Title
Differential Effect of Blocked and Interleaved Study on Category Learningby Classification and Inference

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors
Yao, Yinjie
Deng, Sophia W.

Publication Date
2020

Differential Effect of Blocked and Interleaved Study on Category Learning by Classification and Inference

Yinjie Yao (sb62320@um.edu.mo)
Department of Psychology, University of Macau
Taipa, Macau S.A.R. 999078 China

Sophia W. Deng (wdeng@um.edu.mo)
Department of Psychology and The Center for Cognitive and Brain Sciences, University of Macau
Taipa, Macau S.A.R. 999078 China

Abstract

Previous research has indicated that the way of learning and the sequence of study influence how we learn and represent categories. However, most studies have focused on classification learning and it has been rarely studied how learning sequence influences inference learning. The current study attempted to address this issue. Participants learned four categories by classification or inference in both blocked and interleaved sequence. Then participants completed a transfer task and a feature prediction task. Results showed that classification learners encoded characteristic features and formed similarity-based representations in the blocked study, whereas in the interleaved study, they encoded deterministic features and formed rule-based representations. In contrast, for inference learners, the blocked and interleaved study changed their learning and representation in the same direction. In both sequences, inference learners encoded deterministic features and formed rule-based representations. These results suggest that different mechanisms are likely to be involved for inference and classification learning.

Keywords: category learning; sequence of study; inference; classification; attention; representation; human experiments

Introduction

People learn categories in a variety of ways. Some of this knowledge is taught as part of education, some is learned during daily activities, and some is acquired through extensive experience. There is a wide range of evidence showing that the way of learning (e.g., Yamauchi & Markman, 2000; Yamauchi & Markman, 1998; Markman & Ross, 2003; Anderson, Ross, & Chin-Parker, 2002; Hoffman & Rehder, 2010; Deng & Sloutsky, 2015) and the sequence in which the items are presented (e.g., Clapper, 2014; Sandhofer & Doumas, 2008) affect how we learn and represent categories. However, it has been rarely studied how learning sequence influences category learning and representations under different ways of learning. The main goal of this study is to directly address this issue.

Among various ways of learning categories, two of them-*classification* and *inference*-have theoretical implications and have been studied systematically in laboratory research (e.g., Yamauchi & Markman, 2000; Yamauchi & Markman, 1998; Markman & Ross, 2003; Anderson, Ross, & Chin-Parker, 2002). In classification learning, people predict the category membership of an item. This situation is similar to that of sorting a set of cats and dogs into two distinct groups. Whereas classification learning involves predicting category membership, inference learning involves predicting a feature. In this case, instead of determining whether an animal is a cat or a dog, people predict an unknown, missing, or unobservable feature (e.g., the type of sound the animal makes). There is evidence that classification learning and inference learning result in different allocation of attention to features (Hoffman & Rehder, 2010) and thus different representation of categories (Deng & Sloutsky 2015). For example, using eye tracking techniques, Hoffman & Rehder (2010) found that in contrast to inference learners who allocated attention to multiple dimensions, classification learners optimized attention to the deterministic dimension that distinguish between two categories. This attentional difference between classification learning and inference learning can further lead to differences in memory of features and representation of categories. For example, Deng & Sloutsky (2015) found that whereas classification learners were more likely to attend selectively to the deterministic features and to form rule-based representations of category, inference learners were more likely to attend diffusely, which resulted in similarity-based representations.

Meanwhile, research on learning sequence has shown that the sequence in which items are presented also affects how we attend to, learn, and represent categories (Sandhofer & Doumas, 2008; Zaki & Salmi, 2019; Carvalho & Goldstone, 2014, 2015a, 2015b, 2017). For example, Carvalho & Goldstone (2017) found that blocked and interleaved learning resulted in different allocation of attention to features and representation of categories. Specifically, if categories are learned blocked, learners tend to pay attention to similarities among items of the same category and encode characteristic features that are highly frequent within categories. In contrast, if categories are learned interleaved, learners tend to pay attention to differences between items and encode deterministic features that distinguish between categories.

However, most studies examining the effect of learning sequence have focused on classification learning of categories (Carvalho & Goldstone, 2014, 2015a, 2015b, 2017; Zaki & Salmi, 2019). It is less clear how sequence would affect inference learning. We aim to address the issue in this study.

Overview of The Current Study

The goal of the current study is to examine how learning sequence influences category learning and category representation under classification and inference learning. We adopted a similar design by Carvalho & Goldstone (2017). The current study consisted of three phases: learning, transfer, and feature prediction. During the learning phase, participants learned two sets of two contrasting categories of artificial creatures by either predicting the category membership of an item (i.e., classification learning) or predicting a missing feature, precisely a deterministic feature (i.e., inference learning), and they were provided with feedback. Each category consisted of deterministic features distinguishing between categories and characteristic features that were common with the other category. For both classification and inference learners, two categories were presented in a blocked sequence, whereas the other two categories were presented in an interleaved sequence.

The transfer phase was administered immediately after the learning phase. This phase was designed to examine how well participants learned the categories and on what basis they generalized to novel items. Similar to the learning phase, classification learners were asked to classify items by predicting the category membership, whereas inference learners were asked to predict a missing feature. However, there were three differences compared to the learning phase. First, feedback was not provided. Second, the order of the trials was randomized. And third, the trials consisted of old items shown in the learning phase, and novel items which were not presented before. Importantly, two critical types of novel stimuli were used in the transfer task: Characteristic-Changed creatures and Characteristic-Preserved creatures. Both types of stimuli were composed of deterministic features which distinguished between categories, characteristic features which were common in both category members, and random features which were used to increase the number of unique creatures. There was one critical difference between the two types of stimuli: whereas the characteristic features were the same as the characteristic features in the learning phase for Characteristic-Preserved creatures, they were replaced by novel features for Characteristic-Changed creatures. If participants pay attention to within category similarities and encode the characteristic features, they should exhibit better performance for the Characteristic-Preserved stimuli than the Characteristic-Changed stimuli. Conversely, if participants optimize their attention to differences between categories and have a greater tendency to encode deterministic features, they should exhibit equivalent performance for the Characteristic-Preserved and the Characteristic-Changed stimuli.

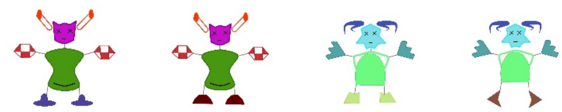

Figure 1. Example of creatures for each category.

A. B.

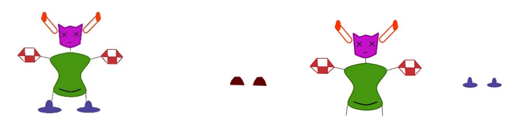

Figure 2. Example of stimuli in classification learning (A) and inference learning (B).

Based on the review above, we predicted that learning sequence would have similar effect on classification and inference learning. Specifically, following blocked learning, classification and inference learners would exhibit better performance for Characteristic-Preserved stimuli than for Characteristic-Changed stimuli. That is, these participants would encode characteristic features better. In contrast, following interleaved learning, classification and inference learners would exhibit equivalent performance, and these participants will encode deterministic features.

Finally, the feature-prediction phase was designed to examine the effectiveness of encoding different features in the course of learning. Participants were presented with learned and novel features and asked to rate the likelihood of features being part of a specific category on a 0-100 scale. For this phase, we also expected that learning sequence would have differential effect on classification and inference learning. Classification learners were expected to encode the characteristic features relatively more effectively in the course of blocked learning than interleaved learning. That is, characteristic features should be rated as category-relevant to a greater extent following blocked learning than interleaved learning. As for inference learners, we expected that they would encode the characteristic features equivalently effectively in both sequences of learning, by rating characteristic features as category-relevant equivalently following blocked and interleaved learning.

Method

Participants

Participants were 60 University of Macau students (43 women). They were tested in a quiet laboratory room on campus and participated for course credit. There were two between-subjects conditions (i.e., classification learning and inference learning), with 30 participants per condition. Informed consent was obtained from each participant.

Stimuli

The stimuli were images of four categories of artificial creatures. The creatures were distinct in their visual appearance. Each creature was composed of five feature-dimensions (i.e., head, body, hands, feet, and antennae) and had different feature values for each dimension. Figure 1 shows examples of the four categories of creatures, and Figure 2 shows examples of stimuli used in classification and inference learning.

Except for the feature design (i.e., shape and color) for each dimension, the structure of creatures was similar to that used in Carvalho & Goldstone's (2017). Each creature was composed of deterministic features, characteristic features, and a random feature. Deterministic features were highly discriminative but lowly characteristic of a category, whereas characteristic features were lowly discriminative but highly characteristic of that category. The distinction reflected how the two kinds of features were distributed among all the creatures presented in the same and the contrasting category, and it was adopted from research of categorization (e.g., Anderson, 1991; Murphy & Ross, 2005). In the learning phase, participants learned creatures with structures in Table 1. As shown in the table, there were three possible feature values in Feature 1-3. Value 2 and Value 3 predicted the category membership and were deterministic features. Value 1 was a characteristic feature frequently shown across both category members and did not predict the category membership. There were also three possible feature values for Feature 5. However, the three values did not predict category membership. These were used as random feature to increase the number of unique creatures. The four categories used in this study shared the same category structure. Two categories of creatures were randomly selected to be learned blocked and the other two categories were learned interleaved. Each category was given a unique novel label.

In the transfer phase, there were three types of stimuli in both learning conditions. 1/3 were creatures previously learned in the learning phase. 1/3 were creatures distinctive from the learned creatures only in the values of Dimension 5 (Characteristic-Preserved stimuli). The rest were creatures distinctive from the learned stimuli in the values of Feature 5 as well as the values of the characteristic features (Characteristic-Changed stimuli). Novel stimuli were composed of Characteristic-Preserved and Characteristic-Changed stimuli. Table 2 demonstrates a category structure of novel stimuli for one category used in the transfer phase. Note that the four categories shared the same category structure of novel stimuli used in the transfer phase.

In the feature prediction phase, participants were presented with images of four types of features: characteristic features shown in the learning phase, deterministic features, features that were presented for the first time during the transfer phase to replace the learned characteristic features, and novel features that had never been presented during the learning or transfer phases.

Table 1. Category structure of the stimuli used in the learning and transfer phases. Note that a feature value represents a specific feature. Feature values are independent across features (e.g., the value 1 in Feature 1 is not the same as the value 1 in Feature 2).

		Feature				
Category	Item	1	2	3	4	5
A	1	2	1	1	1	4
A	2	2	1	1	1	5
A	3	2	2	1	1	3
A	4	1	2	1	1	3
A	5	1	2	2	1	5
A	6	1	1	2	1	4
A	7	1	1	2	2	4
A	8	1	1	1	2	3
A	9	1	1	1	2	5
B	1	3	1	1	1	4
B	2	3	1	1	1	5
B	3	3	3	1	1	3
B	4	1	3	1	1	3
B	5	1	3	3	1	5
B	6	1	1	3	1	4
B	7	1	1	3	3	4
B	8	1	1	1	3	3
B	9	1	1	1	3	5

Table 2. Category Structure for the Novel Stimuli of One of the Categories in Transfer Task

		Feature				
Category	Item	1	2	3	4	5
A	Characteristic-changed	2	4	4	4	6
A	Characteristic-changed	2	4	4	4	7
A	Characteristic-changed	2	2	4	4	6
A	Characteristic-changed	4	2	4	4	7
A	Characteristic-changed	4	2	2	4	6
A	Characteristic-changed	4	4	2	4	7
A	Characteristic-changed	4	4	2	2	6
A	Characteristic-changed	4	4	4	2	7
A	Characteristic-changed	4	4	4	2	6
A	Characteristic-preserved	2	1	1	1	6
A	Characteristic-preserved	2	1	1	1	7
A	Characteristic-preserved	2	2	1	1	6
A	Characteristic-preserved	1	2	1	1	7
A	Characteristic-preserved	1	2	2	1	6
A	Characteristic-preserved	1	1	2	1	7
A	Characteristic-preserved	1	1	2	2	6
A	Characteristic-preserved	1	1	1	2	7
A	Characteristic-preserved	1	1	1	2	6

Design and procedure

Participants were randomly assigned to the classification or inference condition. All participants were instructed to respond as accurately and fast as possible. In both conditions, participants learned four categories. Two categories were learned in blocked sequence and the other two categories were learned in interleaved sequence. The order of learning sequence was counterbalanced across participants. There were three phases (i.e., learning phase, transfer phase and feature-prediction phase). The three phases were presented in the same order for all participants. All participants started by learning two categories in one of the two learning sequences, followed by a transfer task, and then a feature-prediction task. Then, participants completed the three phases with another two categories and the other learning sequence.

In the learning phase, participants were told that there were alien creatures with unique names. In the classification condition, participants were instructed to predict the category membership (i.e., the label) of an item and they were provided with information about all the features. In contrast, in the inference condition, participants were instructed to predict a missing feature of an item. These participants were given information about the remaining features and the label.

In both conditions, there were 36 learning trials (9 trials per category). Specifically, in the blocked learning, participants learned one block containing only one category of creatures, followed by another block with creatures from the contrasting category. In the interleaved learning, participants learned a creature from one category followed by a creature from the contrasting category. Each learning trial was accompanied by feedback. The order of the assigned creature on each learning trial was randomized across participants in both conditions.

The transfer phase was similar to the learning phase, with the participants in the classification condition predicting the category membership of a given item and those in the inference condition predicting a missing deterministic feature of an item. In both conditions, there were 108 trials (27 trials per category). All trials were presented randomly and no feedback was provided.

Finally, in the feature-prediction phase, participants in both classification and inference conditions were asked to rate the likelihood of features being part of a specific category on a 0–100 scale. A rating of 0 indicated that it was impossible that the feature belonged to a certain category, whereas a rating of 100 indicated that the feature definitely belonged to a certain category. There were 80 trials (20 trials per category), and all trials were presented randomly.

Results

Learning Phase All participants were asked to learn four categories by classification or inference in this phase. The learning accuracy was collected. For classification condition, five participants were given wrong task during the learning phase, and one failed to complete the feature prediction task. Data from these participants were excluded from the following analyses. For inference condition, one participant in the transfer task was three standard deviations below the mean accuracy in the last 8 interleaved learning trials, and data from this participant were excluded from the analyses. A 2 (Learning Regime: Classification vs. Inference) × 2 (Learning sequence: Blocked vs. Interleaved) mixed ANOVA revealed a main effect of learning regime, F (1, 52) = 4.90, MSE = 0.22, p = 0.03, with inference learners being more accurate than classification learners. Neither the main effect of learning sequence (p = 0.154) nor the interaction was significant (p = 0.463). We further analyzed the data for classification and inference learning respectively in order to examine the dynamics of learning in two learning regimes.

For classification learning, data were aggregated into two blocks. One block consisted of the first 5 trials for each category, and the other block consisted of the rest 4 trials for each category. As shown in table 3, participants exhibited above chance level of learning in the last 4 trials for blocked sequence, one-sample $t(23)$ = 4.85, p < .001, but not for interleaved sequence, one-sample $t(23)$ = 1.74, p = 0.10. In addition, participants exhibited higher accuracy in the last 4 learning trials than the first 5 trials in blocked learning, paired-samples $t(23)$ = 4.69, p < .001, whereas learning accuracy did not significantly improve between two blocks in the interleaved learning, p = 0.16.

Table 3. Learning Data. Mean accuracy aggregated in 5- or 4- Trial Blocks for Classification and Inference Learning.

Learning Regime	Learning Sequence	Trials 1-5	Trials 6-9
Classification	Blocked	0.49 (0.12)	0.70 (0.20)
Classification	Interleaved	0.55 (0.27)	0.60 (0.28)
Inference	Blocked	0.63 (0.19)	0.79 (0.16)
Inference	Interleaved	0.59 (0.24)	0.68 (0.24)

Similarly, data of inference learning were aggregated into two blocks (i.e., first 5 trials vs. last 4 trials for each category), and these data were shown in table 3. Similar to classification learners, inference learners exhibited above chance level of learning in the last 4 trials for both blocked (one-sample $t(28)$ = 9.48, p < .001) and interleaved sequence (one-sample $t(28)$ = 3.80, p = .001). Furthermore, these participants exhibited higher accuracy in the last 4 learning trials than the first 5 trials in both blocked learning (paired-samples, $t(28)$ = 4.34, p < .001) and interleaved learning (paired-samples, $t(28)$ = 3.64, p = .001).

Taken together, learning data indicated that both classification learners and inference learners learned the two sets of contrasting categories, with inference learners performed somewhat better than the classification learners.

Transfer Phase The transfer task was used to examine the effect of learning sequence on category representations. If participants form similarity-based category representations in the sequence of classification learning, as we predicted, the performance for the Characteristic-Changed stimuli and the Characteristic-Preserved stimuli should be equivalent. In contrast, if participants form rule-based category representations during the learning phase, performance for the Characteristic-Changed stimuli should be worse than the Characteristic-Preserved stimuli. The transfer accuracy was collected and are shown in Figure 5. The average performance across participants for both characteristic-changed and characteristic-preserved stimuli in different learning sequences (i.e., blocked and interleaved learning) were calculated. Data were analyzed with a 2 (Stimuli Type: Characteristic-Changed stimuli vs. Characteristic-Preserved stimuli) × 2 (Learning Sequence: Blocked vs. Interleaved) × 2 (Learning Regime: Classification vs. Inference) mixed ANOVA. There was no significant three-way interaction (p = 0.92). However, a two-way interaction between learning regime and learning sequence was found, F (1, 104) = 20.24, p < .001. We further broke down the analysis by conducting a mixed ANOVA on stimuli type and learning sequence for each learning regime.

For classification learning, the overall effect of learning sequence was not significant (p = 0.69). Neither the type of stimuli (p = 0.61) nor the interaction between learning sequence and type of stimuli (p = 0.27) was significant. Participants' performance for the Characteristic-Changed and the Characteristic-Preserved stimuli following both blocked and interleaved learning did not achieve above-chance level (p = 0.16). These results were unexpected and we further discussed it in the Discussion.

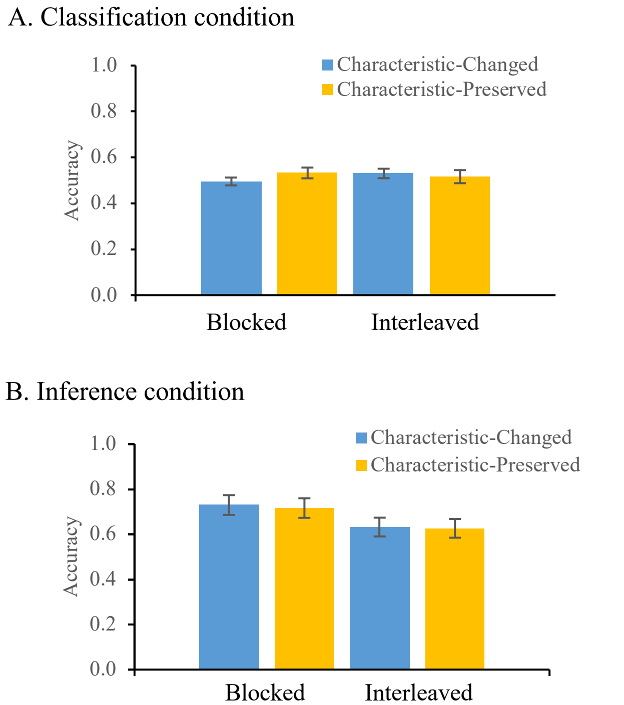

Figure 3. Accuracy of the transfer task in the classification learning condition (A) and the inference learning condition (B). Chance level is 0.5. Error bars represent ±1 standard error of mean.

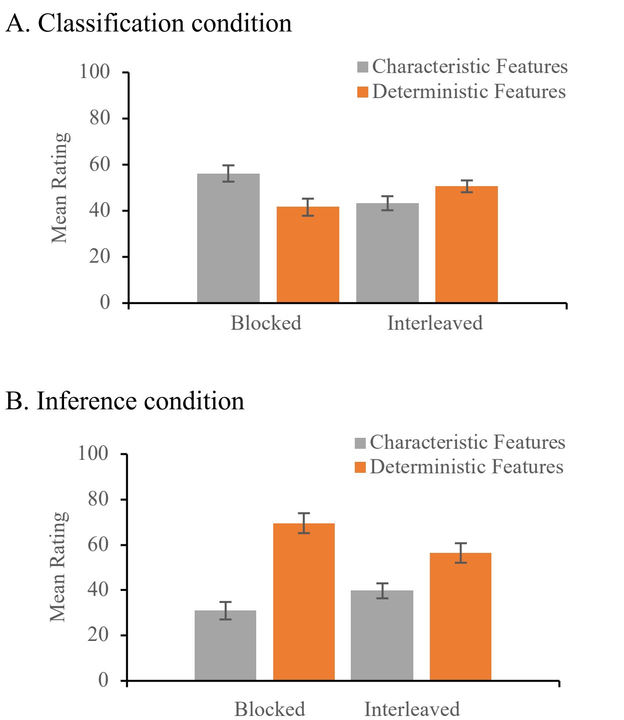

Figure 4. Ratings for features in the feature prediction task in the classification condition (A) and the inference condition (B). Error bars represent ±1 standard error of mean.

For inference learning, the overall effect of learning sequence was significant, F (1, 28) = 4.68, p = 0.03. Participants performed better for the transfer task after blocked learning (M = .73, SD = .24) than that after interleaved learning (M=.63, SD = .23). Neither the type of stimuli (p = 0.88) nor the interaction between learning sequence and type of stimuli (p = 0.98) was significant. In addition, participants' performance for the Characteristic-Changed and the Characteristic-Preserved stimuli was above chance following both blocked (one-sample t (28) = 5.27, p < .001) and interleaved learning (one-sample t (28) = 3.17, p = .004). The results suggested that inference learners formed rule-based category representations in the course of blocked and interleaved inference learning.

Feature Prediction Phase The feature prediction task was used to examine whether the learning sequence changed the perceived relevance of different features. Participants rated how predictive a feature was of a particular category. If learning sequence has an impact on the perceived relevance of features, as discussed in the Introduction, there should be an interaction between types of features and learning sequence. Characteristic features should be rated as relevant for categorization to different extents following different learning sequences. If learning sequence does not influence the perceived relevance of features in inference condition, there should be no interaction between types of features and learning sequence. Participants should rate characteristic features as relevant for categorization equivalently following different learning sequences.

Feature ratings were collected and shown in Figure 4. Data were analyzed with a 2 (Feature Type: Deterministic Features vs. Characteristic Features) × 2 (Learning Sequence: Blocked vs. Interleaved) × 2 (Learning Regime: Classification vs. Inference) mixed ANOVA. There was a significant three-way interaction, F (1, 104) = 27.71, p < .001. We broke down the interaction by conducting a mixed ANOVA on feature type and learning sequence for each learning sequence.

For the classification condition, a significant interaction between learning sequence and type of features was found, F (1, 23) = 11.46, p = .001. Post hoc t test indicated that the average rating for Characteristic Features (M = 54.93, SD = 3.53) was higher than Deterministic Features (M = 41.50, SD = 3.70) after the blocked learning, t (23) = 2.68, p = .01, whereas following the interleaved learning, participants rated higher for Deterministic Features (M = 50.61, SD = 2.57) than Characteristic Features (M = 43.28, SD = 3.05), t (23) = 2.05, p = .03. Therefore, the results indicated that, in classification, the learning sequence changed the perceived relevance of features. These findings, consistent with our predictions, suggested that classification learners encoded characteristic features relatively more effectively following the blocked learning, whereas following the interleaved learning, they encoded deterministic features relatively more effectively.

For the inference condition, a significant interaction between learning sequence and type of features was found, F (1, 28) = 7.28, p = .008. Pairwise t test indicated that participants rated higher for Deterministic Features than Characteristic Features after blocked learning, t (28) = 6.56, p = .001, and interleaved learning, t (28) = 3.04, p = .002. Inference learners encoded the characteristic features equivalently effectively following both sequences of learning, t (28) = 2.05, p = .12. However, participants encoded the deterministic features more effectively following blocked learning (M = 69.53, SD = 4.40) than interleaved learning (M = 56.38, SD = 4.31), t (28) = 2.41, p = .01. Therefore, in contrast to classification learners, for inference learners, learning sequence did not change the perceived relevance of the characteristic features but changed the perceived relevance of the deterministic features.

Taken together, the results in the feature prediction task corroborates findings from the transfer task: whereas classification learners encoded characteristic features more effectively in the blocked learning and encoded deterministic features more effectively in the interleaved learning, inference learners, regardless of the learning sequence, encoded deterministic features more effectively.

Discussion

In the study reported here, we investigated how learning sequence influences category learning and representations under classification and inference learning. The current study reveals one important and novel finding pointing to the differential effect of learning sequence on category learning under different regimes. Specifically, classification learners are more likely to encode characteristic features and form similarity-based representations in the blocked study, whereas in the interleaved study, they tend to encode deterministic features and form rule-based representations, In contrast, for inference learners, the blocked and interleaved study changed their learning and representation in the same direction. That is, in both sequences, inference learners are more likely to encode deterministic features and form rule-based representations. The finding with inference learners is in sharp contrast with previous research (e.g., Hoffman & Rehder, 2010; Deng & Sloutsky, 2015) showing that inference learners tend to encode multiple features and form similarity-based representations.

One explanation for this finding is that, in the current study, inference learners were asked to predict the deterministic features which were closely related to the rule for making inference of the missing deterministic features. This is likely to shift participants' attention to the deterministic features. In the sequence of blocked learning, participants successively learned category members from the same category and encoded deterministic features effectively. They learned the contrasting category in the similar way and then formed rule-based category representation in the course of learning. Conversely, in the sequence of interleaved learning, participants learned categories by contrasting the deterministic features and formed rule-based representation. Therefore, following both blocked and interleaved learning, inference learners encoded the deterministic features better than the characteristic features and formed rule-based representation. Further research is required to directly examine the attentional changes in category learning.

There are several important issues that have not been addressed by current study and will require further research. One main issue pertains to the relatively poor performance of classification learners in this study, and this is likely to result in the inconsistent results for classification learners compared to previous research. Specifically, in the current study, participants in the classification condition did not learn well, which highly contrasted with previous findings (Carvalho, & Goldstone, 2017) in which participants consistently exhibited higher than 60% of accuracy regardless of learning sequences. In addition, the classification leaners did not exhibit equivalent learning performance compared to the inference learners, which contrasted to previous research in which classification learners usually learnt better than inference learners (e.g., Deng & Sloutsky, 2015). One possible explanation for the results is that even though the structure of the categories in the current study was similar to that in previous study, the feature values (i.e., the color and shape of the features) in this study were not highly distinct between each other, which was likely to require more effort for participants to encode features, and hence to decrease the accuracy during the transfer phase.

Another explanation for the results is that the learning task was likely to be difficult for the participants to acquire the rule in classifying items in both learning sequences. Evidence for this explanation comes from studies suggesting that category structure matters to classifying items (Yamauchi, Love, & Markman, 2002). That is, classification learners failed to have a good summary of all category members because of the category structure applied in this study. However, even though during the transfer phase, classification learners did not achieve above-chance level in the transfer phase, the learners exhibited more effective encoding of characteristic features following blocked learning and deterministic features following interleaved learning during the feature prediction phase.

Furthermore, the developmental changes in category learning of different sequences under classification and inference regimes have not been addressed. Research has shown that there is a developmental difference in category learning by classification and inference (Deng & Sloutsky, 2015). In contrast to adults who tend to treat classification and inference learning differently, young children tend to treat these two learning regimes equivalently. Future developmental research will be needed to examine whether the sequence of learning would have similar effect on children's category learning under different learning regimes. By examining developmental changes of these effects, we would be better positioned at characterizing the attentional mechanisms underlying category learning.

References

Anderson, A. L., Ross, B. H., & Chin-Parker, S. (2002). A further investigation of category learning by inference. *Memory & Cognition, 30,* 119–128.

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

Clapper, J. P. (2014). The impact of training sequence and between category similarity on unsupervised induction. *The Quarterly Journal of Experimental Psychology, 68,* 1–55.

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

Carvalho, P. F., & Goldstone, R. L. (2015a). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin & Review, 22,* 281–288.

Carvalho, P. F., & Goldstone, R. L. (2015b). What you learn is more than what you see: What can sequencing effects tell us about inductive category learning? *Frontiers in Psychology, 6*(505), 1–12.

Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 1699–1719.

Deng, W., & Sloutsky, V. M. (2012). Carrot eaters or moving heads: Inductive inference is better supported by salient features than by category labels. *Psychological Science, 23,* 178–186.

Deng, W. S., & Sloutsky, V. M. (2015). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology, 51,* 392–405.

Hoffman, A. B., & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility. *Journal of Experimental Psychology: General, 139,* 319–340.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129,* 592– 613.

Murphy, G. L., & Ross, B. H. (2005). The two faces of typicality in category-based induction. *Cognition, 95,* 175–200.

Sandhofer, C. M., & Doumas, L. A. A. (2008). Order of Presentation Effects in Learning Color Categories. *Journal of Cognition and Development, 9,* 194–221.

Yamauchi, T. (2009). Finding abstract commonalties of category members. *Journal of Experimental & Theoretical Artificial Intelligence, 21(3),* 155–180.

Yamauchi, T., Love, B.C., & Markman, A.B. (2002). Learning Non-linearly Separable Categories by Inference and Classification. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 585–593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language, 39,* 124–148.

Yamauchi, T., and Markman, A.B. (2000). Learning Categories Composed of Varying Instances: The Effect of Classification, Inference and Structural Alignment. *Memory and Cognition, 28,* 64–78.

Zaki, S. R., & Salmi, I. L. (2019). Sequence as context in category learning: An eyetracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45(11),* 1942–1954.