Method

Subjects

The subjects were 304 students from Indiana University who participated in partial fulfillment of an undergraduate psychology course requirement. Subjects were randomly assigned to four training conditions. There were 77 subjects in the low-distortion condition, 78 in the medium-distortion condition, 75 in the high-distortion condition and 74 in the mixed-distortion condition. All subjects had normal or corrected-to-normal vision.

Stimuli and apparatus

The stimuli used in this experiment were dot patterns generated using the Posner-Keele (1968) statistical-distortion algorithm. For each individual subject, prototypes for three different categories were generated by placing 9 dots at random grid positions in the central 30 × 30 area of a 50 × 50 grid.

Different training and transfer patterns of each category were generated using the statistical-distortion procedure of Posner et al. (1968). Each pattern was constructed from the prototype of its category by displacing each dot by a random direction and distance in accordance with the Posner et al. procedure. Low-level, medium-level and high-level distortions were generated by moving the individual dots, on average, 4, 6 and 7.7 Posner-levels away from their prototype. Each individual subject was presented with a unique set of randomly generated prototypes and training and transfer patterns.

Procedure

A standard learning-transfer paradigm was used in this experiment. In the learning phase, subjects were trained to classify a set of training patterns into three categories. On each trial, a dot pattern was presented at the center of the computer screen and remained visible until a subject responded with a key press. After the response, corrective feedback appeared for 2s below the presented pattern. A different set of training patterns was presented in each of the 10 training blocks. The learning phase was followed by a transfer phase where subjects classified selected novel patterns as well as a subset of training patterns into the same three categories. No corrective feedback was given on any test trial.

For each individual subject, prototypes for three different categories were randomly generated. Subjects were randomly assigned to one of the four training conditions that differ in terms of the variability of the training patterns. In each condition, 90 training patterns (9 per block) were randomly generated around each of the three category prototypes (270 patterns in total). The category prototypes were distorted by various levels using the Posner-Keele (1968) statistical-distortion algorithm to generate the training patterns for the four training conditions: all low-distortions, all medium-distortions, all high-distortions, and mixture (equal number) of the three distortion levels, respectively. The mixed-distortion training condition contained 90 patterns of low, medium, and high-distortions each, with an equal number of patterns for each distortion level presented in each individual block and category. The test patterns consisted of 27 old patterns that were presented in the training phase (9 per category, with at least 2 of the 27 patterns from each training block), 3 prototypes (1 per category), 9 new low-level distortions (3 per category), 18 new medium-level distortions (6 per category), 27 new high-level distortions (9 per category). Each pattern was presented once in a random order for each subject for a total of 84 trials.

Results

Learning. Figure 1 shows the average proportion of correct classification responses over the training blocks for each of the four training conditions. Across all the training conditions, the classification accuracy gradually improves over the course of training. The low-distortion training condition shows the highest accuracy, the medium- and mixed-distortion conditions show intermediate levels of accuracy, while the high distortion condition shows the lowest accuracy.

To confirm these observations, we conducted a 4x10 mixed-model ANOVA using training conditions (low, med, high, mixed) and blocks as factors. The analysis revealed a significant main effect of blocks, F(6.35, 1905.07) = 84.44, p < .001, η2 = .220\*. The main effect of training conditions was also significant, F(3,300) = 82.85 , p < .001, η2 = .453, as was the interaction effect between learning condition and blocks, F(19.05, 1905.07) = 2.865, p < .001, η2 = .028. To verify the differences in the classification performance across training conditions near the end of the training phase, we compared the mean proportion of correct responses for the last 3 training blocks for each condition. The analysis showed that the mean proportion of correct responses is higher in the low condition (M = 0.885) than in the medium condition (M = 0.691), t(141.2) = 7.55, p < .001\*\*, and higher in the medium condition than in the high condition (M = 0.499), t(149.4) = 6.92, p < .001.

Transfer. Figure 2 shows the mean proportions of correct responses for the different types of test patterns across training conditions. The general trend is that, across training conditions, the classification accuracy is the highest for the prototypes, and decreases in the order of low-, medium- and high-level distortion test patterns. Moreover, it seems that in each of the low, medium, and high-distortion training conditions, the old patterns are classified with virtually the same accuracy as the novel patterns with the same level of distortion. As in the end of the training phase, the overall classification accuracy is the highest for the low-distortion condition, followed by the medium and mixed-distortion conditions, and is the lowest for the high-distortion condition.

More importantly, the main goal of the study is to find out whether training on patterns with a higher level of distortions will lead to better generalization performance on novel patterns with the same level of distortions. A closer inspection of Fig. 2 revealed that, contrary to the general expectation, the novel high distortions were classified with notably lower accuracy in the high-distortion training condition than in the other conditions with lower levels of distortions. Indeed, the novel high distortions were classified with virtually the same accuracy in the low-distortion and medium-distortion conditions, suggesting that there is little advantage in training with a higher level of distortions in order to improve generalization performance. Similarly, the novel medium distortions were classified with lower accuracy in the medium-distortion and mixed-distortion conditions than in the low-distortion condition.

To confirm the observed patterns of generalization, we conducted a 4x4 mixed-model ANOVA, using condition (low, medium, high, mixed) and novel pattern type (prototype, new-low, new-medium, new-high) as factors. The analysis revealed a significant main effect of pattern type, F(2.62, 779.36) = 128.5, p < .001, η2 = .092; a significant main effect of training condition, F(3,300) = 15.35, p < .001, η2 = .091; and a significant interaction between the two factors, F(7.79, 779.36) = 4.4, p < .001, η2 = .010. For the novel high-distortion patterns, the mean proportion of correct responses is significantly lower in the high condition (M = .512) than in the medium condition (M = .631), t(150.7) = 4.024, p < .001, the mixed condition (M = .593), t(146.8) = 2.655, p = .009, and the low condition (M = .637), t(135.5) = 4.786, p < .001. For the novel medium-distortion patterns, the mean proportion of correct responses is significantly lower in the medium condition (M = .692) than in the low condition (M = .771), t(146.8) = 2.631, p = .009.

Additionally, to confirm the observed null effect of old vs. new patterns on the test performance, we conducted a 4 x 2 mixed-model ANOVA with training conditions (low, medium, mixed, high) and old/new pattern types (the new pattern types are restricted to consist of only low, medium and high-distortions for the respective training conditions) as factors. Again, the analysis captured the significant main effect of training condition, F(3,300) = 45.46, p < .001, η2 = .313. On the other hand, there is neither a significant main effect of old/new pattern types (p = .309) or interaction between the two factors (p = .347). As reported in Table 1, there is very small difference in the mean classification accuracies of the old patterns and the corresponding new patterns in each condition.

Figure 3 shows the frequency distributions of individual test accuracies that are computed by averaging the proportions of correct responses over all novel pattern types. As can be seen, there is a high level of variability in the individual accuracies across the training conditions. To ensure that the qualitative patterns observed are not contaminated by the data from the subjects with idiosyncratic test performance, we repeated the same statistical analyses on a selection of subjects with adequate overall test performance. Specifically, we defined the subject with adequate test performance as the proportion of subjects with the highest 90% of individual test accuracies in each condition. As a result, there remained 70 subjects in the low training condition, 71 in the medium training condition, 68 in the high training condition and 67 in the mixed training condition. Even after removing the lower performing subjects from all conditions, the classification accuracy for novel high distortions in the high- training condition (M = 0.538) remained significantly lower than in the medium training condition (M=0.661), t(135.0)=4.345, p< .001, the mixed training condition (M=0.626), t(132.7) = 3.045, p = .008, and the low training condition (M=0.657),t(119.4) = 4.660, p< .001. Moreover, there still seems to be practically no difference in the classification accuracy of novel high distortions in the low and medium training conditions. For the novel medium-distortion patterns, the classification accuracy in the medium training condition (M = 0.732) also remained significantly lower than the low training condition (M = .809), t(125.4) = 3.210, p = .002. As shown in Figure 4, the overall patterns of transfer results stayed the same as in the analysis in which all subjects were included.

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

\*\* Welch t-tests were conducted which assume unequal population variances

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table 1 |  | **Pattern Type** | |  | |
| **Condition** |  | Old | New | |
| Low |  | 0.860 | 0.873 | |
| Medium |  | 0.698 | 0.692 | |
| Mixed |  | 0.697 | 0.683 | |
| High |  | 0.532 | 0.512 | |

Note. For the mixed condition, the results are averaged across the patterns with different levels of distortions.

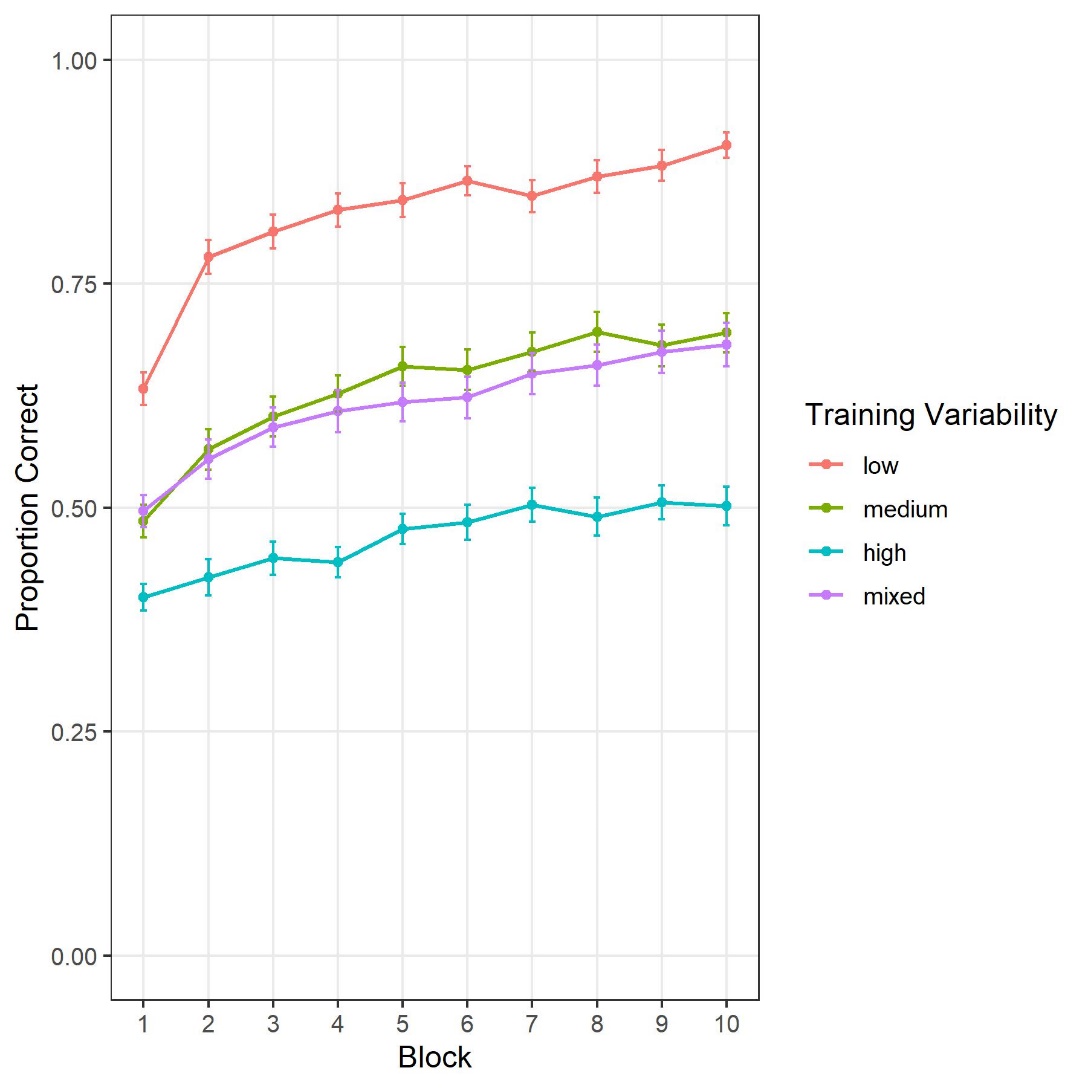


Figure 1

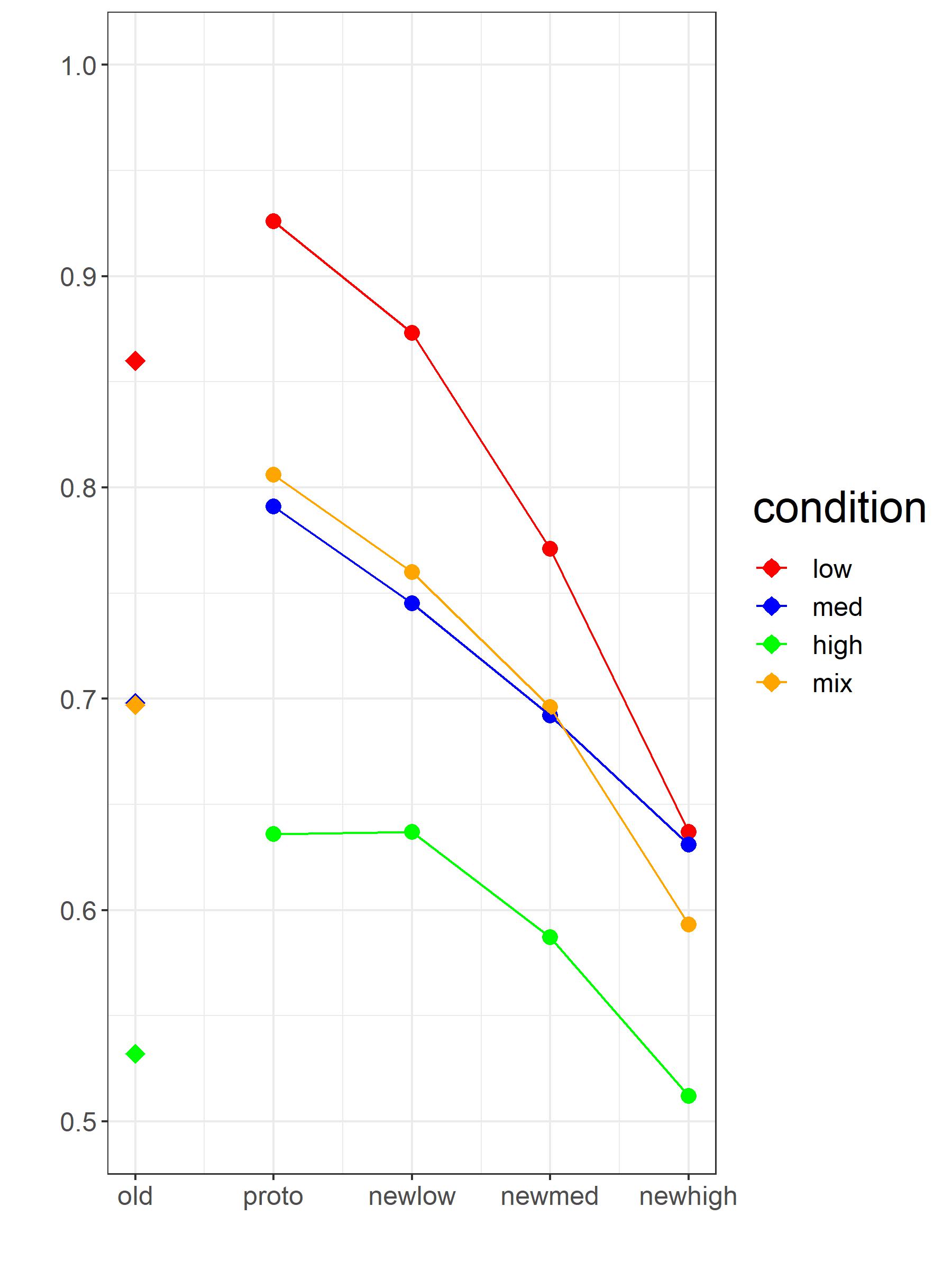


Figure 2

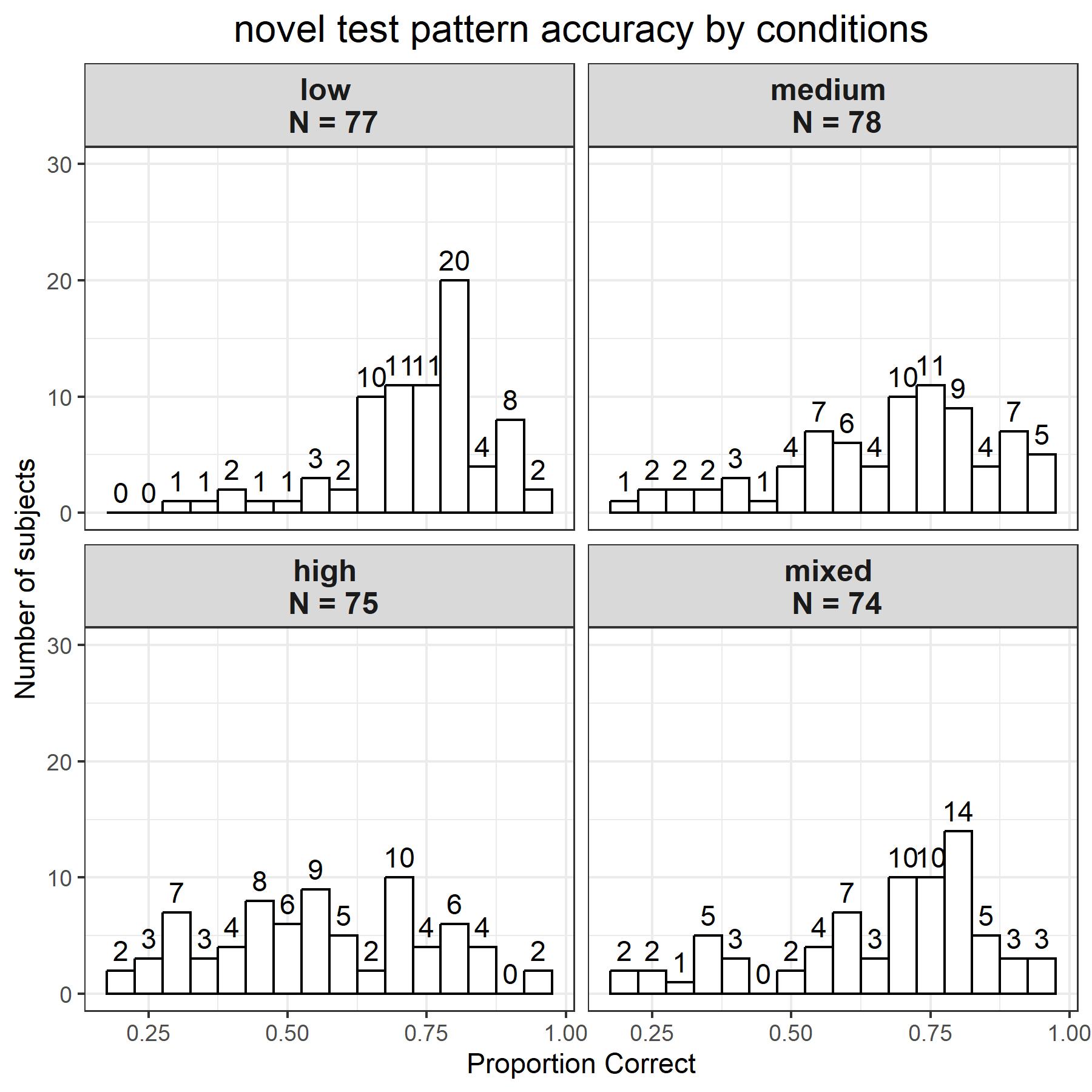


Figure 3

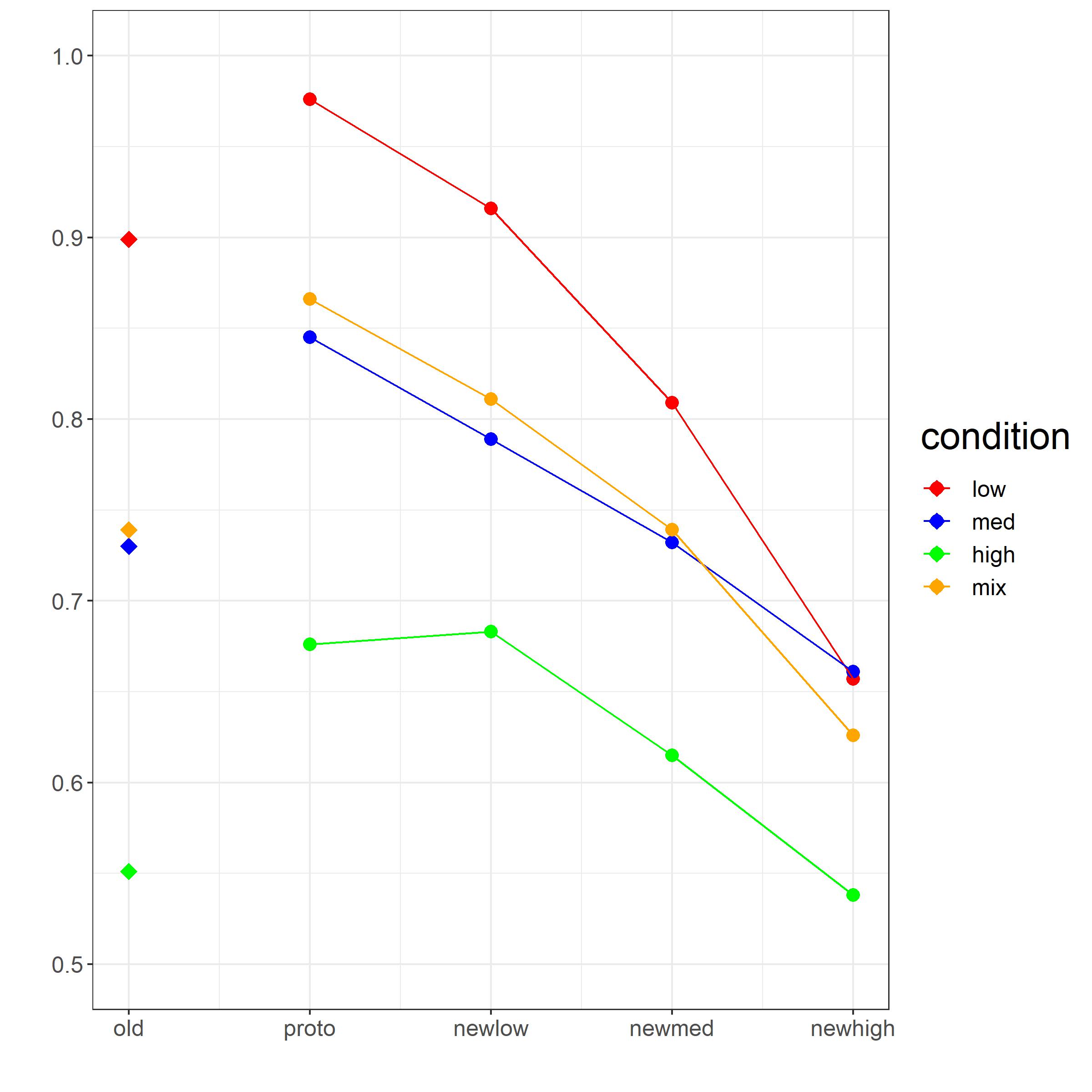


Figure 4