Title
Is focusing enough in category learning?

Permalink
https://escholarship.org/uc/item/7600g9n2

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Yim, Hyungwook
Yoon, Sejin

Publication Date
2024

Peer reviewed

Is focusing enough in category learning?

Hyungwook Yim (hwyim@hanyang.ac.kr)
Department of Cognitive Sciences, Hanyang University, Seoul, Republic of Korea

Sejin Yoon (ysj2021@hanyang.ac.kr)
Department of Cognitive Sciences, Hanyang University, Seoul, Republic of Korea

Abstract

We examined whether selective attention, which is mainly theorized as the ability to focus on the category-relevant dimension, is a sole construct in understanding category learning. As the attention literature dissociates selective attention into focusing and filtering, we argue that filtering is another component that should be considered to fully understanding category learning. In the study, we provide an experimental paradigm that can dissociate filtering from focusing. By utilizing the paradigm along with collecting individual attention control measures, we show that filtering is related to the ability to inhibit irrelevant information. We also present that the current computational models that incorporate selective attention only as an ability to focus can not explain the results from the current study.

Keywords: category learning; selective attention; distraction; executive function; individual difference

Introduction

Selective attention has been an important component in theorizing the mechanisms underlying category learning (e.g., Shepard, Hovland, & Jenkins, 1961; Kruschke, 1992; Nosofsky, 1986). Selective attention enables one to attend to the category-relevant dimension, and filter out the category-irrelevant dimensions. Therefore, selective attention not only aids efficient learning, but also helps future generalization as category-relevant dimensions are prioritized when processing the stimulus (Mackintosh, 1965).

Computational models mathematically formalize selective attention by assuming that the exemplars (or the prototype) are represented in a multi-dimensional psychological space (Love, Medin, & Gureckis, 2004; Kruschke, 1992; Nosofsky, 1986; Medin & Schaffer, 1978; Smith & Minda, 1998). Then selective attention acts as stretching the dimension that is attended. The stretched dimension, therefore, becomes easier to discriminate than other dimension, and lead better learning and generalization.

Selective attention has also been an important keyword for explaining the development of category learning. As the functionality of selective attention has been linked to the maturation of the prefrontal cortex (Squire, Noudoost, Schafer, & Moore, 2013; Diamond, 2002), and since the maturation of the prefrontal cortex has been known to be relatively slow (e.g., Kolk & Rakic, 2022), developmental studies have provided crucial insights into how selective attention effects category learning. For example, Best, Yim, and Sloutsky (2013) examined infants and adults in a category learning task using novel visual categories while tracking their eye gaze. Although both groups learned the categories, only adults selectively attended to the category-relevant dimension and optimized their attention as they learned the category. Infants, in the other hand, did not soley attend to the category-relevant dimension, but rather *distributed* their attention across dimension.

A study by Deng and Sloutsky (2016) supports the idea of *distributed attention* as a hallmark of low functioning selective attention during category learning. In their study, four-to-five year-olds and adults were examined in a novel visual category learning task. With a comparably more mature prefrontal cortex, children do selectively attend to the category-relevant dimension as adults do. However, children still show a *distributed attention* pattern as infants do. As a consequence when given a surprise memory test about the category examplars after the learning task, children retain the category-irrelevant information of the examplars (information that should be filtered out in order to learn the category efficiently), while adults do not.

Additionally, not attending to the irrelevant dimension (i.e., learned inattention) affects learning future categories when the irrelevant dimension becomes relevant. For example, Hoffman and Rehder (2010) examined adults in a classification task while tracking their eye gaze. Unbeknownst to the participants the category-relevant dimension changed midway as they learned the initial category. Results showed that participants struggled to learn the second category if they attended to the category-relevant dimension in the first category. Moreover, Best et al. (2013) examined infants and adults in a similar experiment, where they find that adults but not infants struggle to learn the second category. Therefore, the ability to attend to the category-relevant dimension is helpful for highlighting the category-relevant dimension. However, at the same time, it decreases the relative sensitivity of the unattended dimension (or the category-irrelevant dimension), and elicits a *Cost of selective attention* when learning new information.

What is selective attention in category learning?

Although previous research has shown that selective attention has a major effect in category learning, the majority of the studies have defined selective attention as the ability to *focus* on the category-relevant dimension. Computational models of category learning also formalize selective attention as the ability to flexibly shift one's attention (e.g., λ_α parameter in

ALCOVE; Kruschke, 1992).

However, focusing may not be the sole construct of selective attention. Previous studies in attention and executive function has supported that attention is a multi-component construct (e.g., Miyake et al., 2000). Neuroimaging studies show that there two separate components of selective attention, where one is the ability to focus on the relevant information while another is filtering out irrelevant dimension (Bridwell & Srinivasan, 2012). Moreover, developmental studies show that focusing and filtering may be separate constructs by showing that the ability to filter matures later in development (Unger & Sloutsky, 2023). Similarly, developmental studies in category learning have provided evidence of the low filtering ability through a behavioral pattern called *distributed attention* (e.g., Best et al., 2013; Deng & Sloutsky, 2016). In these studies, infants or children learn an artificial category while showing a gaze pattern which does not efficiently focus on the category-relevant dimension, but is *distributed* across all dimensions.

Although the term *distributed attention* implies that attention is controlled intentionally (or endogenously), given that the distributed attention pattern is usually observed in infants and children, the pattern may not be rooted in the ability to intentionally distribute one's attention (e.g., Chong & Treisman, 2005). Therefore, the term *distracted attention* may be a more precise term to describe the behavioral patterns of infants and children during category learning, and a more passive selective attention component (such as *filtering*) may be the ability that is responsible for the behavioral pattern.

In the current experiment, therefore, we (1) developed an experimental paradigm that can dissociate the focusing component and filtering component during category learning, (2) examined how the filtering component may interact during category learning, and (3) tried to specify which attentional control (or executive function) component is responsible for the filtering ability using an individual differences approach. For the main category learning task, we used the cost of selective attention task (Best et al., 2013; Hoffman & Rehder, 2010), where the category relevant dimension changed mid way during the task unbeknownst to the participant. The paradigm has been known to capture the a signature of attention optimization during category learning (Yim, Best, & Sloutsky, 2011). We additionally manipulated the presentation sequence of the examplars to control trial-to-trial bottom-up distraction by changing the number of features that changed on the next trial. Finally, we measured each participant's attentional control (or exeutive function Engle & Kane, 2004; Miyake et al., 2000), and process speed. For attentional control we used the anti-saccade task to measure inhibition, symmetry span task for working memory span, and the color-shape switch task for switching ability. For process speed, we used the the simple reaction task, and choice reaction task, which is known to measure both process speed and some aspects of attentional control (Deary, Johnson, & Starr, 2010). Finally, we simulated a category learning model (ALCOVE; Kruschke, 1992), which assumes selective attention as the ability to focus on category-relevant information, and examined whether the model can explain our experimental results.

Experiment

The order of experiments was fixed as follows: (1) category learning task, (2) simple reaction task, (3) anti-saccade task, (4) choice reaction task, (5) symmetry span task, and (6) color-shape task.

Methods

Participants Fifty-two participants (33 females, M = 23.13yrs, SD = 2.45yrs) were recruited through an online community page of Hanyang University, Seoul, Republic of Korea. Each participant was compensated with 20,000 KRW (approximately 15 USD) for their participation. The research was approved by the Institutional Review Board at Hanyang University.

Materials, Design & Procedure

Category learning task: Participants were introduced to the task as they were observing an alien pond, which can predict the typhoon of the next day. The category stimuli were pictures of a pond with six different animals and plants displayed around the pond. Each animal/plant changed in a binary fashion rendering a total of 64 unique examplars (see Figure 1A, B). Participants were presented with an image of the pond one at a time and were asked to classify whether the typhoon would come or not. The participant learned the category until they achieved more than 14 trials in a block of 16 trials or when they reached 15 blocks, where each block included 16 trials. Unbeknownst to the participant the category-relevant dimension changed once they learned the category (more 14 correct trials in a block of 16 trials), which was designed to generate the 'cost of selective attention'. Therefore, participants learned two categories using the same stimuli set. We will refer the first category set the AB Learning set, and the second category set the CD Learning set. The category-relevant dimension (a location around the pond) was randomly selected for each learning set, with a constraint that the two dimensions are not adjacent to each other. In a trial, a fixation cross was displayed for 1sec in the center on a black background. Then the stimulus was presented on a black background until the participants made a decision using the button box ("Yes (a typhoon will come)" or "No (a typhoon will not come)"). The left/right side of the correct response was counterbalanced. After the decision has been made, the stimulus disappeared and the feedback("Correct", or "Incorrect") was presented on the screen for 1sec. Most importantly, there were two conditions in the experiment regarding the presentation sequence of the exemplars. In Condition 2, two dimensions changed its feature in the following trial, while in Condition 4 four dimensions changed its feature. The participants were

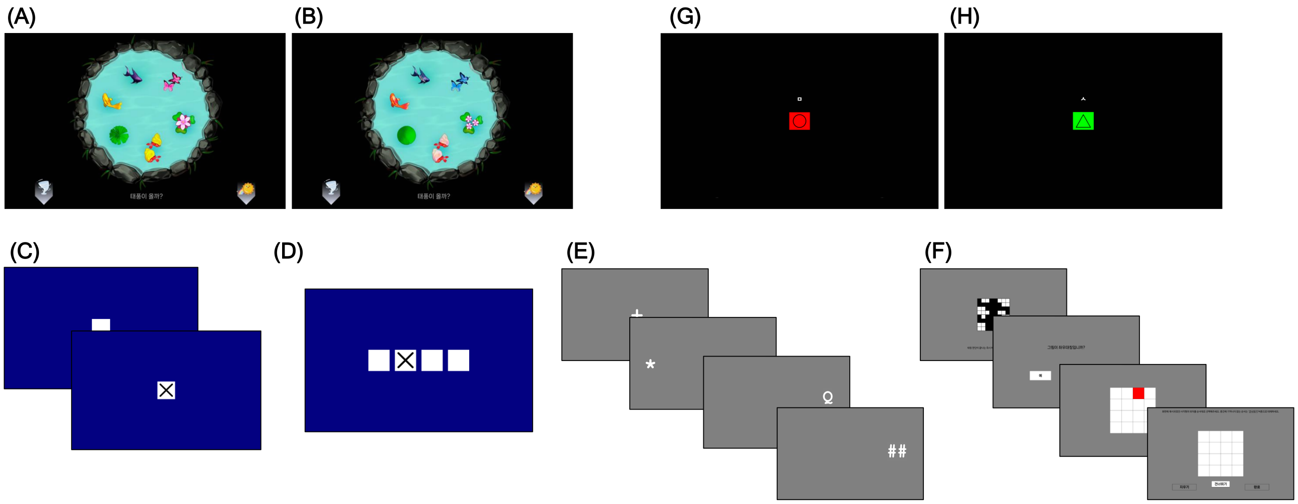

Figure 1: Stimuli and Tasks used in the Experiment. (A) An image of the alien pond that was used as a category exemplar (prototype 1), (B) An image of the alien pond that was used as a category exemplar (prototype 2), (C) Simple reaction task, (D) Choice reaction task, (E) Anti-saccade task, (F) Symmetry span task, (G) Color-shape switch task (when the cue was for the shape and the target was a red circle), (H) Color-shape switch task (when the cue was for the color and the target was a green triangle)

randomly assigned to one of the conditions. Each block had an equal number for exemplars that predicted the typhoon and that did not. The sequence of the trials were also constrained, where the same category-type did not appear more than three time in a row. The category learning task and the following tasks were presented on a 27 inch monitor, and the experiment was controlled by Psychtoolbox3. For the category learning task, an Eyelink 1000 Plus (SR Research Ltd.) was used for eye-tracking with the sampling rate of 1000Hz (Kleiner, Brainard, & Pelli, 2007; Brainard, 1997; Cornelissen, Peters, & Palmer, 2002). Moreover, the luminance was controlled to 12 lux at the point of view of the participants in front of a gray (RGB: 220, 220, 220) screen.

Simple Reaction task: Both simple reaction task and choice reaction task were adapted from Deary et al. (2010). The letter 'X' appeared in the square with random intervals between 1 to 3sec and remained until the subject responded. Participants were asked to press the space key as quickly as possible when they saw the 'X' on the screen (see Figure 1C). There were eight practice trials and 20 main trials.

Choice Reaction task: In a trial, four white squares were presented horizontally in the center of the screen until the participant's response. The four squares were assigned to the letter x, c, b, and n on the keyboard. Participants were asked to press the corresponding key as quickly as possible when black 'X' was shown on the square (see Figure 1D). The inter-trial-interval ranged from 1 to 3sec, and there were 8 practice trials followed by 40 main trials.

Antisaccade task: The anti-saccade task was adapted from Draheim, Tshukara, and Engle (2023). In this task white fixation cross appears on the gray background for either 1sec or 2sec (see Figure 1E). Then a white asterisk(*) appeared at 12.3° visual angle on the right or left side of the screen for 300ms. Then a letter (O or Q) appeared on the opposite side of the asterisk for 100ms and was rapidly be masked by '##'. Participants sitting 60cm from the monitor, and were asked to quickly look the opposite direction of the asterisk, and decide whether the letter was a Q or O within 5sec. Feedback was given with a cyan 'O' and magenta 'X' for 1sec. There were 16 practice trials followed by two blocks, which had 36 trials each.

Symmetry Span task: The symmetry span task was based on Unsworth and Brewer (2009). Participants were asked to recall the sequence of the positions of the red squares in 4×4 matrix while the distraction task were given between the presentation of red squares (see Figure 1F). In the distraction task, the participants were asked to judge if the black and white pattern drawn on the 16×16 matrix is left-right symmetry or not. There were three practice phases to help participants understand the task. In the first practice phase, only red squares appeared one after another for four sets with set-size two and three for twice each. In the second practice phase, only symmetry judgements were executed with 15 trials, where feedback given for 1s right after the response. The mean plus 2.5 standard deviation of the reaction time in the second phase was use as a time limit for the main following phases. In the third practice phase, the participant had to solve the symmetry quiz their personalized time limit

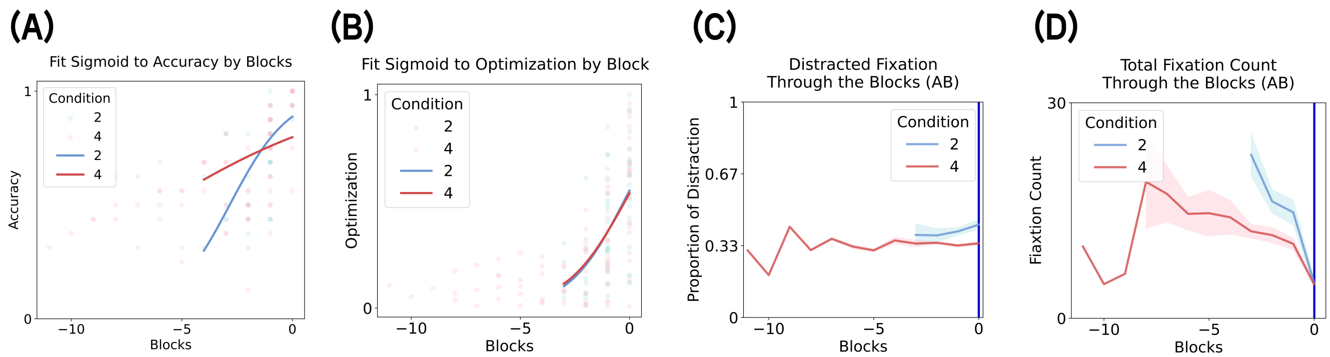

Figure 2: Category learning results plotted by a lag plot were all data was shift to the point where learning was accomplished (lag 0 point). (A) accuracy data by block, (B) eye gaze (fixation) data by block, where the y-axis represents proportion of fixation to the category-relevant dimension calculated by dividing the number of fixations to the category-relevant dimension by the total number of fixations. (C) proportion of distraction across block, and (D) total fixation across block.

(i.e., distractor task), then a red square was presented for the participant to remember the location for 650msec (i.e., main span task). Participants were informed that there was a time limit for symmetric quizzes, and if the subject failed to move on to the next page for symmetric quiz responses in time, they were immediately shown the red square. Three sets with set-size 2(two pairs) were executed as the third practice and participants had to answer correctly at least 85% in the symmetry quiz. In the main trials, the procedure was identical to the last practice phase except that the set-size ranged from two to five, and the trials consisted of three trials of each set-size. This made a total of 42 symmetry quizzes and 42 square tasks.

Color-Shape Switch task: The color-shape switch task which was adapted from Miyake et al. (2000). In a trial, a cue letter ㅅ or ㅁ (each corresponding to the first consonant for the word 'color' and 'shape' in Korean) appeared for 150ms, followed by a presentation of the cue and a shape below the cue (see Figure 1G, H). The shape was either a triangle or a circle in a color of green or red. The participants were asked to decide whether the color was green or red if the cue was ㅅ, and whether the shape was a triangle or a circle if the cue was ㅁ using the Z and M keys on the keyboard. The interval between trials was 2.5s and feedback was provided for 1sec after each trial. There were two blocks, each with 12 trials for the practice phase, and two blocks with 24 trials per block for the main phase. There were 24 non-switch trials and 23 switch trials, and the sequence was controlled so that the switching or non-switching trials did not appeared more than four times in a row.

Behavioral Results

Category learning task

We first examined how many participants learned the category in each condition (the first AB Learning set) within 15 blocks. There were 24 out of 25 participants in Condition 2 who learned the first category, while 22 out of 27 participants in Condition 4 who learned the first category, where a logistic regression showed a marginal difference between the conditions ($\chi^2(50) = 2.91, p = .087$). When comparing the total trials to learn the first category, Condition 4 ($M = 61.23 trials, SD = 51.05$) took longer to learn than Condition 2 ($M = 42.67 trials, SD = 39.25$) with showing a statistically marginal difference (one-tailed independent t-test, $t = 1.39, p = .086$). When we included the non-learners with assuming that they required a full 15 blocks to learn the category, the results showed a statistically significant difference ($t = 2.20, p = .016$).

Cost of selective attention was evaluated to see if there were any differences in selectively attending to the relevant dimension between the two Conditions by comparing the performance in the second CD Learning set. There was only one participant who did not learn in Condition 2, and two in Condition 4 ($\chi^2(44) = .46, p = .50$). The number of trials to learn the CD Learning set also did not differ statistically (Condition 2: $M = 45.96 trials, SD = 34.66$; Condition 4: $M = 42.14 trials, SD = 23.73$; $t = .422, p = .68$). As the difference in the cost of selective attention was not evident, we focused our analyses on the 1st learning (AB Learning set). All analyses hereafter, we only included the data from the learners.

To investigate the learning trajectory we first examined the learning accuracy by block. Instead of looking at the data in a chronological order, we re-arranged the data backwards from the point where learning was accomplished and generated a backward lag plot (see Figure 2A). Then, we fit the accuracy data with a sigmoid function as follows: $P(correct) = 1/(1 + exp(-k \times (block - x0)))$, where $P(correct)$ is the accuracy, $block$ is the number of blocks, and k, $x0$ were free parameters that controls the x-axis shift and scale of the sigmoid function respectively. A randomization task with 1,000 samples showed a statistically significant difference between the two conditions for the scale parameter (Condition 2:

$x0 = -2.84, k = .73$; Condition 4: $x0 = -5.97, k = .23$; $p_{empirical-x0} = .73, p_{empirical-k} < .001$). Using the same method as above, we also examined eye gaze by calculating the fixations on the category-relevant dimension divided by the total fixations for each block (see Figure 2B). Results did not show a statistically significant difference between the two conditions (Condition 2: $x0 = -.15, k = .85$; Condition 4: $x0 = -.16, k = .74$; $ps_{empirical} > .21$).

To examine the amount of distraction during learning, we examined the proportion of fixations to the changed features in each trial. As there are two features that change every trial in Condition 2, and four in Condition 4, we expect that a greater amount of distraction will be generated by Condition 4. However, interestingly, results show that there was a statistically greater amount of distraction for Condition 2 than for Condition 4 (Condition 2: $M = .38, SD = .06$; Condition 4: $M = .32, SD = .02$; Independent t-test; $t = 4.11, p < .001$; the proportion of distractions for Condition 4 was weighted by .5 as there are twice as many dimensions that change in Condition 4).

We also examined the number of fixation across blocks. If the participants selectively focused on the category-relevant dimension, there would be lesser amount of fixation across blocks, and the number of fixations would decrease rapidly. Results are shown in Figure 2D, where the average number of fixations per block did not differ across the conditions (Condition 2: $M = 11.59, SD = 5.68$; Condition 4: $M = 10.04, SD = 3.01$; $p = .27$). However, Condition 2 showed a more rapid decrease in the number of fixations than Condition 4 as learning progressed (linear regression with BLOCK and CONDITION showed an statistically significant effect for BLOCK, $t = 8.39, p < .001$, and BLOCK×CONDITION interaction, $t = 6.71, p < .001$, but not for CONDITION, $t = 1.29, p < .20$).

Table 1: Correlation coefficients between measure of distraction and attentional control measures. Values in bold represent Pearson correlation coefficients and the values in the parenthesis represent p-values. * represents $p < .05$ and + represents $p < .1$.

	Distraction	Fixation
Color-Shape	-.09 (.56)	.27 (.094+)
Anti-saccade	.14 (.38)	-.04 (.80)
Symmetry span	-11 (.49)	-.12 (.44)
Simple reaction	-.09 (.57)	.08 (.62)
Choice reaction	.33 (.035*)	.19 (.23)

Source of distraction

In order to investigate which aspect of the attention control ability is related to the distraction during learning we examined the correlation between the two distraction measures (proportion of fixations to the changed features (Proportion of Distraction), number of fixations (Fixation count)) and the results from five attentional control tasks. For each participant's data, we took the mean reaction time for the Color-Shape task, accuracy for the Anti-saccade task, accuracy for the Symmetry span task, and median reaction times for the Simple reaction task, and Choice reaction task. Then we calculated the Pearson correlation coefficients for each pair. Results are shown in Table 1, where the values represent correlation coefficients and the values in the parenthesis show p-values. A statistically significant correlation was found between the Proportion of Distraction and the Choice reaction tasks ($p = .35$), and a marginally significant relation between the Fixation count and Color-shape task ($p = .94$).

ALCOVE Simulation

To examine whether the current experimental results (different learning patterns between Condition 2 and Condition 4) can be explained by assuming that selective attention is majorly the ability to focus on the category-relevant dimension, we simulated a category learning model. We used ALCOVE as it incorporates selective attention as an ability to flexibly focus one's attention, and that it can simulate the learning trajectory (Kruschke, 1992).

For all simulations the parameters were fixed to $C = 5$, $\phi = 2$, $\lambda_\alpha = 1.0E\text{-}2$, and $\lambda_w = 5.0E\text{-}2$. For each condition, thirty experimental trial sequences were randomly generated as in the Experiment. However, we only generated four blocks for the AB Learning set followed by four blocks for the CD (switch) Learning set, which resulted in 128 trials (8 blocks $\times$ 16 trials) for each randomly generated sequence. We modified the catlearn library (Wills, O'Connell, Edmunds, & Inkster, 2017) that implements the ALCOVE model in R, while modifying the model to accommodate discrete features instead of continuous ones (Lee & Navarro, 2002).

Results are shown in Figure 3. The top row shows accuracy results between the conditions. Both conditions show an increase in accuracy before the switch, while a decrease in accuracy when the new category was introduced as the experimental results. However, although the current experimental results show that Condition 4 required more trials to learn compared to Condition 2, the simulation results showed an identical learning trajectory. The bottom row shows how the attention weights in the model changed through learning. Overall, attention was gradually optimized in the first category (AB learning), where the relevant dimension was D1. After the switch when learning the second category (CD learning), attention gradually optimized to the relevant dimension (D4) with some lingering attention to the previously relevant dimension (D1), which shows the*cost of selective attention*. However, again, there was no difference between the two conditions regarding attention deployment.

Discussion

The current study examined whether *focusing* can be distinguished from *filtering* in category learning. We

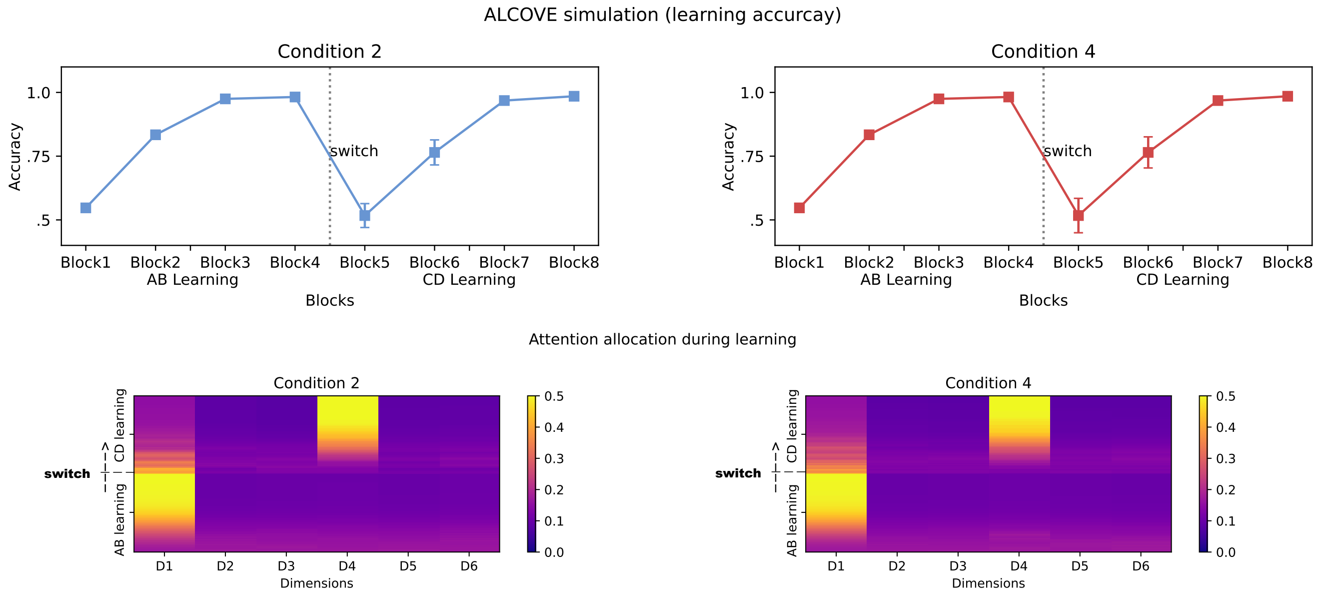

Figure 3: Simulation results using ALCOVE. The first column shows results from Condition 2, and the second column shows results from Condition 4. The top row shows mean accuracy for each block, and the bottom row shows attention allocation, which can range from 0 to 1, of the model for each trial during learning. The relevant dimension for the first learning category (AB learning) was D1, and for the second switched category (CD learning) was D4. Note that the error bars in the first row represents *SD* of the simulated results.

provided a novel experimental design that can distinguish the two components by manipulating the amount of new features that change in the following trial. Results showed that when changing more features in the following trial learning was hindered. We also showed that more distraction during the trials were related to attentional control tasks that were related to inhibition (Color-shape task, and Choice reaction task). Finally, we showed that a computational model that only formalizes focusing can not explain the behavioral differences generated by the manipulation.

Most importantly, the increasing the amount of new features in the following trial increased the length of the required trials to learn the categories. However, the shape of learning (e.g., slope of learning) did not diff across the two conditions supported by the accuracy data and eye gaze optimization data. Moreover, there was no difference in the CD Learning set, which implies that the amount of the "Cost of selective attention" was similar. Therefore, the manipulation seems to hinder the initial process of learning such as searching the relevant information, which would correspond to the definition of filtering.

On the other hand, it was interesting to observe that the two distraction measures (i.e., proportion of distraction, and fixation count) were both higher in Condition 2 than in Condition 4. One possible explanation is that one's attention in Condition 4 was stuck to a few category-irrelevant dimensions due to overflowing information. The notion corresponds to Engle's theory of attention control (Burgoyne & Engle, 2020), where attention control is categorized into 'Maintenance', and 'Disengagement'. The two term may possibly be related to 'Focusing' and 'Filtering', but further investigation should be required for a conclusion.

The correlation results provide additional information about the source of the distractions. The fact that distraction was correlated with the Choice reaction task[1], but not with the anti-saccade task implies that the ability related to distraction is not simply a low-level inhibition process. Instead it seems that the ability is related to the inhibition mechanism involved during the response process, which concur with the interpretation of the choice reaction task (e.g., Deary et al., 2010).

The simulation results apparently showed that defining selective attention as only 'focusing' can not explain the current results, which was generated by manipulating variable that would affect the filtering mechanism. Selective attention in most of the formal models of category learning only considers the category-relevant information, and can not consider category-irrelevant information. Importantly, incorporating a filtering mechanisms that also monitors category-irrelevant information would be beneficial for understanding the development of category learning as the effect of filtering is more crucial in infants and children (e.g., Unger & Sloutsky, 2023).

[1] It is noteworthy that even though Condition 2 showed more distraction than Condition 4, the individual correlations (calculated regardless of the conditions) showed that more distraction is linked to worse performances in the two tasks as both attention tasks were evaluated by reaction time.

Acknowledgments

This work was supported by the National Research Foundation of Korea funded by the Ministry of Science and ICT (No. RS-2022-00166828) to HY. We thank Jiyeon Park for data collection. Correspondence concerning this article should be addressed to Hyungwook Yim (E-mail: hwyim@hanyang.ac.kr).

References

Best, C. A., Yim, H., & Sloutsky, V. M. (2013, October). The cost of selective attention in category learning: developmental differences between adults and infants. *Journal of Experimental Child Psychology*, *116*(2), 105–19. (Publisher: Elsevier)

Brainard, D. H. (1997, January). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. (Publisher: Brill)

Bridwell, D. A., & Srinivasan, R. (2012, October). Distinct attention networks for feature enhancement and suppression in vision. *Psychological Science*, *23*(10), 1151–1158.

Burgoyne, A. P., & Engle, R. W. (2020, December). Attention Control: A Cornerstone of Higher-Order Cognition. *Current Directions in Psychological Science*, *29*(6), 624–630. (Publisher: SAGE Publications Inc)

Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*(1), 1–13.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002, November). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, *34*(4), 613–7.

Deary, I. J., Johnson, W., & Starr, J. M. (2010). Are processing speed tasks biomarkers of cognitive aging? *Psychology and Aging*, *25*(1), 219–228.

Deng, W. S., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. Stuss & R. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London, UK: Oxford University Press.

Draheim, C., Tshukara, J. S., & Engle, R. W. (2023, May). Replication and extension of the toolbox approach to measuring attention control. *Behavior Research Methods*.

Engle, R. W., & Kane, M. J. (2004). *Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control*. Retrieved from

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.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3? *Perception*, *36*, 14.

Kolk, S. M., & Rakic, P. (2022, January). Development of prefrontal cortex. *Neuropsychopharmacology*, *47*(1), 41–57.

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

Lee, M. D., & Navarro, D. J. (2002, March). Extending the ALCOVE model of category learning to featural stimulus domains. *Psychonomic Bulletin & Review*, *9*(1), 43–58.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004, April). SUSTAIN: A network model of category learning. *Psychological review*, *111*(2), 309–32.

Mackintosh, N. J. (1965). Selective Attention in Animal Discrimination Learning. *Psychological Bulletin*, *64*(2), 124˜150.

Medin, D. L., & Schaffer, M. M. (1978). Context Theory of Classification Learning. *Psychological Review*, *85*(3), 207–238.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000, August). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive psychology*, *41*(1), 49–100.

Nosofsky, R. M. (1986). Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology: General*, *115*(1), 39˜57.

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

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

Squire, R. F., Noudoost, B., Schafer, R. J., & Moore, T. (2013). Prefrontal Contributions to Visual Selective Attention. *Annual Review of Neuroscience*, *36*(1), 451–466.

Unger, L., & Sloutsky, V. M. (2023, February). Category learning is shaped by the multifaceted development of selective attention. *Journal of Experimental Child Psychology*, *226*, 105549.

Unsworth, N., & Brewer, G. A. (2009). Examining the relationships among item recognition, source recognition, and recall from an individual differences perspective. *Journal of experimental psychology. Learning, memory, and cognition*, *35*(6), 1578–1585.

Wills, A. J., O'Connell, G., Edmunds, C. E. R., & Inkster, A. B. (2017, January). Chapter Three - Progress in Modeling Through Distributed Collaboration: Concepts, Tools and Category-Learning Examples. In B. H. Ross (Ed.), *Psychology of Learning* *and Motivation* (Vol. 66, pp. 79–115). Academic Press.

Yim, H., Best, C. A., & Sloutsky, V. M. (2011). Cost of Attention as an Indicator of Category Learning. In *Proceedings of the 33th Annual Conference of the Cognitive Science Society* (pp. 1763–1768).