Decoupling levels of learning: behavioral evidence for dissociable low- and high-level structure learning

Beáta Tünde Szabó[1,2], József Fiser[1,2]
[1] Department of Cognitive Science, Central European University, Vienna, Austria,
[2] Center for Cognitive Computation, Central European University, Vienna, Austria

Abstract

Hierarchical Bayesian models offer a unified framework for understanding both learning and meta-learning — the transfer of abstract knowledge across tasks. We investigate whether these two processes are dissociable through a novel statistical learning paradigm that combines low-level shape pair structures with a higher-order color-based rule. Participants viewed shape scenes organized into covert pairs with consistent color contrast patterns (the *pepita* rule), followed by tests assessing recognition of both pair-level and meta-structural regularities. Subject-level analyses revealed three learner profiles: (1) those who acquired both low- and high-level structures, (2) those who learned only low-level pairs, and (3) non-learners. Notably, strong low-level learning was a prerequisite for successful meta-learning, aligning with predictions of hierarchical models. These findings support a behavioral dissociation between learning and meta-learning and highlight individual differences in abstract knowledge acquisition and transfer.

Introduction

Hierarchical Bayesian models offer a powerful framework for understanding how people acquire and apply abstract knowledge in inductive reasoning (Lake et al., 2015; Kemp & Tenenbaum, 2008). Within this framework, *meta-learning* — or "learning to learn" — describes how abstract knowledge transfers across tasks. We test the hypothesis that while a unified hierarchical Bayesian model can capture both learning and meta-learning (Székely et al. 2024), these two processes — (1) learning within a task and (2) using abstract knowledge to guide learning across tasks — are dissociable across observers. If correct, this hypothesis further implies that — assuming a hierarchical model with a meta-structure (MS) guiding subordinate task-level structures (Si) — the relative influence of the meta-structure on learning varies across individuals, potentially reflecting different learning strategies (Holton et al., 2025). As a first step in exploring this, we introduce a novel statistical learning paradigm paired with subject-level analyses of low- and high-level structural representations, without transfer. Preliminary results reveal distinct participant clusters: some learn both the meta- and lower-level structures, others focus on the latter, and a third group shows minimal structure engagement. These consistent groupings support our proposed dissociation and highlight meaningful individual differences in learning low- and higher level structures.

Experimental design

Our experiments employed a standard statistical learning design (Fiser & Aslin, 2001) with two phases: familiarization and test, with the main difference in the shapes being colored instead of black as shown in Fig. 1.

Familiarization phase: participants (recruited online, via Prolific, n=50 with 1 technical exclusion and 4 exclusions for failing the attention tests) viewed scenes, each composed of shapes rendered in multiple colors. The 12 shapes were covertly organized into 6 distinct pairs (inventory). In the first half of the familiarization phase, all shapes were shown in grey, while in the second half, they were pseudo-randomly colored in blue or green in equal proportions. Crucially, each pair always featured contrasting colors, resulting in a distinctive pepita pattern. In total, participants were exposed to 288 scenes over a 15-minute training period.

Test phase: Participants completed sequential 2AFC trials, choosing the more familiar scene.

Pair test. Assessed learning of the pairs (in grey). True scenes had a real pair, foils had a distractor pair from the same shapes in different pairing.
Colored scene test. Assessed learning of the pepita color rule. Both scenes had true pairs, but only the true scene followed the rule. In *blocked* tests, colors were grouped into two blocks; in *unblocked*, they formed a checkerboard.
Colored control tests. Assessed bias toward color layouts. Both scenes followed the color rule and had true pairs, but in one scene the colors were blocked, in the other unblocked.

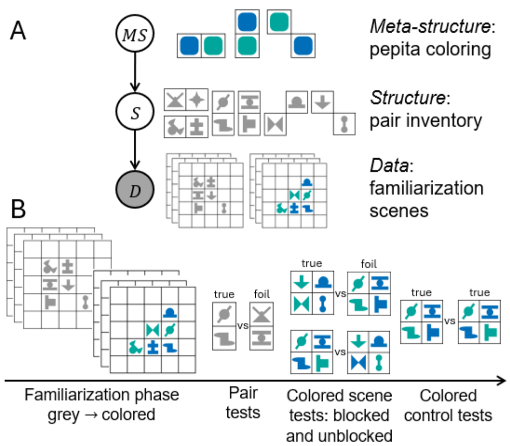

Figure 1: A) Experimental design B) Experimental procedure

Results

Subjects answering based purely on pair familiarity can be distinguished from ones extracting the meta-structure (pepita-rule) within a pair structure (Fig. 2). For the majority of subjects, item-wise performance in the *pair test* predicts their performance

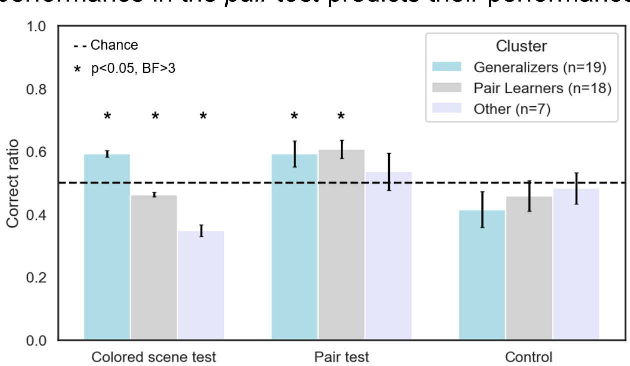

Figure 2: Cluster performances on the *colored scene tests*, *pair tests*, and *control tests* (mean+SEM). Clustering was performed using agglomerative hierarchical clustering with Ward's linkage method (Robidoux et al., 2014) based on only the *colored scene test* performance.

in the *colored scene tests* with additional gains attributable to learning the color distribution (Fig. 3). The shape information dominates color as shown by the different patterns in the *colored scene test* (Fig 3). The *control test* results show no performance differences across clusters in Fig. 2, but large inter-individual variability indicates differing layout preferences (e.g., blocked vs. checkerboard patterns). The control results signal wide variation in biases over the metastructure. Also, the *Pair-learner* cluster and the *Other* cluster's performance in the *colored scene test* is below chance, suggesting additional systematic biases specific to those clusters.

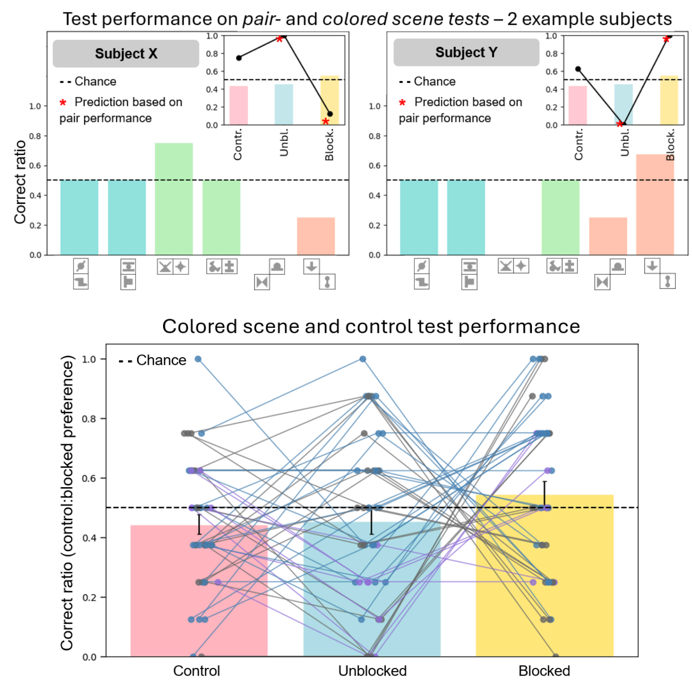

Figure 3: Top. Example subjects with distinctively different color test signatures (inset), explained by their detailed pair performance. Bottom. Same signature is present for the majority of subjects, signaling a strong reliance on the low-level structure learning.

Discussion

Our results show that learning low- and high-level structures can be dissociated and vary across individuals — even without requiring transfer. Some participants grasp abstract meta-structures, others rely on surface familiarity, and some engage minimally, with biases that need further investigation. The clear clustering underscores distinct learning strategies, pointing to how abstract knowledge may shape and potentially transfer across tasks.

References

Fiser J, Aslin RN (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. Psychol Sci 12:499–504.

Holton, E., Braun, L., Thompson, J., Grohn, J., & Summerfield, C. (2025). Humans and neural networks show similar patterns of transfer and interference during continual learning.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. Proceedings of the National Academy of Sciences, 105(31), 10687-10692.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. Science, 350(6266), 1332-1338.

Robidoux, S., & Pritchard, S. C. (2014). Hierarchical clustering analysis of reading aloud data: A new technique for evaluating the performance of computational models. *Frontiers in Psychology*, *5*, 267.

Székely, A., Török, B., Kiss, M., Janacsek, K., Németh, D., & Orbán, G. (2024). Identifying transfer learning in the reshaping of inductive biases. *Open Mind*, *8*, 1107-1128.