Cognition 155 (2016) 23–29

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Optimal sequencing during category learning: Testing a dual-learning systems perspective

Sharon M. Noh [a,*], Veronica X. Yan [b], Robert A. Bjork [b], W. Todd Maddox [a]

[a] University of Texas, Austin, United States
[b] University of California, Los Angeles, United States

ARTICLE INFO

Article history:
Received 11 May 2015
Revised 10 June 2016
Accepted 14 June 2016

Keywords:
Category learning
Schedules
Sequencing
Blocking
Interleaving
Dual-systems

ABSTRACT

Recent studies demonstrate that interleaving the exemplars of different categories, rather than blocking exemplars by category, can enhance inductive learning—the ability to categorize new exemplars—presumably because interleaving affords discriminative contrasts between exemplars from different categories. Consistent with this view, other studies have demonstrated that decreasing between-category similarity and increasing within-category variability can eliminate or even reverse the interleaving benefit. We tested another hypothesis, one based on the dual-learning systems framework—namely, that the optimal schedule for learning categories should depend on an interaction of the cognitive system that mediates learning and the structure of the particular category being learned. Blocking should enhance rule-based category learning, which is mediated by explicit, hypothesis-testing processes, whereas interleaving should enhance information-integration category learning, which is mediated by an implicit, procedural-based learning system. Consistent with this view, we found a crossover interaction between schedule (blocked vs. interleaved) and category structure (rule-based vs. information-integration).

1. Introduction

When learning new categories, how should the study of category exemplars be sequenced so that learners can accurately classify new exemplars on a later test? When an art student, for example, must learn to recognize the styles of different artists so as to be able to identify the artist responsible for a never-before-seen painting, should he or she study examples of artists' paintings one artist at a time, or should the paintings by the different artists be intermixed? Recent findings suggest that in this case, and in the inductive learning of other naturalistic categories, such as butterflies and birds, interleaving exemplars of different categories yields better category learning than does blocking exemplars by category (e.g., Birnbaum, Kornell, Bjork, & Bjork, 2013; Kang & Pashler, 2012; Kornell & Bjork, 2008; Wahlheim, Dunlosky, & Jacoby, 2011). More recent work using artificial stimuli suggests, however, that interleaving is only superior when between-category discriminability is low, and that blocking is superior when between-category discriminability is high (e.g., Carvalho & Goldstone, 2014; Zulkiply & Burt, 2013). The important implication of these studies is that there may be no single "optimal" method of sequencing, but rather, the optimal method may depend on various factors (e.g., the nature of the to-be-learned categories).

Although category discriminability can play an important role in determining whether interleaved or blocked study enhances category learning, we argue that another, yet unexplored, factor may be important: the learning system that mediates performance. An extensive body of behavioral, neuropsychological, and neuroscience literature suggests that optimal learning of different category structures is mediated by at least two neurobiologically grounded and competing learning systems (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Maddox, 2011; Maddox & Filoteo, 2005; Nomura & Reber, 2008). One is a frontally mediated hypothesis-testing system that relies on working memory and executive attention to develop and test verbalizable rules that are used to optimally solve rule-based (RB) categories. The second is a striatally mediated procedural-based learning system that does not rely on working memory and executive attention but, instead, learns non-verbalizable stimulus-response mappings that are used to solve information-integration (II) categories. These two systems compete and previous research show that there is an initial bias toward the hypothesis-testing system, with control being passed to the procedural-based learning system only when the category structure warrants (e.g., with information-integration categories;

* Corresponding author at: 108 E. Dean Keeton Stop A8000, Department of Psychology, University of Texas at Austin, Austin, TX 78712-1043, United States.
E-mail address: smnoh@utexas.edu (S.M. Noh).

http://dx.doi.org/10.1016/j.cognition.2016.06.007
0010-0277/

Ashby & Maddox, 2011; Ashby et al., 1998; Maddox & Ashby, 2004). Dual-learning-systems research suggests that learning in each system is optimized under different training conditions. For instance, rule-based category learning is optimized when full feedback is provided (e.g., "Wrong, that was a B") whereas information-integration category learning is optimized when immediate, minimal feedback is provided (e.g., "Wrong"; Maddox, Love, Glass, & Filoteo, 2008).

We hypothesize that the optimal schedules for category learning are also dependent on the underlying category structure. In the current study, we tested this hypothesis directly. With respect to rule-based categories, blocking exemplars by category should allow individuals to more easily generate, test, and adjust their working hypotheses, particularly when there is a relatively demanding working memory load. To introduce a working memory load, we used a four-category variant of the rule-based and information-integration learning structures (from Maddox, Filoteo, Hejl, & Ing, 2004), rather than the more typical two-category learning variant found in many dual-learning systems studies. An interleaved schedule, on the other hand, would hurt rule-based learning by introducing a more demanding working memory load, as individuals would have to generate and test multiple rules for each category simultaneously. While interleaving would allow learners to compare exemplars that do and do not fit into a given category, the working memory load involved in holding multiple dimensions in mind for multiple categories would make using rule-based hypothesis testing difficult. We predict that blocked study should better should facilitate rule-based category learning than interleaved study in our experiments. Following the same reasoning, we also hypothesize that interleaved study should be beneficial for information-integration category learning because it discourages the use of rule-based strategies and speeds the transition to the procedural based learning system.

2. Experiment 1

2.1. Method

2.1.1. Participants and design

One hundred and thirty-two participants (mean age = 30.0, age range = 19–57, 71 females) were recruited from Amazon Mechanical Turk (MTurk) and paid $1.00 for their participation. Category structure (rule-based vs. information-integration) and study schedule (blocked vs. interleaved) were manipulated in a 2×2 between-subjects design. An *a priori* power analysis determined that for a medium effect size ($f = 0.25$), we would need 32 participants per condition to reach a power of 0.80.

2.1.2. Materials

The four-category rule-based and information-integration category structures are displayed in Fig. 1. Each stimulus was comprised of a line of varying length and orientation at a fixed distance from center (that varied in position) on the computer screen. The stimuli were constructed from three continuous-valued dimensions: line orientation (between 0 and 90°), line length (0–200 pixels), and position (between 0 and 100° offset from fixation). Each dimension has eight values at equal intervals, but only line length and line orientation values defined category membership. Each of the eight line length values were paired with each of the 8 line orientation values, for a total of 64 unique lines of varying length and orientation. These 64 lines were randomly paired with one of 8 different positions, so that each unique line could be shown in one of 8 positions on the screen. In the rule-based condition, the stimulus space was divided into four categories using decision bounds that were verbalizable (e.g., "all members of category A contain a short, steep line"). To generate the information-integration condition, the category boundaries and stimuli from the rule-based condition were rotated 45° so that no simple verbalizable rule could define category membership. This transformation allows us to both differentiate rule-based and information-integration category-learning strategies while keeping the category structures and stimulus distributions mathematically equivalent.

2.1.3. Procedure

Participants were asked to learn to distinguish exemplars from four different categories. A cover story was provided, suggesting that these were images generated by four different robots and the task was to learn each robot's way of generating images. During the study phase, participants observed each of the 64 images (constructed from the factorial combination of all 8 line lengths with all 8 line orientations) once with a randomly selected (but without replacement) position. Each item was presented with the appropriate category label (A, B, C, or D) for 3.5 s each. In the blocked condition, participants saw the 16 exemplars from one category before moving on to the next, whereas in the interleaved condition, all 64 exemplars were presented in a randomized order. Fig. 2 shows examples of the sequencing and stimuli used in the study phase. Following this passive study phase, participants moved on to the test phase, where they were shown the same 64 length-orientation pairings. The test stimuli were randomly presented, and following each stimulus presentation, participants were asked to select what they believed to be the appropriate category label by clicking on one of four buttons (labeled A, B, C, and D) arranged horizontally below each stimulus display. This final test was self-paced and without feedback.

2.2. Results and discussion

2.2.1. Classification performance

Average final test performance for each condition is presented in Fig. 3. A 2×2 between-subjects ANOVA revealed a main effect of category structure such that accuracy was higher for information-integration category structures ($M = 0.58$, $SD = 0.17$), relative to rule-based structures ($M = 0.51$, $SD = 0.19$), $F(1,128) = 5.36$, $MSE = 0.03$, $p = 0.022$, $\eta_p^2 = 0.04$. There was no significant main effect of schedule, $F(1,128) = 0.30$, $MSE = 0.03$, $p > 0.20$. There was, however, a significant interaction, $F(1,128) = 5.34$, $MSE = 0.03$, $p = 0.022$, $\eta_p^2 = 0.04$. Post-hoc *t*-tests revealed that for rule-based categories, accuracy following blocked study ($M = 0.55$, $SD = 0.19$) was marginally higher than accuracy following interleaved study ($M = 0.46$, $SD = 0.19$), $t(66) = 1.96$, $p = 0.055$, $d = 0.47$. The pattern was reversed, however, for information-integration categories: Accuracy following interleaved study ($M = 0.61$, $SD = 0.17$) was higher than accuracy following blocked study ($M = 0.55$, $SD = 0.17$), but this difference was not found to be significant $t(62) = 1.30$, $p = 0.20$, $d = 0.33$.

2.2.2. Model fits

The accuracy-based analyses suggest that blocking enhances RB learning, whereas interleaving helps II learning. We hypothesized that this effect would occur because blocking may facilitate hypothesis-testing and the rule-discovery process, whereas interleaving may discourage rule use (perhaps by introducing a working memory load). To examine this possibility, we fit a number of different decision bound models (e.g., Ashby & Gott, 1988; Maddox & Ashby, 1993) to the data from each individual participant in order to understand the kind of strategy each participant used to classify the stimuli. For each of the four experimental conditions, the relevant models were fit separately to the data from the 64-trial test block.