successively presented stimuli and an observer's ability to compare them. Furthermore, we propose that, if comparison is important for the advantage of interleaving, blocking stimuli from the same category might, under certain situations, also provide an opportunity for key comparisons – resulting in equally good performance.

For this purpose, two experiments were conducted. In both experiments, morphed pictures of human faces were used (see Figure 1). Similarity between these is easily manipulated because stimuli closer in the matrix are more similar than those farther apart. In Experiment 1 we manipulate both the between- and within-category similarity among the stimuli and presented them in either a blocked condition or an interleaved condition. We predict that performance will be improved for interleaved presentation in the high similarity condition. According to Mitchell et al. (Mitchell, Kadib, et al., 2008; Mitchell, Nash, et al., 2008), interleaving accentuates between-category differences, resulting in an improved ability to distinguish the categories. There are two ways in which the similarity between objects could affect categorization difficulty. First, as objects within a category become more similar, categorization is expected to become easier because it is easier to see what the category members have in common. Second, as objects from different categories become more similar, categorization is expected to become more difficult because it is harder to distinguish the categories. If interleaving accentuates features that discriminate between categories, then this would be expected to particularly benefit situations in which all objects are similar to one another because of the difficulty in identifying discriminating features.

In Experiment 2 we approach the importance of comparisons by contrasting performance in the interleaved condition with two blocked conditions. In both conditions, stimuli from the same category are presented in separate blocks but in one condition only highly similar stimuli are presented successively and in the other only more dissimilar stimuli are presented successively. We propose that these three conditions offer opportunities for three different kinds of comparisons: while interleaving allows for comparisons between the two categories, blocking allows for comparisons within the category. Moreover, within-category comparisons might be more informative when the successively presented stimuli are highly similar (R. Goldstone, 1996). This situation will allow the creation of a good representation of the common features of that category and performance similar to interleaving the two categories.

Experiment 1

Method

Participants Seventy-one Indiana University undergraduate students participated in this experiment in partial fulfillment of a class requirement. Participants were randomly distributed on an on-arrival basis to one of two similarity conditions: high similarity vs. low similarity. All participants completed all the phases of this experiment and 9 were excluded for not reaching 50% correct responses during category learning.

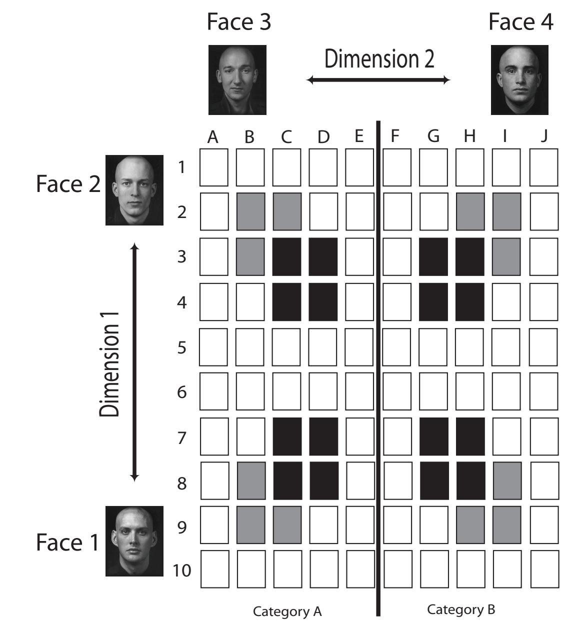

Figure 1: Example of a set of stimuli used in Experiments 1 and 2; two other matrices were created using the same method. These stimuli varied along 2 arbitrary dimensions creating a 10 x 10 matrix of morphed faces. Stimuli shaded in dark grey were only presented in the low similarity condition of Experiment 1 and stimuli shaded in black were only presented in the high similarity condition of Experiment 1 and in Experiment 2. Notice that high and low similarity conditions shared 4 stimuli that, although shaded in black, were also presented in the low similarity condition (C3, C8, H3 and H8). The line between columns E and F represents one of the possible category boundaries used during the experiments (the other possibility was a line between rows 5 and 6).

Stimuli The stimuli were morphs of bald male faces selected from Kayser (1997) using the blending technique outlined in Steyvers (1999). The faces varied along two arbitrary dimensions (see Figure 1). Each dimension was achieved by creating negative contingencies between two faces. For example, along Dimension 1, the more of face 1 is present in each morphed face, the less there is of face 2. These two contingencies are independent, i.e., the proportion of face 1 relative to face 2 is independent of the proportion of face 3 relative to face 4, resulting in a set of blended faces that differed only in the proportion of each of the 4 original faces. Using 8 different original faces, two different 10 x 10 matrices were created. Not all faces from each 10 x 10 grid were presented to participants. Indeed, for each group only 16 faces (8 from each category) were presented to participants. Which faces from the 10 x 10 grid were presented depended upon the experimental group. Participants in the High Similarity condition saw faces that were closer together in the matrix (black squares in Figure 1). Conversely, participants in the Low Similarity condition saw only faces that were farther apart in the matrix (grey squares in Figure 1). Notice that this similarity manipulation was both within and between categories, thus there are two clusters of 4 faces in each category for each similarity condition. This manipulation will be relevant for the modifications introduced in Experiment 2.

Procedure This experiment had 2 phases. Each phase was composed of two tasks: a category learning task and a generalization task regarding the categorization learned. Each phase differed only in the face matrix used and in the way stimuli were presented during category learning.

Category Learning Each category learning task was composed by 4 blocks with 64 training trials each. On each trial, a face was presented in the center of the computer screen for 500 ms. After the face was removed, the participant was tasked with classifying it into one of the two possible 'Clubs', by pressing the corresponding key (each key was attributed to one Club: Q vs. P and A v.s L) in the computer keyboard. After the participant's response the face was presented again for 2000 ms together with the presentation of feedback relative to the accuracy of participants' response and the category of the face. Afterward, there was a 1000 ms inter-trial interval and then a new trial would start.

The two category learning tasks differed in the order of presentation of category exemplars. In the interleaved condition, the exemplars alternated 75% of the time between one category and the other. In the blocked condition, alternation between categories occurred only 25% of the time. Thus, while in the interleaved condition the probability of a trial with a face from one category being followed by a trial with a face from that same category was low, in the blocked condition this probability was high. We used this probabilistic approach rather than creating purely interleaved or blocked conditions in order to diminish the possibility that participants noticed the pattern of alternation in responses, which would affect categorization accuracy. Furthermore, if a purely blocked condition was used there would be no way to guarantee subjects' attention to the task, as there would be no uncertainty in response and all faces would belong to the same 'Club'. This approach has been used before, in similar tasks with successful results (R. Goldstone, 1996). Which condition was presented first in the experiment was counterbalanced across participants, as was the allocation of one of the 10 x 10 matrices to each condition.

Generalization Each category learning task was followed by a generalization task regarding the two categories just learned. The generalization task was composed of 100 trials, corresponding to the presentation of all the faces in that 10 x 10 matrix. On each generalization trial a face was presented for 500 ms after which participants had to indicate the Club to which the face belonged. There was a 1000 ms inter-trial interval, after which a new face was presented. There was no response feedback in this task.

Results

The graph in Figure 2 depicts the main results from Experiment 1 regarding category learning across the 4 learning blocks and generalization accuracy, for each combination of presentation schedule (interleaved and blocked) and similarity (high similarity and low similarity).

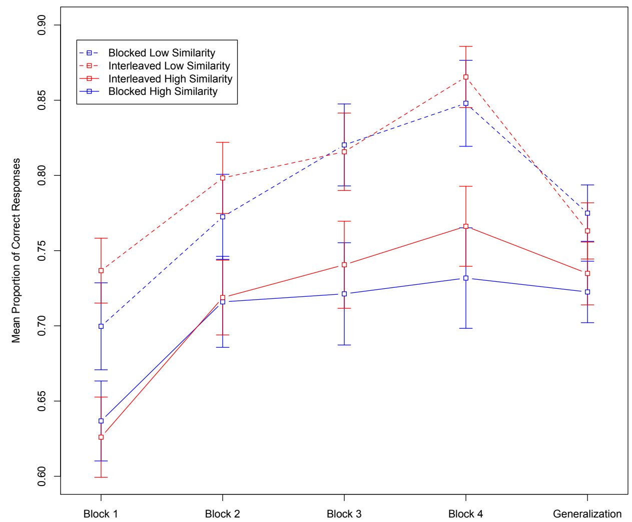

Figure 2: Mean proportion of correct responses in Experiment 1. Performance improves across learning blocks and is overall better for the High Similarity condition with no effect of presentation schedule or interaction. There is no effect of either similarity or schedule of presentation in the generalization task. Error bars indicate standard errors.

A 2 x 2 x 4 mixed ANOVA with similarity condition (high vs. low similarity) as a between-subjects factor and presentation schedule (interleaved vs. blocked) and learning block (1 vs. 2 vs. 3 vs. 4) as within-subject factors revealed a significant improvement in accuracy during category learning, regardless of presentation schedule or similarity condition, F (3,180) = 45.33, $p < .0001$, MSE = 0.02. Likewise, a significant main effect of similarity revealed that participants' accuracy is higher for the low similarity condition when compared to the high similarity condition, F (1, 60) = 13.02, p = .0006, MSE = 0.15. However, participants' accuracy is comparable between interleaved and blocked presentation schedules, with no main effect of presentation schedule found, F (1,60) = 0.22, p = .64, MSE = 0.26.

Likewise, there is no interaction between similarity condition and presentation schedule, F (1,60) = 0.01, p = .91, MSE = 0.26, similarity condition and learning block, F (3, 180) = 0.99, p = .40, MSE = 0.02 or presentation schedule and learning block, F (3,180) = 0.21, p = .89, MSE = 0.02. Additionally, the interaction between the 3 variables was also not significant, F (3, 180) = 1.01, p = .39, MSE = 0.02.

Regarding the generalization task, a 2 x 2 mixed ANOVA was performed with similarity condition as a between-subjects factor and presentation schedule as a within-subjects factor, considering only trials with novel stimuli. This analysis revealed that participants perform equally well regardless of presentation schedule or similarity condition. There is no main effect of similarity, F (1, 60) = 2.44, p = .12, MSE = 0.04, or schedule of presentation, F (1, 60) = 0. Likewise, no interaction between these two variables was found, F (1,60) = 0.2, p = 0.67, MSE = 0.05. However, a 2 x 2 x 5 mixed ANOVA with similarity condition as between-subject factor and presentation schedule and distance of the stimuli to the category border (i.e., the position in the grid relative to the category limit: 1 vs. 2 vs. 3 vs. 4 vs. 5) as within-subjects factors, revealed an improvement in performance with the increase in distance, F (4, 276) = 126.20, p < .0001, MSE = 0.02, as well as an interaction between distance and similarity condition, in which the advantage of the low similarity condition over high similarity increases with the increase in distance to category boundary, F (4, 176) = 3.43, p = .01, MSE = .02. All the other effects were non-significant.

In sum, these results show that decreasing the similarity between exemplars of the two categories increases participants' performance, even if within-category similarity also decreases. While an advantage for categorization might be expected from decreasing between-category similarity because it makes discrimination easier, the decreased within-category similarity could have equally detrimental results by making categories very heterogeneous. Thus, one might expect the advantage of decreasing between-category similarity to be nullified by the increase in within-category similarity. However, our results show that the beneficial influence of decreasing between-category similarity outweighs the detrimental influence of decreasing within-category similarity. One might also have expected that interleaving exemplars of both categories in the high similarity condition would be particularly beneficial by accentuating the differences between the two categories and thus reducing the difficulty in discriminating very similar categories. Yet, the results of Experiment 1 do not show any improvement in performance when exemplars of both categories are interleaved, which is surprising given the wealth of evidence for the advantage of interleaving.

One of the possible reasons for this null result might be the opportunity for comparisons that can take place during blocked presentation. Although interleaving allows more opportunity for inter-category comparisons, there has been some evidence for the importance of within-category comparisons (R. Goldstone, 1996). Looking specifically at the interleaving-blocking phenomenon, one might argue that not all blocked presentation schedules are equivalent. Consider, for example, a situation in which (a) the comparison is possible between two stimuli that, although in the same category have many differences in contrast to (b) the comparison is only possible between two stimuli that are in the same category and are very similar. The two within-category comparisons offer the opportunity to extract different kinds of information. While the first might highlight irrelevant information about within-category dissimilarities, the latter might emphasize relevant within-category similarities. It is possible that one of these comparison processes approaches the advantageous effect of interleaving stimuli from two similar categories. Given that we did not control for these two possible comparisons opportunities during blocked presentation in Experiment 1, Experiment 2 will approach this question in more detail.

Experiment 2

Method

Participants Seventy-six undergraduate students from Indiana University participated in this experiment in partial fulfillment of a class requirement. All participants completed all the phases of this experiment and 17 were excluded for not reaching 50% correct responses during category learning.

Stimuli In this experiment we used the same two 10 x 10 matrices used in the first experiment and also a third one, constructed in the same way as the other two, using 4 new faces from Kayser (1997). As in Experiment 1, not all 100 faces from each matrix were presented during category learning. Only faces that were closer to each other and to the category boundary were presented (i.e., those corresponding to the high similarity condition from Experiment 1).

Procedure This experiment had 3 phases. Each phase was composed of a Categorization task followed by a generalization task based on the categorization done before. Each phase differed only in the way stimuli were presented in the Categorization Task as well as which matrix of faces was used.

Category Learning Each category learning task was composed of 3 blocks of 64 trials each (8 exemplars of each category repeated 4 times). There were 3 different category learning conditions: interleaved, blocked-close and blocked-far. Each of these consisted in the categorization of the faces in one of 3 pairs of 'Clubs' (P vs. Q, A vs. L, and Z vs. M). The interleaved condition consisted in high alternation rate of the categories within each block (75% category change from trial to trial). Blocked-close and blocked-far both consisted in low alternation conditions (25% category alternation). They differed, however, in the proximity of the stimuli presented in sequence. In the blocked-close condition each stimulus from one category would only be followed by a stimulus of the same category if that stimulus was from the same within-category cluster (see Figure 1 for an illustration). In the blocked-far condition each stimulus would only be followed by a stimulus from another within-category cluster. This resulted in a manipulation of the within-category similarity between the stimuli being shown successively in the blocked conditions. The order of these conditions was randomized across participants as well as which 10 x 10 matrix was used in each one of them. Every other detail not stated here was the same as in Experiment 1.

Generalization Each category learning task was immediately followed by a generalization task in which participants classified each presented face in one of the two groups previously learned, without feedback. This generalization task was similar to the one in Experiment 1 except for the number of stimuli presented. In this Experiment, generalization consisted of 84 trials corresponding to the faces from the 10 x 10 matrix that had not been presented during category learning.

Results

The graph in Figure 3 depicts the main results from Experiment 2 regarding category learning and generalization tasks for each one of the learning conditions: interleaved, blocked-close and blocked-far.

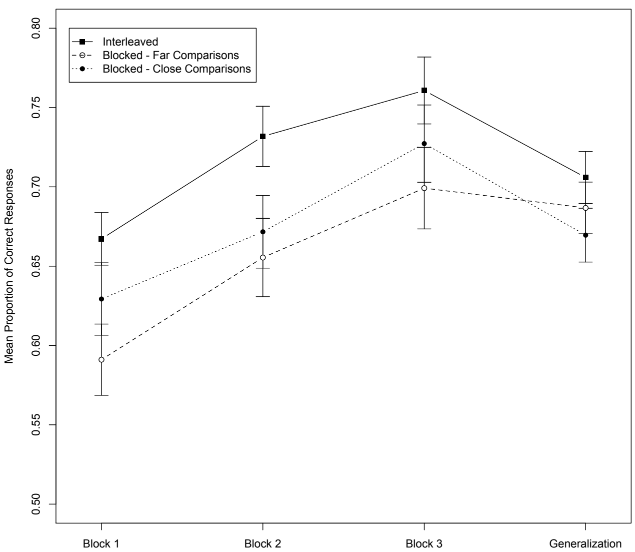

Figure 3: Mean proportion of correct responses for Experiment 2. Participants' performance improves across learning and is overall better in the interleaved condition than in the blocked-far condition. There is no difference between the interleaved and blocked-close conditions. During generalization there are no differences between any of these conditions. Error bars indicate standard errors.

A 3 x 3 within-subjects ANOVA with both learning block (1 vs. 2 vs. 3) and presentation schedule (interleaved vs. blocked-close vs. blocked-far) as within-subjects factors was performed. This analysis revealed a main effect of learning block, F (1,57) = 83.54, $p < .0001$, $MSE = 0.02$. As it can be seen in Figure 3, there is an overall improvement in categorization accuracy throughout category learning. Moreover, there is also a main effect of presentation schedule during category learning, F (2,114) = 4.25, $p = .02$, $MSE = 0.11$. A set of planned comparisons further shows that response accuracy is significantly higher for the Interleaved condition as compared with the Blocked-Far condition, F (1,57) = 8.74, $p = .005$, $MSE = 0.7$, but not significantly higher than for the Blocked-Close condition, F (1,57) = 2.91, $p = .09$, $MSE = 0.8$. Finally, no interaction between categorization improvement and presentation condition was found, F (2, 114) = 0.20, $p = .83$, $MSE = .02$.

For the results of the generalization task, a within-subjects ANOVA with presentation schedule as the only factor was performed. This analysis revealed that participants are equally good at categorizing new stimuli into one of the categories they had just learned, regardless of how the categories had been presented before, with no main effect of presentation schedule, F (2,114) = 0.85, $p = .43$, $MSE = 0.05$. However, analyzing performance as a function of the stimuli distance to the border revealed a significant improvement in performance with greater distances, regardless of presentation schedule, F (4, 228) = 58.72, $p < .0001$, $MSE = 0.04$, but no main effect of presentation schedule, F (2, 114) = 0.54, $p = .58$, $MSE = 0.23$, or interaction between the two variables, F (8, 456) = 0.90, $p = .51$, $MSE = 0.03$.

These results demonstrate that interleaving the presentation of exemplars from two categories results in better categorization performance than blocked presentation, but only if just dissimilar exemplars from the same category are presented in successive trials. This result contributes to the notion that successive comparisons play an important role in the interleaved advantage and category learning as a whole.

General Discussion

In Experiment 1, although there was no effect of presentation schedule, a strong effect of similarity in categorization accuracy was found. More precisely, the advantage of decreasing between-category similarity in the low similarity condition outweighs the disadvantage of the accompanying decrease in within-category similarity, resulting in better performance relative to the high similarity condition, in which both within- and between-category similarities were high. These results are, however, contrary to previous results showing that a high similarity condition results in better categorization accuracy than a low similarity condition (R. L. Goldstone, Lippa, & Shiffrin, 2001). This opposing pattern of results might be related to the greater overall similarity between the stimuli in our experiments when compared with Goldstone et al. (2001), in which the low similarity condition involved the presentation of four completely different faces. In a matrix like the ones we used here, between-category similarity seems to have a greater impact on categorization accuracy than within-category similarity. This greater overall similarity is also expected to result in more interrelated concepts, each category being defined in opposition to the other (R. Goldstone, 1996). This proposal is also consistent with the results of Experiment 2 showing that interleaving resulted in better performance than blocked-far, but not blocked-close. Moreover, it is also consistent with the results from the generalization task in both experiments, showing that regardless of similarity condition, participants are better in classifying faces more distant from the category border. This result suggests the creation of a caricature representation of each category, i.e., a representation that is not the central tendency of that category but departs from it in the opposite direction from the central tendency of the concept simultaneously learned, characteristic of interrelated concepts (R. Goldstone, 1996).

One of the possible reasons for the similar results between interleaved and blocked-close conditions in Experiment 2 might be related to the kind of comparisons they allow. As stated in the Introduction, current theories of order effects in perceptual learning propose that interleaving stimuli from different categories allows the participants to more directly identify differences between the two categories. However, we argued that blocked presentation might also offer an opportunity to establish informative comparisons. When two very similar stimuli from the same category are presented in a sequence, their similarities will be highlighted (R. Goldstone, 1996), resulting in performance that approaches that seen after interleaved study of the categories. Unfortunately, the results presented here do not allow for a direct comparison between the two blocked conditions in Experiment 2. Nonetheless, a case could be made that the two are less favorable for category learning than interleaving – thus the proximate lower values of accuracy. This disadvantage can, in turn, be eased by presenting only very similar stimuli successively in the blocked condition.

In sum, the characteristics inherent to the categorization tasks presented here make between-category similarity the better overall predictor of good categorization performance, contrary to some previous evidence using similar stimuli. However, performance in the high similarity condition can be improved by alternating exemplars from each category or by sequentially presenting very similar stimuli from the same category. These results point to the creation of interrelated concepts and the representation of each category in opposition to one another. The work presented here also shows that whether interleaving or blocking is more beneficial is probably the result of an interaction between the characteristics of the stimuli and the comparisons that the observer is able to establish. Interleaving two similar stimuli seems to result in good performance, presumably because it allows for a better contrast between the distinguishing features of each stimulus. Similarly, blocking very similar stimuli approaches that performance by allowing for a better identification of the features of that category.

Acknowledgments

Research supported in part by a Fulbright Research Fellowship to the first author, and by National Science Foundation REESE grant 0910218. The authors would like to thank Akshat Gupta for his help collecting the data.

References

Dwyer, D., Hodder, K., & Honey, R. (2004). Perceptual learning in humans: roles of preexposure schedule, feedback, and discrimination assay. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 57*(3), 245-259.

Gibson, E. (1969). *Principles of perceptual learning and development*. New York: Apleton-Century-Crofts.

Goldstone, R. (1996). Isolated and interrelated concepts. *Memory & Cognition, 24*(5), 608-628.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. (2001). Altering object representations through category learning. *Cognition, 78*(1), 27-43.

Kang, S. H. K., & Pashler, H. (in press). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*.

Keyser, A. (1997). *Heads*. New York: Abbeville Press.

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology, 23*, 1297-1317.

Kornell, N., & Bjork, R. (2008). Learning concepts and categories: is spacing the "enemy of induction"? *Psychological Science, 19*(6), 585-592.

Lavis, Y., & Mitchell, C. (2006). Effects of preexposure on stimulus discrimination: an investigation of the mechanisms responsible for human perceptual learning. *The Quarterly Journal of Experimental Psychology, 59*(12), 2083-2101.

Mitchell, C., Kadib, R., Nash, S., Lavis, Y., & Hall, G. (2008). Analysis of the role of associative inhibition in perceptual learning by means of the same-different task. *Journal of Experimental Psychology: Animal Behavior Processes, 34*(4), 475-485.

Mitchell, C., Nash, S., & Hall, G. (2008). The intermixed-blocked effect in human perceptual learning is not the consequence of trial spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 237-242.

Mundy, M., Honey, R., & Dwyer, D. (2007). Simultaneous presentation of similar stimuli produces perceptual learning in human picture processing. *Journal of Experimental Psychology: Animal Behavior Processes, 33*(2), 124-138.

Mundy, M., Honey, R., & Dwyer, D. (2009). Superior discrimination between similar stimuli after simultaneous exposure. *The Quarterly Journal of Experimental Psychology, 62*(1), 18-25.

Rohrer, D., & Pashler, H. (2010). Recent Research on Human Learning Challenges Conventional Instructional Strategies. *Educational Researcher, 39*(5), 406-412. doi: 10.3102/0013189X10374770

Steyvers, M. (1999). Morphing techniques for manipulating face images. *Behavior Research Methods, Instruments, & Computers, 31*(2), 359-369.

Taylor, K., & Rohrer, D. (2010). The Effects of Interleaved Practice. *Applied Cognitive Psychology, 24*(6), 837-848. doi: 10.1002/acp.1598

Vlach, H., Sandhofer, C., & Kornell, N. (2008). The spacing effect in children's memory and category induction. *Cognition, 109*(1), 163-167.