The Role of Category Structure in Category Learning

Hyungwook Yim (yim.31@osu.edu)
Department of Psychology, The Ohio State University
267 Psychology Building, 1835 Neil Avenue
Columbus, OH 43210, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Department of Psychology, The Ohio State University
239 Psychology Building, 1835 Neil Avenue
Columbus, OH 43210, USA

Abstract

Two category-learning experiments were conducted to examine the role of category structure and learning regime in category learning. We particularly focused on effects of these factors on selective attention, which was measured by eye-tracking methods. Results show that even though supervision was weaker than in previous studies, attention optimization and cost of attention were observed during category learning (Experiment 1). Moreover, there were faster learning and stronger attention optimization when statistically denser categories were learned (Experiment 2). At the same time, there were weaker costs of selective attention when learning denser categories than when learning sparser categories. Results are discussed in relation to theories of category learning.

Keywords: category learning, cost of selective attention, category structure, eye tracking

Introduction

Selective attention is one of the key components in category learning (Kruschke, 1992; Nosofsky, 1986; Shepard, Hovland, & Jenkins, 1961). The ability to selectively attend to category-relevant dimensions aids the learner to ignore category-irrelevant information and makes learning more efficient. For example, when learning how to distinguish Siberian Huskies from Alaskan Malamutes, which look very similar, the color of the eyes is one of the relevant features one should look for (most Huskies have blue eyes and Malamutes have brown eyes). Therefore, learning to focus on the color of the eyes while ignoring other irrelevant features (e.g. color of the fur or markings) would aid learning the two categories. Selective attention could be captured in category learning tasks that involve eye-tracking as attention optimization, where looking to category-relevant information increases and looking to irrelevant information decreases (Hoffman & Rehder, 2010).

However, optimizing one's attention to the current category-relevant dimension may result in learning to ignore the category-irrelevant dimension, which results in learned inattention to the irrelevant dimension (Kruschke & Blair, 2000). Therefore, if a new to-be-learned category has a category-relevant dimension that was previously irrelevant, learning may become more difficult, which represents a cost of selective attention. For example, when learning to distinguish meerkats from prairie dogs, which again look

very similar, the shape of the ears is one of the good dimensions to look. However, if one has previously learned how to distinguish Huskies from Malamutes, where eyes were attended and ears were ignored, learning to attend to the once-ignored ears would be hindered.

The close link between attention optimization and the cost of selective attention has been demonstrated in previous research (e.g., Hoffman & Rehder, 2010). In their study, participants were given either a supervised classification task (e.g. classifying a stimulus into category A or B) or a supervised inference task (e.g. inferring the missing feature of a stimulus that belongs to a certain category) and their eye movements were recorded. Since the classification task (e.g. focusing on the color of the eyes to classify Huskies and Malamutes) required attention optimization to the relevant dimension, results showed cost of selective attention when learning a new category. On the other hand, since the inference task (e.g. figuring out whether a Malamute has blue eyes or brown eyes) does not require attention optimization, the cost did not occur when learning the next category. Therefore, the study showed that (a) the characteristics of the task affect allocation of attention and (b) when attention optimization occurred, the cost of selective attention also followed.

Although attention may be affected by the characteristics of the task (i.e., classification vs. inference) it can also be affected by category structure. Categories that have multiple correlated dimensions (or statistically dense categories) may be learned without selective attention, whereas categories that have few relevant dimensions (or statistically sparse categories) may require selective attention (Kloos & Sloutsky, 2008; Sloutsky, 2010). For example, when learning the category dog, many dimensions are relevant (e.g. nose, fur, four-legs, etc.) and therefore it is relatively easy to learn. However, when learning abstract concepts such as friction, very few dimensions are relevant among many irrelevant dimensions (e.g. a car trying to stop at the red light and a person trying to open a jar both shows friction). Therefore, to learn a sparse category one has to "selectively attend" to the relevant dimension among many other irrelevant dimensions.

Finally, the deployment of selective attention may be also affected by learning regime. Since supervised learning provides information about the relevant dimension, it is more likely to recruit selective attention than unsupervised

learning (Kloos & Sloutsky, 2008). Kloos & Sloutsky (2008) showed that sparse categories could largely benefit from supervision, while it could sometimes hinder dense categories. Since selective attention filters irrelevant information and allocate attention to the relevant on information (Kruschke, 2001; Mackintosh, 1975), trying to attend to multiple correlated information (i.e. dense categories) could be harder than attending to a few.

In the current study, we examined the effects of category structure on selective attention in the course of category learning. In all experiments, a supervised category learning task was used while the participants' eye movements were recorded. Moreover, cost of attention and attention optimization were observed to infer the attentional mechanism in category learning.
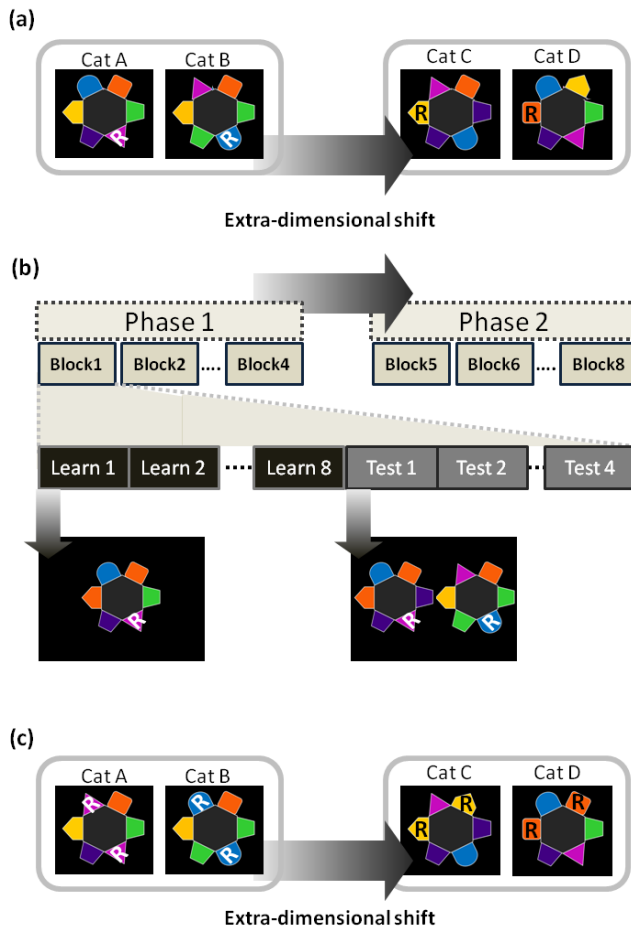

Figure 1. Description of the stimuli structure and experimental design. (a) stimuli used in Experiment 1 – sparse category, (b) experimental design of Experiment1, and (c) stimuli used in Experiment 2 – dense category (note that 'R' represents the location of the relevant dimension in each exemplar which was not visible to the participants)

Experiment 1

Experiment 1 examined the cost of attention when an extra-dimensional shift occurred between two sparse categories with supervision. As shown in previous studies, extra-dimensional shift maximizes cost of attention, therefore making it easy to observe the attentional dynamics during category learning (Hall, 1991; Hoffman & Rehder, 2010).

Methods

Participants Thirty-three adults with normal or corrected to normal vision participated in the experiment for course credit. An additional 8 participants were excluded from the analysis due to not exceeding the learning criterion (see Procedure).

Stimuli Flower-like artificial categories were used in the experiment (see Figure 1a). Each exemplar had a gray hexagon in the middle with six colored shapes on every side. Among the six colored shapes, five changed their color/shape in a binary fashion, whereas one was constant, serving as a category relevant dimension. Therefore, there were 32 unique stimuli for each category with two categories having the relevant feature on the right-bottom side of the hexagon (i.e., category A: purple triangle, category B: blue semi-circle) and two categories having the relevant feature on the left side of the hexagon (i.e., category C: yellow pentagon, category D: orange square). Therefore, the relationship between A or B and C or D was an extra-dimensional shift.

Procedure The experiment had 2 phases and in each phase there were 4 blocks. Within each block there were 8 learning trails followed by 4 test trials. After the first 4 blocks (Phase 1), unknown to the participants, the category had an extra-dimensional shift (see Figure 1b). Therefore if the first half of the blocks were presented with category A, the second half of the blocks were presented with category C in the learning trials. In the learning trials, exemplars were presented for 1.5 seconds, one at a time in the middle of the screen. At the beginning of each block, participants were told that they would see flowers that have one common feature they had to find, which served as a supervision signal.

In the test trials two category exemplars were presented side by side until the participant made a response. One exemplar was a novel exemplar from the category that was used in the learning trials. The other exemplar was a new category where the relevant feature was in the same dimension as the learned category but had a different feature (e.g. Cat A and Cat B in Figure 1a). Participants were told to choose the exemplar that they thought was a member of the category they saw in the learning trials by pressing a left or right response button. When the response was made, the stimuli disappear without any feedback. Also before each learning and test trial, a fixation point (i.e. red cross) was presented on a random-dot background, and the participants were told to look at the fixation to proceed with the experiment. Moreover, a Tobii T60 eye tracker was used to

collect eye gaze with the sampling rate of 60Hz during the whole experiment.

Results

Before analyzing the data, participants who did not learn the first category were excluded. To be considered as a learner one had to have 3 correct responses out of 4 test trials in the last block of Phase 1 (i.e. block 4). To determine whether a cost was incurred, accuracy, reaction time, and eye gaze data were analyzed by block. Especially by comparing the blocks before and after the unknown category switch (i.e. block 4 vs. block 5).

The overall accuracy for the test blocks was .90, $SD = .21$ (Phase 1: $M = .92$, $SD = .18$, Phase 2: $M = .87$, $SD = .23$), with all test trials being significantly higher than chance performance, $p < .001$ (see Figure 2a). Results of a 2 × 4 (Phase × Block) within-subjects ANOVA conducted on accuracy scores at test showed a main effect for Block, $F(2.3, 73.61) = 8.14$, $p < .001$, indicating that accuracy differed by block, but there was no significant main effect for Phase or a interactions ($ps > .05$). Moreover, a significant cost of attention was demonstrated between the last block of learning phase 1 (block 4) and the first block of learning phase 2 (block 5) by a significant decrease in accuracy from block 4 to block 5, $t(32) = 5.07$, $p < .001$.

Before analyzing the reaction time (RT), all incorrect responses were excluded, and for each individual the median RT for each block were used in the analysis. The mean reaction time for all test blocks was 1160 ms, $SD = 892$ ms (Phase 1: $M = 1199$ ms, $SD = 922$ ms, Phase 2: $M = 1121$ ms, $SD = 863$ ms) (see Figure 2b). A 2 × 4 (Phase × Block) within-subjects ANOVA conducted on RT showed a main effect for Block, $F(1.77, 54.91) = 9.58$, $p < .001$, but there was no significant main effect for Phase or a interaction ($ps > .05$). Statistical difference between block 4 and block 5 were also found, $t(32) = 2.78$, $p < .005$, demonstrating a cost of attention.

Eye gaze data were also analyzed for each block by calculating the weighted proportion of looking to the relevant spatial dimension. This value was calculated by taking looking time (fixation) to the relevant features divided by looking time (fixation) to the irrelevant and relevant features combined. However, since there was greater spatial area for irrelevant features (5 shapes) than the relevant features (1 shape), looking time to the relevant features was multiplied by five to equate the spatial area. Therefore, .50 in the analysis represents an equal amount of looking to the relevant and irrelevant features at a given block. Fixations were calculated by using an I-DT algorithm with a minimum duration threshold of 100 ms and a dispersion threshold of 1° of visual angle (Salvucci & Goldberg, 2000).

The overall weighted proportion of looking to the relevant dimension was for all test blocks was .63, $SD = .30$ (Phase 1: $M = .63$, $SD = .30$, Phase 2: $M = .64$, $SD = .31$). All blocks except the first blocks in each phase (i.e. block1 and block 5) showed a significantly higher proportion of looking

to the relevant spatial dimension (paired t-test, $ps < .05$). A 2 × 4 (Phase × Block) within-subjects ANOVA only showed a main effect for Block, $F(2.68, 80.37) = 5.59$, $p < .001$. Moreover, a marginal drop was demonstrated after block 4, which indicated a cost of attention, $t(30) = 1.83$, $p = .07$ (see Figure 2c).

In sum, both behavioral and eye gaze patterns indicated a cost of attention for participants who learned the first category. Both phases showed an evidence of attention optimization (i.e. increased accuracy, decreased RT, and increased looking time to the relevant dimension). The indication of attention optimization followed by a cost of attention was evident even though supervision was not provided as strong as in previous studies. (Note that explicit feedback was given after every trial in Hoffman & Rehder (2010)).
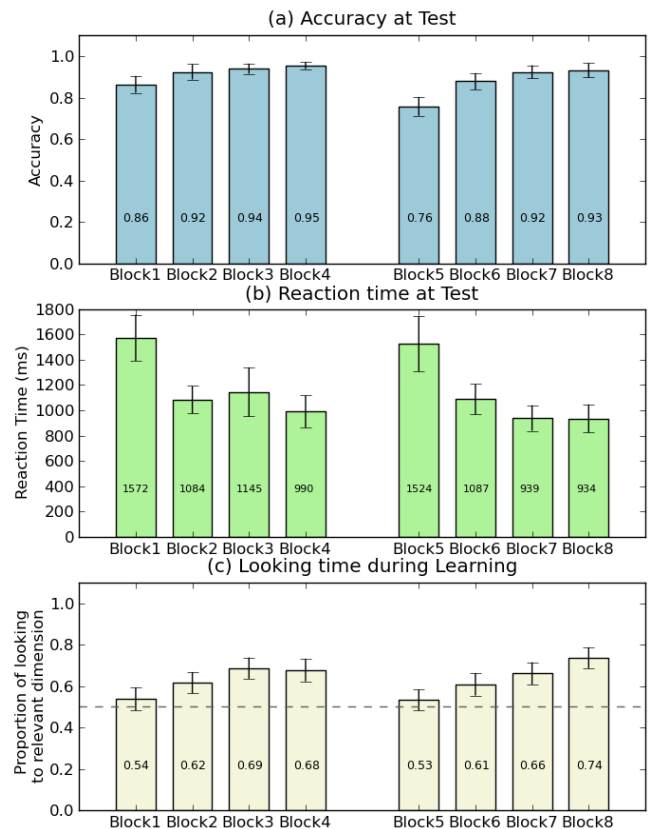

Figure 2. Results from Experiment 1. (a) accuracy at Test, (b) reaction time at Test, and (c) looking time during Learning. The proportion of looking to the relevant dimension are weighted values in that the dotted line at .5 indicate chance level of equally looking to the relevant and irrelevant dimensions. Note that all error bars represent +/- one standard error.

Experiment 2

Experiment 2 examined the cost of attention when an extra-dimensional shift occurred between two dense categories with supervision.

Methods

Participants Forty-two adults with normal or corrected to normal vision participated in the experiment. In addition, one participant was excluded from the analysis due to not exceeding the learning criterion.

Stimuli & Procedure The stimuli and procedure were identical to Experiment 1 except that dense categories were used. In contrast to sparse categories, dense categories had two category-relevant spatial dimensions instead of one (see Figure 1c). For category A and B, in addition to the bottom-right relevant dimension, the upper-left location had a constant shape/color as the bottom-right location had. For category C and D, in addition to the left location, the upper-right location had a constant shape/color identical as the left location.

Results

The overall accuracy for the test blocks was .98, $SD = .11$ (Phase 1: $M = .97$, $SD = .12$, Phase 2: $M = .98$, $SD = .11$), with all test trials being significantly higher than chance performance, $p < .001$ (see Figure 3a). A 2×4 (Phase $\times$ Block) within-subjects ANOVA did not show any main effect or interactions ($ps > .05$). Moreover, there was no significant difference between block 4 and block 5, indicating the absence of cost.

The mean reaction time for all test blocks was 838 ms, $SD = 528$ ms (Phase 1: $M = 860$ ms, $SD = 385$ ms, Phase 2: $M = 838$ ms, $SD = 528$ ms) (see Figure 3b). A 2×4 (Phase $\times$ Block) within-subjects ANOVA with RT only showed a main effect for Block, $F(2.26, 90.36) = 6.86$, $p < .001$. Also, the difference between block 4 and block 5 was not significant ($p > .05$).

In a dense category, there were two relevant dimensions and four irrelevant dimensions. Therefore, the weighted proportion of looking to the relevant dimension was calculated by multiplying two to the numerator instead of five as in Experiment 1. The overall weighted proportion for all learning blocks was .65, $SD = .23$ (Phase 1: $M = .63$, $SD = .23$, Phase 2: $M = .65$, $SD = .23$). All blocks showed a significantly higher proportion of looking to the relevant spatial dimension, paired t-test, $ps < .005$ (see Figure 3c). A 2×4 (Phase $\times$ Block) within-subjects ANOVA did not show any main effects or interactions, $ps > .05$. Also, a significant drop was not found between block 4 and 5, $p > .05$.

The results show no evidence of cost for the looking time data. Also there was no evidence of attention optimization (i.e. increased looking to the relevant dimension). However the accuracy is very high compared to the sparse condition, indicating that learning the dense category was easier than learning sparse category. Therefore it is possible that attention optimization occurred quickly, and the cost of attention was weak early in the block.
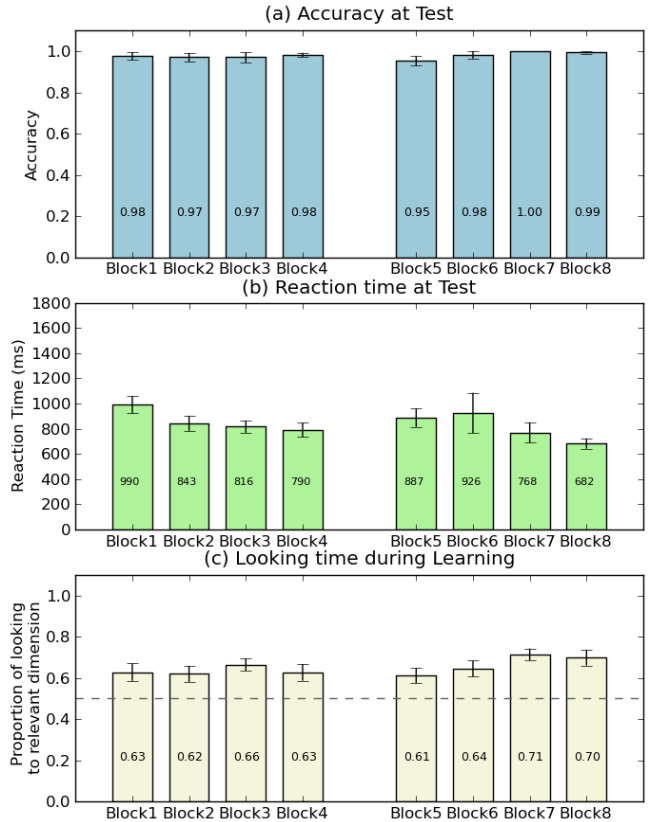

Figure 3. Results from Experiment 2. (a) accuracy at Test, (b) reaction time at Test, and (c) looking time during Learning. The proportion of looking to the relevant dimension are weighted values in that the dotted line at .5 indicate chance level of equally looking to the relevant and irrelevant dimensions. Note that all error bars represent +/- one standard error.

To capture the early attention optimization in block 1 a moving window of 3 trials were used to calculate the proportion of looking to the relevant dimension, instead of using the whole block. Then a one-sample t-test was conducted against the chance level of .5. Results show that attention optimization occurred around the window 3, which would be around the 4th trial and lasted throughout the block (see Figure 4a). The same method could be applied to Block 5 where the second category was introduced. Results show that attention optimization occurred around the window 3, which would be around the 4th trial (see Figure 4b).

On the other hand, the cost of attention could be captured by comparing the last trial of block 4 and the first trial of block 5 instead of comparing the whole block. Results showed marginally significant drop from the last trial of block 4 ($M = .59$, $SD = .44$) to the first trial of block 5 ($M = .43$, $SD = .36$), $p = .068$, indicating a cost of attention.

In sum, dense categories were learned quicker than the sparse categories (faster attention optimization), and the cost of selective attention was weaker.
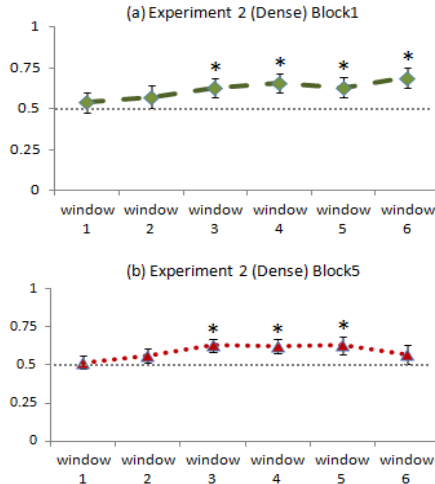
(a) Experiment 2 (Dense) Block1

(b) Experiment 2 (Dense) Block5

Figure 4. Attention optimization of block 1, block 5 in Experiment 2. Note that the asterisks represent $p < .05$, and all error bars represent +/- one standard error.

Then what would have made dense categories have lesser cost and stronger attention optimization? One possibility is that since dense categories have multiple category-relevant dimensions, attention allocation is much more distributed than sparse categories. Therefore, with limited amount of attention there will be smaller attention allocated to a dimension in the dense categories than in the sparse categories (Sutherland & Mackintosh, 1971), which would lead to an easier/faster attention shift to a newly relevant dimension. On the other hand, it could also be possible because dense categories have more category-relevant dimensions, and thus there is a higher probability of spotting a relevant dimension. In this case, one could perfectly learn the dense categories with attending only one dimension instead of both.
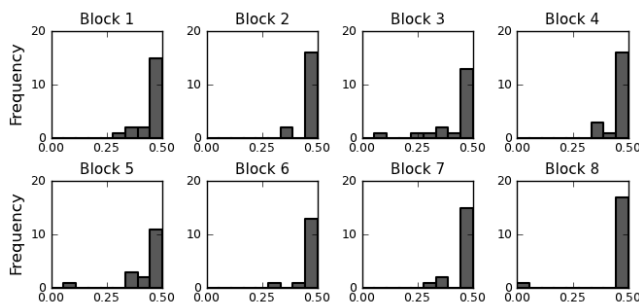

Figure 5. The distribution of looking time between the two category-relevant dimensions in Experiment 2. Values closer to 0 indicate looking equally to the two relevant dimensions, whereas values closer to 1 indicate looking to only one dimension in a trial.

To investigate the latter possibility, the distribution of looking time between the two relevant dimensions was calculated. For each trial, the proportion of looking to one of the dimensions was calculated, where .5 represents equal looking to both dimensions. Then the absolute difference from .5 was taken. Therefore, the value close to .5

represents looking to only one dimension, and 0 represents looking to both dimension. Figure 5 shows the calculated values across subjects by block. Results indicate that subjects relied on a single dimension in most of the trials when learning the dense categories.

General Discussion

The current study manipulated category density in the course of supervised category learning. Results show that even though supervision was weaker than in previous studies using sparse categories, attention optimization and cost of attention were observed during category learning (Experiment 1). Moreover, the dense categories were learned faster than sparse categories, and even with a stronger attention optimization, dense categories (Experiment 2) had a weaker cost of attention.

In Experiment 1, sparse categories were learned with weaker supervision than in previous studies using similar sparse categories. Note that when the sparse categories used in the current experiment were presented without supervision, participants failed to learn them (Yim, Best, & Sloutsky, 2011). Supervision in the current experiment consisted of a hint that there is one dimension that is consistently relevant. However, the majority of participants learned the category. Also compared to previous studies where feedback was given on every trial (Hoffman & Rehder, 2010; Rehder & Hoffman, 2005), supervision here was only given at the start of each block. However, attention optimization and cost of attention were observed.

First, attention optimization should be closely related to the specific supervision signal. Category learning has mainly assumed that error signals from feedback mediates selective attention (Blair, Watson, & Meier, 2009; Kruschke, 2001). However, the current task does not provide any feedback. A possible explanation would be that the supervision helps reduce the hypothesis space for the participants. Although knowing that there will be only one relevant dimension does not provide direct error signal, it drastically reduces the hypothesis space of possible category-relevant information. Although the effects of supervised and unsupervised learning on category formation has been discussed (Gureckis & Love, 2003; Love, 2002), the effects of various kinds of supervision has not been investigated systematically, which should be examined in future research.

Second, although it is known that attention optimization is a precursor of cost of attention, it is possible that the greater cost in the current study originates from the difference of density between the current and previous research. The stimuli in Hoffman & Rehder (2010) had 2 out of 3 irrelevant dimensions whereas the current study has 5 out of 6 irrelevant dimensions. The sparser the category is the harder it would be to learn the relevant dimension. However, once selective attention is engaged, the cost would be greater for sparser categories. This is because there are more irrelevant dimensions in a sparser category, which means that there will be more unattended dimensions

during learning (i.e. learned inattention). Therefore, when an extra-dimensional shift occurs, the probability of figuring out a newly relevant dimension among the previously irrelevant dimensions will be lower than in a less sparse category. Although it is not possible to directly examine this hypothesis from the current study, the relationship between category density and cost of attention could be examined with controlling the amount of attention optimization through manipulating the number of irrelevant dimensions.

In Experiment 2, most of the participants optimized to one dimension instead of distributing their attention to all relevant dimensions (see Figure 5). Although the categories used in the current study are deterministic and do not require an information integration process (Ashby, Alfonso-Reese, Turken, & Waldron, 1998), there is evidence that adults distribute their attention to all relevant dimensions when learning dense categories that had a similar category structure as the current one (Kloos & Sloutsky, 2008). One main difference between the previous study and the current study is the presentation time during learning. In Kloos & Sloutsky (2008), participants observed the category exemplars in a self-paced maner, whereas the current study presented the exemplars for 1.5sec. Since the category could be learned by using both distributed and non-distrubuted attention, it is highly possible that the fast presentation time leaded the participants to attend to only one dimesion.

Finally, the results may have implications for understanding the development of category learning. Since it is known that children gradually gain the ability to selectively attend (Hanania & Smith, 2010), it would be hard for them to learn sparse categories, which requires the ability to selectively attend to a small number of category-relevant dimensions. Therefore, the role of supervision would be crucial for learning spare categories early in development. If the interaction among the category structure, learning regime, and category learning is well established, it would help to understand the developmental trajectory of category learning.

Acknowledgments

References

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

Blair, M. R., Watson, M. R., & Meier, K. M. (2009). Errors, efficiency, and the interplay between attention and category learning. *Cognition*, *112*(2), 330–336.

Gureckis, T. M., & Love, B. C. (2003). Human unsupervised and supervised learning as a quantitative distinction. *International Journal of Pattern Recognition and Artificial Intelligence*, *17*(5), 885–901.

Hall, G. (1991). *Perceptual and Associative Learning*. New York, New York, USA: Oxford University Press.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: towards a unified developmental approach. *Developmental Science*, *13*(4), 622–635.

Hoffman, A. B., & Rehder, B. (2010). The Cost of Supervised Classification: The Effect of Learning Task on Conceptual Flexibility. *Journal of Experimental Psychology: General*, *139*(2), 319–340.

Kloos, H., & Sloutsky, V. M. (2008). What's Behind Different Kinds of Kinds: Effects of Statistical Density on Learning and Representation of Categories. *Journal of Experimental Psychology: General*, *137*(1), 52–72.

Kruschke, J. K. (1992). ALCOVE: An Exemplar-Based Connectionist Model of Category Learning. *Psychological Review*, *99*(1), 22–44.

Kruschke, J. K. (2001). Toward a Unified Model of Attention in Associative Learning. *Journal of Mathematical Psychology*, *45*, 812~863.

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, *7*, 636–645.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, *9*(4), 829–835.

Mackintosh, N. J. (1975). A Theory of Attention: Variations in the Associability of Stimuli with Reinforcement. *Psychological Review*, *82*(4), 276–298.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–61.

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*(1), 1–41.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the symposium on Eye tracking research & applications - ETRA '00*, 71–78.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*(13), 1–41.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*, 1244–1286.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.

Yim, H., Best, C. A., & Sloutsky, V. M. (2011). Cost of Attention as an Indicator of Category Learning. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1724–1729). Austin, TX.