Experiment 1

Subject

198 undergraduates at Indiana University were recruited for this experiment. There were 91 valid subjects in repeating condition (REP) and 95 valid subjects in non-repeating condition (NREP). Every subject was randomly assigned to the REP or NREP condition. 12 subjects (7 in REP condition and 5 in the NREP condition) were excluded for data analysis due to overly poor performance. The exclusion criterion in this experiment was based on the two distributions of classification accuracy rate across all subjects in the 8-15 blocks of the learning phase and the complete transfer phase. Every subject whose accuracy rate fell below mean – 2.5 \* standard deviation of the accuracy distribution in either phase was excluded from subsequent data analyses.

Stimuli and apparatus

The stimuli used in this experiment are dot patterns generated using Posner, Goldsmith, and Welton's (1967) procedure. Each pattern consists of 9 dots positioned in a 32 x 32 grid and connected with lines. Prototypes for six different categories were randomly generated, and three of them were used to generate foils only for the recognition transfer in experiment 1. Other patterns used in the learning and transfer phases were generated using the statistical-distortion procedure by Posner et al. (1967). Each pattern was constructed from the prototype of its category by displacing each dot by an arbitrary distance. Low-level, medium-level and high-level distortions are, on average, displaced 4, 6 and 7.7 Posner levels away from their prototype. Patterns with different point configurations were presented for each subject, with the only constraint being that the distortion patterns fell within a pre-specified range around their prototypes. We used Dell Computers to display the stimuli and control the experiment.

Procedure

Repeating condition (REP). A standard learning-transfer paradigm is used. In the learning phase, subjects were instructed to classify dot patterns into three categories A, B and C by pressing corresponding keys. The learning phase consisted of 15 blocks, each of which had 15 trials. On each trial, a medium-level distortion was presented and feedback about the correct category was given after subject responded. The same 15 learning patterns were repeated across the blocks but the order of presentation was randomized. After 225 learning trials, subjects were instructed to recognize whether a transfer item was old (presented in the learning phase) or new (not presented in the transfer phase) by pressing two corresponding keys. The transfer patterns comprised of the 15 old distortions presented in the training phase, 3 prototypes (1 per category), 15 new medium-level distortions (5 per category), and 6 foils or medium-level distortions from 3 other prototypes (2 per category). Each pattern was presented for a total of 39 trials. The order of presentation was randomized for each subject.

No-repeating condition (NREP). The training phase was similarly structured as in the repeating condition, but 225 different medium-level distortions were presented over 15 blocks. In other words, no single learning pattern was repeated in any two trials. In the transfer phase, 15 old distortions were randomly sampled from the 225 learning patterns, with the constraint that no two patterns had been presented in the same training block and that equal number of patterns from each category was presented.

Results

Learning. The learning performance across the 15 blocks in the training phase for the REP and NREP conditions are shown in figure 1. The substantial learning effect over blocks was significant, F(8.66,1593.89) = 140.37\* , p < .001, η2 = .433, MSe = . The main effect of learning conditions was significant, F(1,184) = 16.26 , p < .001, η2 = .081, MSe = , as was the interaction effect between learning condition and blocks, F(8.66,1593.89) = 2.463 , p = .01, η2 = .013, MSe = . As can be seen from figure 1, on average, the classification accuracy of subjects in the REP condition improved at an apparently faster rate than subjects in the NREP condition, even only after 5 blocks of training. The average classification accuracy at the end of the learning phase was also higher for the REP than the NREP condition.

\*Greenhouse-Geisser correction applied for violation of the sphericity assumption.

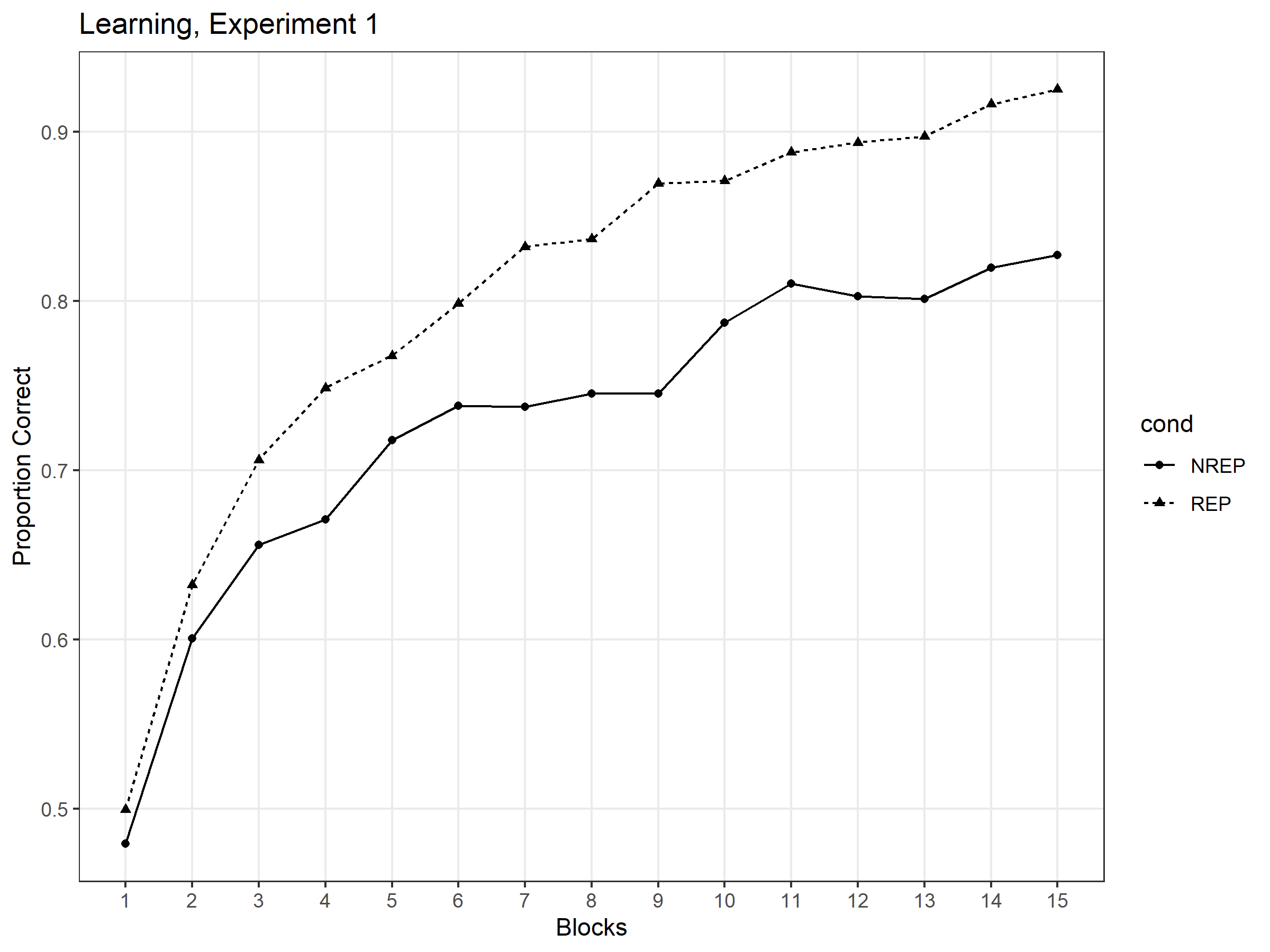


Figure 1 Mean proportion of correct classifications as a function of the number of blocks for REP and NREP conditions, Experiment 1.

Transfer-Recognition. The probability with which each type of the transfer patterns were judged as old in REP and NREP conditions was shown in figure 2. In both conditions, the proportion of old responses for the old (REP: 84.5%, NREP:69.3%), new medium distortion (REP: 34.3%, NREP:63.2%), prototype (REP: 78.4%, NREP:91.6%) and foil (REP: 5.3%, NREP:15.1%) are denoted by four bars from the left to right on the figure. The main effect of the item type, F(2.67,490.68) = 883.93, MSe = , of the learning condition, F(1,184) = 54.85, MSe = , and the interaction between the two, F(2.67,490.68) = 64.66, MSe = , were all significant, all p <.001. In the REP condition, the old distortion was judged as old significantly more than the new medium distortion, t(90) = 24.51, p <.001, cohen’s d = 2.569, and was judged as old marginally significantly more than the prototype, t(90) = 2.21, p = .059\*. In the NREP condition, the old distortion was also judged as old significantly more than the new medium distortion with the difference much smaller than in the REP condition, t(94) = 3.59, p = .001, cohen’s d = .368, but was judged as old significantly less than the prototype, t(94) = 10.21, p < .001. Apparently from the figure, the probability of mistaking a foil for an old distortion was very low in both conditions.

\*In this paper, p values of multiple t tests conducted on the same data set were adjusted for Bonferroni correction. If any p value is less than .05 before the correction but greater than .05 after the correction, the effect of difference is considered marginally significant

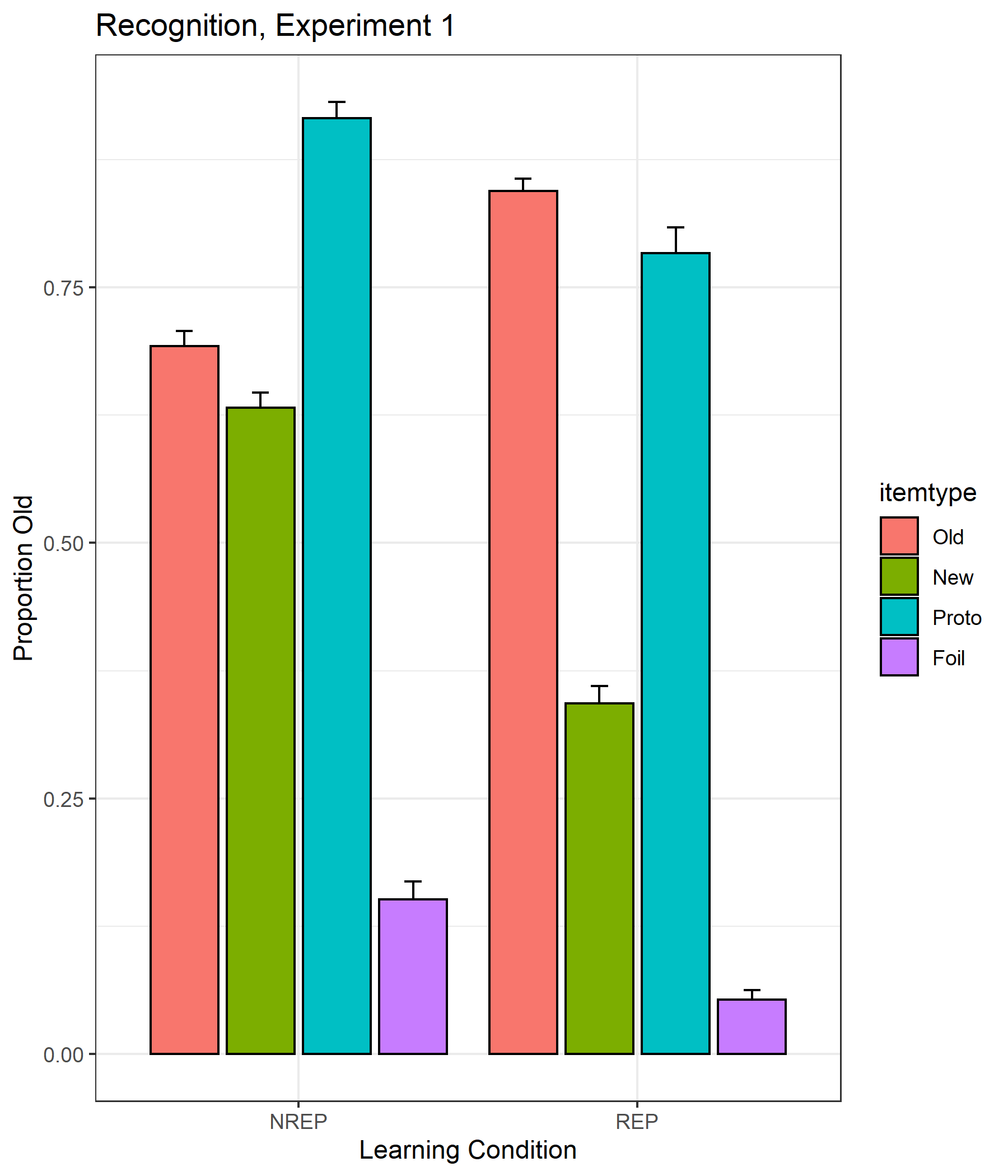


Figure 2 Mean proportion of old responses (with standard error bars) to four different types of transfer patterns (old, new medium, prototype, foil) for REP and NREP conditions, Experiment 1.

Experiment 2

Subject

89 undergraduates at Indiana University were recruited for this experiment. There were 39 valid subjects in the repeating condition (REP) and 44 valid subjects in the non-repeating condition (NREP). Every subject was randomly assigned to the REP or NREP condition. 6 subjects (4 in REP condition and 2 in the NREP condition) were excluded for data analysis due to overly poor performance. The exclusion criterion is similar to that in the experiment 1, except that we now define the accuracy in the recognition transfer phase as the proportion of correctly recognizing an old item minus the proportion of misrecognizing a foil item as old.

Procedure

Repeating condition (REP). The learning phase was the same as in experiment 1. In the transfer phase, subjects were instructed to classify the transfer patterns into the same three categories. The transfer patterns comprised of 15 old distortions, 3 prototypes (1 per category), 15 low-distortions (5 per category), 15 new medium-level distortions (5 per category), and 15 high-level distortions (5 per category). Each pattern was presented for a total of 63 trials. The order of presentation was randomized for each subject.

Non-repeating condition (NREP). The same change from REP to NREP condition was made as in the experiment 1

Results

Learning. The learning performance across the training blocks for the REP and NREP conditions are shown in figure 3. The learning data showed a similar pattern of results to the pattern in the experiment 1. The main effect of learning over blocks was significant, F(7.14,578.08) = 56.78 , MSe = , p < .001, η2 = .412. The main effect of learning conditions was significant, F(1,81) = 18.09 , MSe = , p < .001, η2 = .183. As can be seen from figure 1, the overall learning rate of subjects in the REP condition was higher than subjects in the NREP condition, and the classification accuracy at the end of learning was higher for the REP condition than the NREP condition.

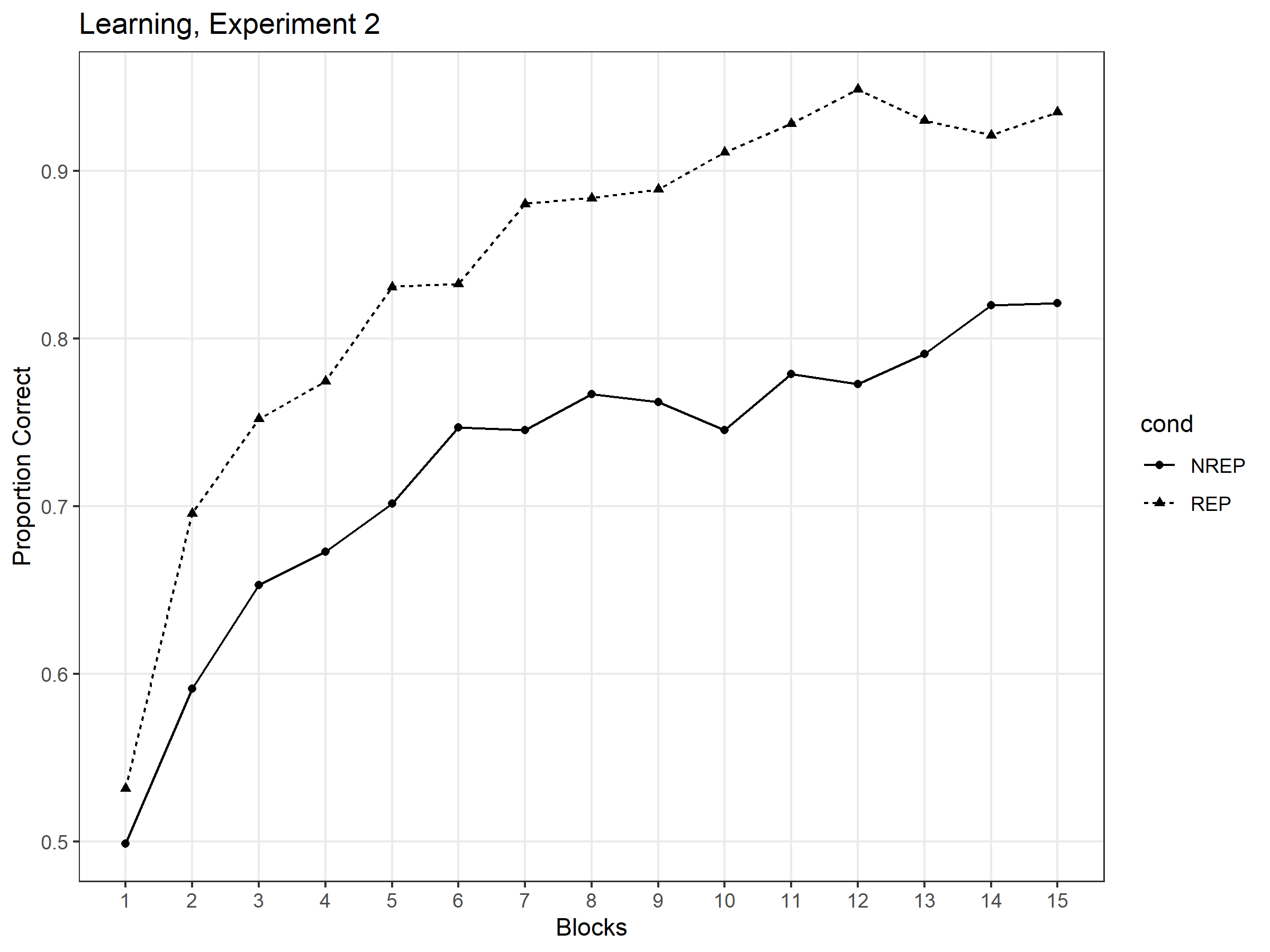


Figure 3 Mean proportion of correct classifications as a function of the number of blocks for REP and NREP conditions, Experiment 2

Transfer– classification. The probability with which each type of the transfer patterns were correctly classified in REP and NREP conditions was shown in figure 4. First, a 2 x 4 mixed-model ANOVA was conducted with learning conditions (REP and NREP) as a between-subject factor and different levels of distortions (prototype, low, new medium and high distortions) as a within-subject factor. There was a main effect of distortion levels, F(2.3,186.67) = 46.08, MSe = .696, p < .001, η2 = .363. Indeed, the rate of classification accuracy decreased as a function of the level of distortions away from the prototype in the REP condition (prototype = 94.0%, low = 92.3%, medium = 86.5%, high = 76.6%), as was the case in the NREP condition (prototype = 93.9%, low = 88.6%, medium = 84.8%, high = 74.4%). However, there was neither an effect of learning condition nor a condition x distortion level interaction, both ps > .05. Second, a 2 x 3 mixed-model ANOVA was conducted on learning conditions and three item types (old, new medium distortion and prototype). The main effect of item type was significant, F(1.62,131.04) = 13.61, MSe = .183, p < .001, as was the condition x item type interaction, F(1.62,131.04) = 4.72, MSe = .064, p = .016. The main effect of condition was not significant, F(1,81) = 1.82, MSe = .085, p = .181. Subsequent test showed that the old distortion was classified significantly more accurate than the new medium distortions in the REP condition, t(38) = 5.50, p < .001, but the difference in classification accuracy was not significant in the NREP condition, t(43) = 1.00, p = .646. Moreover, the old distortion was not classified significantly more accurate than the prototype in the REP condition, t(38) = .98, p = .670, but was classified significantly less accurate than the prototype in the NREP condition, t(43) = -2.78, p = .016.

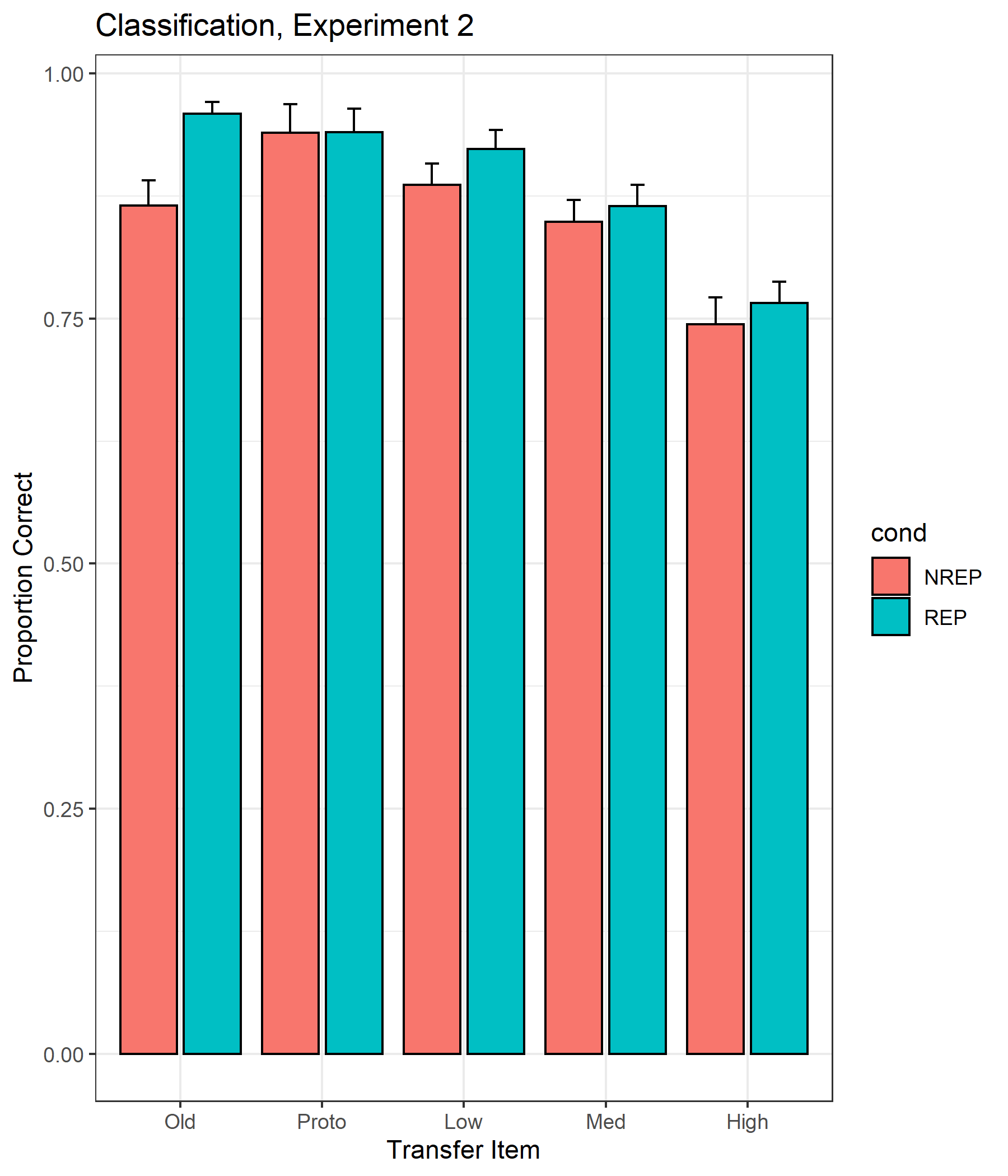


Figure 4 Mean proportion of correct classifications (with standard error bars) to five different types of transfer patterns (old medium distortion, prototype, low distortion, new medium distortion, high distortion) for REP and NREP conditions, Experiment 2.

GCM model

According to an MDS analysis of the similarity ratings of dot patterns (Shin & Nosofsky, 1992), six psychological dimensions can well account for variability in the perceived similarity among dot patterns. In this simulation, we represented each dot pattern as a point in a six-dimensional feature space. The six-coordinates of the prototypes and medium distortions in the learning phase were set in a way analogous to the statistical-distortion procedure by which they were generated. Within-category and between-category distances were two freely varying parameter in our model. The distances of Low, medium and high distortions were set to be the within-category distance multiplied by 1.2, 2.8 and 4.6, respectively, to map the dot displacements used by Homa et al. (2019). Each coordinate of the prototype was randomly sampled from a uniform distribution from 0 to the between-category distance. Each coordinate of a training medium distortion in a certain category was the sum of the corresponding coordinate of the prototype in its category and a random distance sampled from a zero-centered normal distribution with standard deviation equal to the medium-level distortion distance. The training patterns in the two conditions were separately generated. 100 different training patterns were generated for each category in the REP condition, and 5 different training patterns, each to be repeated once in every training block, were generated for each category in the NREP condition. For each experiment, a single exemplar was constructed to represent each item type of the transfer patterns. The transfer patterns in the experiment 1 consisted of prototype, low, medium and high-level distortions. Each level of distortion was constructed from the prototype plus a random distance based on the corresponding distortion distance. The transfer patterns in the experiment 2 consisted of old medium distortion, new medium distortion and foil. For each of two learning conditions, the old distortion was chosen to be the first exemplar constructed for the first category in that condition. A new medium distortion was randomly generated with medium distortion distance. Each coordinate of a foil was defined as the sum of a uniformly distributed random number (0 – between-category distance) and a normally distributed random number (sd = medium distortion distance). The transfer patterns in the experiment 3 were similar to those in the experiment 2, except that the prototype in the first category was used instead of the foil.

In the experiment 1, the evidence in favor of category A, given the presentation of a transfer pattern i, is defined as the summed similarity of the transfer pattern i to each of the learning patterns in category A. The conditional probability with which the pattern i is classified in category A is then computed as the evidence favoring category A plus some background noise, normalized by the sum of evidences for all three categories, which can expressed by equation (1).

(1)

where is the similarity between examplars i and j, is the set of three categories involved in this experiment, and is the number of times each exemplar was repeated during the learning phase in this experiment ( for the REP condition and for the REP condition). denotes the background noise present at the beginning at the learning phase, and its effect on the summed similarity fades away as more exemplars were stored in memory after the learning phase. denotes the response-scaling factor (subjects make classification decisions exactly based on the probability matching when , but towards the category with the largest probability more deterministically when ).

The inter-exemplar similarity is derived from a Euclidean distance between the two dot patterns i and j as represented by points in the six-dimensional psychological space, as in equation (2).

(2)

where represents the coordinate of exemplar i on dimension m in the six-dimensional psychological space. The similarity measure is an exponential decay function of the distance, as in equation (3)

(3)

where the sensitivity parameter c reflects subjects’ overall discriminability in the psychological space.

In the experiment 2 and experiment 3, the overall familiarity for a transfer pattern i is defined as the summed similarity of the pattern i to all the training patterns in all three categories, with background noise and response-scaling parameter included to mirror the same cognitive processes as in the experiment 1. The probability with which the pattern i is judged as an old item is then computed as the overall familiarity divided by the sum of the overall familiarity and a criterion parameter, which can expressed by equation (4).

(4)

where represents four free parameters to be estimated for the two learning conditions in the two experiments separately (i.e. , , and ). It reflects the criterion in making old recognition judgement. Larger means more strict criterion and smaller means more lenient criterion. is also changed to 20 in these experiments as there were 20 learning blocks.

Across the three experiments, 9 freely-varying parameters were estimated: between-category distance, within-category distance, sensitivity parameter c, background noise , response-scaling parameter , and criterion parameters .

Probabilities of the correct classifications in the experiment 1 and of the old recognitions in the experiment 2 by learning conditions and item types were predicted by simulating the experimental processes once. To obtain a reliable result of the mean proportions across subjects as in Homa et al. (2019), the experimental processes were iterated by 10000 times, and the predicted probabilities were averaged across iterations. For simplicity, the prototype of the first category as defined in each iteration was always selected to construct test patterns from. We believe that our model doesn’t lose generality as the prototypes were arbitrarily assigned in each iteration.