Strategy Use and Latent Cognitive Regimes in Perceptual and Relational Classification

Nadejda Bocheva[0000-0001-7792-7742] and Bilyana Genove[0000-0002-2170-298X]

Institute of Neurobiology, Bulgarian Academy of Sciences, Bulgaria

Abstract. Understanding how individuals adapt their categorization strategies requires methods that detect changes in internal processing, rather than relying solely on accuracy or response times. This study addressed this challenge through three experiments by investigating how the brain learns categories under structural ambiguity. All experiments used the same large set of unique stimuli, while varying the classification rules to increase cognitive complexity: a perceptual rule (Experiment 1), a structural rule (Experiment 2), and a relational rule (Experiment 3). As no stimulus was repeated and feature values increased monotonically within categories, participants could not use exemplar memorization or simple perceptual clustering to infer category boundaries. Data were analyzed using Hidden Markov Model (HMM) segmentation, which identifies latent behavioral regimes directly from trial-level performance. The resulting states were independent of low-level and spatial stimulus features, indicating that they reflected internal processing modes rather than stimulus properties. Boundary uncertainty further contributed to this stimulus independence, as it could not be reliably inferred from perceptual similarity alone. The perceptual and structural rules produced stable, efficiency-based regimes, whereas the relational rule revealed a near-absorbing relational regime alongside several short-lived heuristic regimes. Cross-experiment trajectories demonstrated substantial individual differences in cognitive flexibility, with only a minority of participants discovering the relational rule. These findings demonstrate that HMM segmentation can identify hidden cognitive states, with transitions between them serving as a mechanism to adaptively reduce model-based computations under uncertainty.

Keywords: Categorization, Hidden Markov Model, Latent Cognitive Regimes, Low-level and Spatial Image features, Uncertainty

1 Introduction

Categorization allows humans to reduce sensory complexity and respond adaptively by grouping objects, events, and actions according to relevant features. Classic theories emphasize perceptual similarity and rule-based processing [9, 13, 1], whereas the application of complex or relational rules requires greater cognitive control and working memory [7, 3]. Rather than relying on a single approach, individuals flexibly shift between heuristics, rule application, and mixed strategies depending on task demands or internal fluctuations in attention and control [2, 4]. Traditional metrics such as accuracy and reaction time capture overall performance but may overlook dynamic variations and individual differences. Recent research underscores the value of analyzing behavioral time series to reveal transitions between cognitive modes, although identifying these modes without strong assumptions remains challenging.

To address this problem, a data-driven Hidden Markov Model (HMM) was applied to categorize trial-level behavior across three experiments with varying cognitive demands. Each experiment used the same set of 196 unique stimuli, which varied in color dominance and spatial grouping. This design ensured that perceptual input remained constant while the rule defining category membership changed across experiments: color dominance (Experiment 1), number of clusters (Experiment 2), and their relationship (Experiment 3). Stimuli were constructed so that feature values increased monotonically within each category and no item was repeated, requiring participants to rely on rule-based inference rather than memorization.

HMM segmentation provides a principled framework for modeling trial-wise behavior as a sequence of latent states that differ in accuracy, response time, or other behavioral signatures [5, 8, 11]. This approach has been used to characterize strategy switching, learning dynamics, and fluctuations in cognitive control across a range of tasks [14]. We used this approach alongside behavioral analyses to answer three questions: (1) Do participants use the intended rule, and how does performance depend on stimulus strength? (2) Are categories confounded by low-level or spatial features? (3) What behavioral regimes emerge in each experiment, and how do they differ across participants?

Across experiments, participants exhibited distinct perceptual, heuristic, and relational styles that systematically changed in response to boundary proximity and sequential trial history. Some individuals remained in a single state, while others transitioned between states as the rules became more complex. By integrating a structured modeling approach with a carefully designed set of stimuli, this study offers a comprehensive perspective on how categorization behavior organizes into cognitive modes and varies across individuals and tasks under uncertainty.

2 Methods

2.1 Participants

Twenty-eight adults (16 females, 12 males; ages 21–55 years) participated in three experiments carried out on separate days, with at least 1 week between sessions. The study protocol was approved by the Ethical Board of the Institute of Neurobiology, Bulgarian Academy of Sciences. All participants provided written informed consent in accordance with the Declaration of Helsinki.

2.2 Stimuli

All stimuli consisted of 24 colored squares (red or yellow) presented on a black background within a 10 × 10° visual angle region. Red square counts (3–9, 15–21, red–yellow ratios from 3/21 to 21/3) and cluster numbers (3–9, 15–21) yielded 196 unique stimuli by combining all ratio and cluster options. Figure 1 shows examples of extreme color dominance and clustering.

2.3 Procedure

All 196 stimuli were used in each experiment, presented centrally on a 21″ Dell Trinitron monitor (60 Hz, 1280 × 1024 px) for 2 s at a viewing distance of 57 cm.

Experiments differed only in the classification rule: color dominance (Experiment 1), number of clusters (Experiment 2: low = 3–9, high = 15–21), or a relational rule (Experiment 3: one category where color dominance and clustering varied in the same direction; the other, incongruent variation). Stimuli within each experiment were organized so that feature values progressed monotonically within each category, with a fixed discontinuity between them, and no item was repeated—requiring rule-based inference from feedback rather than recall. The trial order was randomized, and the experiment sequence was counterbalanced across participants. As neither the distribution of stimuli nor their order marked the category transition, participants' internal estimates of the boundary could vary across trials. Responses were made via a joystick; auditory feedback was provided after each trial. Accuracy and response time (RT) were recorded.

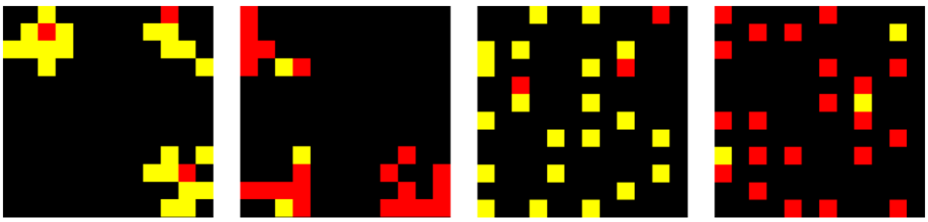

Fig. 1. Examples of stimuli illustrating the most extreme combinations of color dominance and spatial clustering. Each panel shows a pattern of 24 colored squares (red or yellow) arranged within a 10 × 10° region. The examples reveal cases with strong red or yellow dominance and cases with minimal or maximal clustering, representative of the full stimulus space used in the experiments.

2.4 Behavioral Measures

Accuracy was coded as correct/incorrect. RTs were trimmed (0.2–5 s) and log-transformed. These metrics supported descriptive and HMM analyses. For Experiments 1 and 2, stimulus strength was (number of red squares − 12)/12 and (number of clusters − 12)/12, with the category boundary at 0 (range: −0.75 to 0.75).

2.5 Low-level and Spatial Feature Extraction

Although only color dominance and clustering were manipulated, participants might rely on additional visual properties. To assess reliance on non-target features, we quantified brightness, contrast, spatial frequency, and entropy for each image, along with geometric descriptors: horizon-tal/vertical centroids (cx/cy), spreads (LR/UD), cluster count (nC), and the size of the largest cluster (largest), computed for all, red, and yellow pixels. These variants were denoted _all, _red, _yellow.

2.6 Hidden Markov Model (HMM) Segmentation

We applied HMMs to trial-wise responses in each experiment, using accuracy (binary) and log-RT as emissions. HMMs with 5 states (Experiments 1/2) and 6 states (Experiment 3) were selected based on BIC, log-likelihood, and explainability. Models were fit with the Baum–Welch algorithm (20 starts), keeping the best log-likelihood. The Viterbi algorithm decoded state sequences; we computed state-level accuracy, RT, dwell time, runs, transitions, and regime usage. For cross-experiment comparisons, states were grouped into higher-level regimes (perceptual, heuristic, relational) based on behavioral metrics and transitions. These labels informed the Sankey diagram.

2.7 Statistical Analyses

Differences in low-level and spatial features among HMM states were assessed with Welch's ANOVA and the Kruskal–Wallis test due to violations of normality and homogeneity of variance. To evaluate whether unintended low-level or spatial properties could provide an alternative basis for category learning, we quantified category differences for each of the 24 features using independent-samples t-tests, computed Cohen's d, and performed hierarchical clustering ($1 - |r|$, average linkage) on the full feature set. To test if the unintended features could support discrimination, we trained logistic regression and random forest classifiers on the 24 features (excluding rule-defining ones) and assessed accuracy and stability. All analyses used R [10], depmixS4 [15] for HMMs, and dbscan [6] for clustering.

3 Results

3.1 Behavioral Performance

Accuracy declined as task difficulty increased: high in Experiment 1, intermediate in Experiment 2, and lowest and most variable in Experiment 3. Correct responses were faster than incorrect ones in Experiments 1 and 2, but in Experiment 3, RTs for correct and incorrect responses overlapped, suggesting correctness was not linked to speed in the relational task. All experiments showed multimodal RT distributions, indicating heterogeneous processing modes (Figure 2).

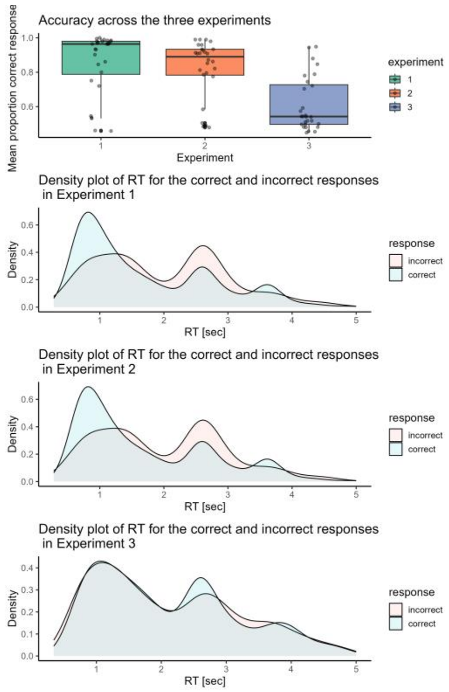

Fig. 2. Behavioral performance across experiments. Response-time (RT) distributions are shown separately for correct and incorrect responses. Accuracy decreases, and RT increases with task difficulty, and Experiment 3 shows substantial overlap between correct and incorrect RTs.

3.2 Accuracy as a Function of Stimulus Strength (Experiments 1 and 2)

Accuracy was analyzed as a function of stimulus strength (Figure 3). In Experiments 1 and 2, accuracy increased with distance from the category boundary, confirming reliance on the intended rule (color dominance or clustering). Near the boundary, performance was more variable, reflecting ambiguity. Experiment 3 lacked a one-dimensional stimulus strength, so no analogous analysis was done.

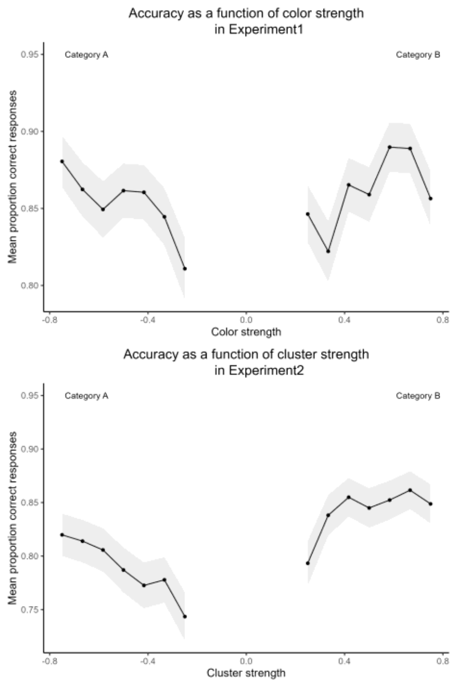

Fig. 3. Accuracy as a function of stimulus strength in Experiments 1 and 2. Stimulus strength is expressed relative to the category boundary (0), with negative values belonging to Category A and positive values to Category B. Accuracy increases with distance from the boundary in both experiments, indicating better performance for stimuli farther from the category boundary.

3.3 Low-level and Spatial Feature Differences between Categories

To assess whether unintended low-level or spatial properties of the stimuli could provide an alternative basis for category learning, we analyzed the structure of 24 low-level and spatial features extracted from the stimuli. Hierarchical clustering (Figure 4) revealed three groups—color (e.g., LR/UD/cx/cy/nC/largest for red and yellow channels), spatial/geometric (e.g., LR_all, UD_all, cx_all, cy_all, nC_all, largest_all), and mixed features comprising entropy, edges, symmetry, and spatial frequency —but none aligned with category structure.

In each experiment, we quantified the differences in features between categories. Across all three experiments, several low-level and spatial features showed statistically significant differences between the categories at the stimulus level. However, these differences were associated with the rule-defining dimensions—such as color dominance or clustering—rather than providing independent information for distinguishing the categories. The effect sizes were significantly larger for features directly related to the rule dimensions in Experiments 1 and 2, specifically for color dominance in Experiment 1 and clustering in Experiment 2. In contrast, the effect sizes in Experiment 3 were small to moderate. Importantly, classifiers trained on the unintended features performed at chance levels across all experiments, with logistic regression accuracy hovering around 0.50 and random forests exhibiting instability, indicating that these low-level or spatial features did not provide a reliable basis for classification. Therefore, participants were unable to solve the task using these cues.

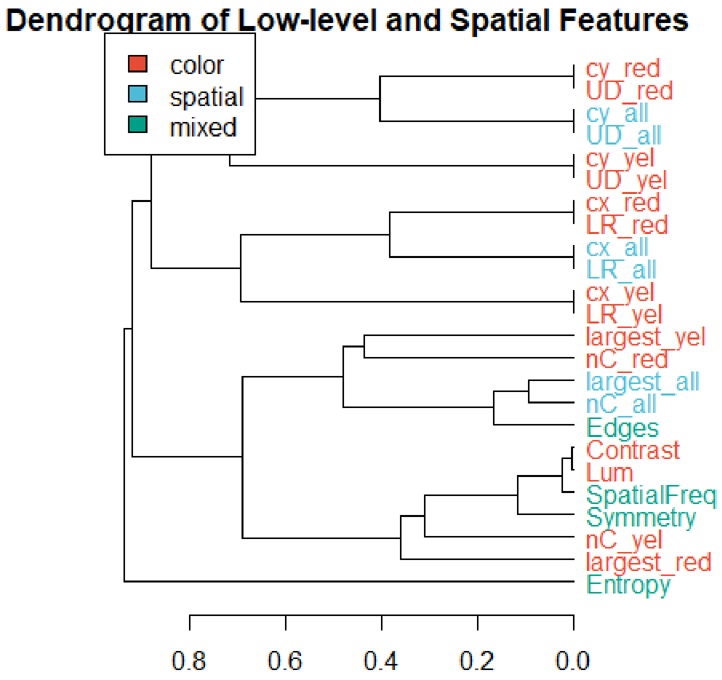

Fig. 4. Hierarchical clustering of the 24 low-level and positional features computed across the full stimulus set. The dendrogram reveals three coherent groups: color-related features (red labels), spatial/shape features (blue labels), and mixed features (green labels). The clustering reflects correlations among the features but does not correspond to the category structure. Combined with the classification analyses, the dendrogram confirms that the unintended features do not provide category-diagnostic information across the three experiments.

3.4 State-level Behavioral Regimes

The hidden Markov model (HMM) detected distinct behavioral regimes (states) within each experiment (see Figure 5). In Experiment 1, the states differed primarily in processing effectiveness. Two efficient perceptual states (States 1 and 2) demonstrated high accuracy and fast reaction times (RTs); one showed intermediate performance (State 3); and the other two showed lower accuracy with fast or slow responses (States 4–5). In Experiment 2, the states formed a gradient ranging from a low-efficiency (State 1) to an intermediate mixed state (State 4) and finally to a high-efficiency state (State 5). The broader range of accuracy-RT combinations in this experiment reflects increasing task complexity.

In Experiment 3, accuracy and RT were partially decoupled: one state showed high accuracy with moderate RTs (State 2), several (States 3–5) showed varying efficiency (e.g., fast but inaccurate, slow but accurate, or variable), and one showed chance-level accuracy with short RTs (State 6), denoting a non-cognitive state. These patterns indicate that the latent states capture distinct behavioral modes, but efficiency alone does not determine their cognitive identity.

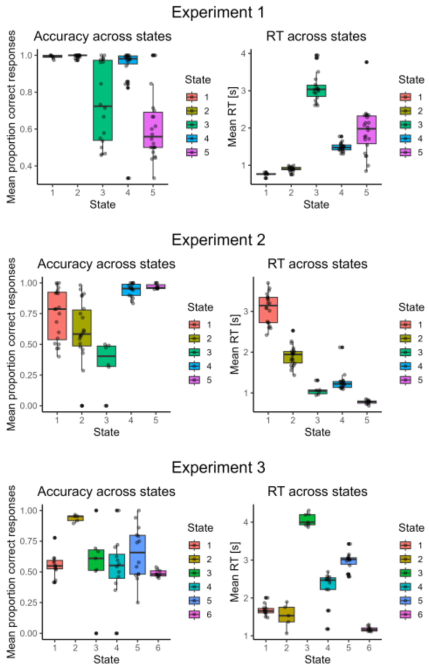

Fig. 5. State-level behavioral signatures. Accuracy and response time (RT) are shown for each HMM state. Experiments 1 and 2 show accuracy–RT covariation, whereas Experiment 3 shows a decoupling between accuracy and RT.

3.5 Low-level Stimulus Features across HMM States

Although the HMM was fit exclusively to behavioral variables (accuracy and response time), stimulus features can correlate with latent states if they systematically influence behavior. To test this possibility, we examined whether low-level geometric or spatial properties of the stimuli differed across states. In Experiment 1, entropy varied reliably across states (Kruskal–Wallis $p < .001$), consistent with its link to color dominance. In Experiments 2 and 3, no low-level or spatial features differed significantly across states (all $p > .18$), despite small numerical differences in some features, indicating that the regimes reflected behavioral strategies rather than stimulus-driven segmentation.

We next examined whether each state's accuracy depended on the relevant features. In Experiment 1, only State 1's accuracy tracked color strength, identifying it as a perceptual regime. In Experiment 2, States 1, 4, and 5 followed the grouping rule, while State 3 showed a reversed or irregular pattern, indicating an incorrect grouping strategy. State 2 was unrelated to cluster strength. In Experiment 3, only State 2 was sensitive to the relationship between color strength and grouping, identifying it as a relational regime. The other states across experiments showed no clear link to the intended features but did display consistent behavioral patterns, so we classified them as strategy-based. State 6 in Experiment 3 showed neither feature dependence nor a consistent behavioral pattern, so we labeled it non-cognitive.

3.6 State Stability and Transitions

Dwell times and transition likelihoods revealed clear differences in regime stability between experiments (Figure 6). In Experiment 1, most participants occupied one of two stable attractor regimes (States 3 and 5), corresponding to consistent perceptual or cognitive strategies. A smaller subset relied on stable strategy-based regimes (State 1). Only a minority showed substantial occupancy of the unstable switching regimes (States 2 and 4). Overall, Experiment 1 exhibited the highest degree of regime stability, consistent with the relative simplicity of the task.

In Experiment 2, dwell times were more evenly distributed across the low-efficiency, intermediate mixed, and high-efficiency strategy-based states (States 1, 4, 5). Only two participants occupied the incorrect grouping-rule regime (State 3). A thin but consistent band of unstable switching (State 2) was present but not dominant. All states exhibited high self-transition probabilities, indicating the presence of multiple stable strategies rather than a single dominant regime.

In Experiment 3, the relational state (State 2) showed long dwell times and high self-transition probability, indicating that relational reasoning, once discovered, becomes a stable processing mode. In contrast, State 6 also exhibited near-absorbing dynamics, but its behavioral profile (chance accuracy, idiosyncratic RTs) indicates that it reflects a non-cognitive absorbing state rather than a meaningful regime. The remaining strategy-based states (States 3–5) were shorter-lived and more variable.

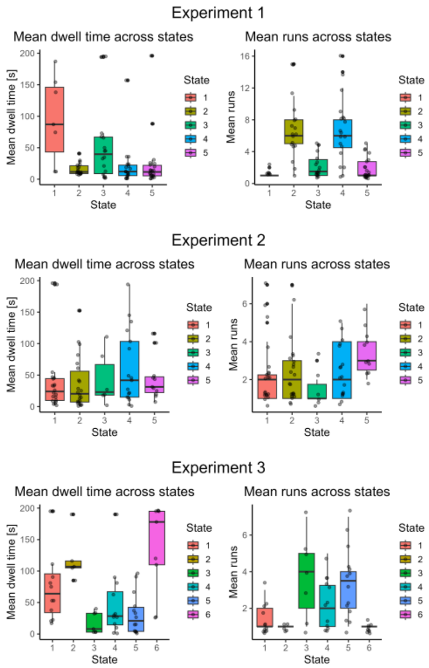

Fig. 6. State stability and transitions. Mean dwell times and state-to-state transition chances are shown for each experiment. Experiment 1 exhibits a dominant stable state; Experiment 3 shows one near-absorbing relational state.

3.7 Regime usage

Heatmaps of regime usage revealed clear differences in how participants allocated their time across latent regimes in the three experiments (Figure 7). In Experiment 1, most participants tended to occupy either efficient or less efficient perceptual states, with a smaller group relying on strategy-based states. Only a minority exhibited considerable occupancy of unstable switching states. This pattern reflects the relative simplicity of the perceptual task and the presence of strong attractor regimes. In Experiment 2, regime usage was more heterogeneous. Most participants relied upon strategy-based regimes (low-efficiency, intermediate mixed, or high-efficiency), whereas only a small subset predominantly occupied the incorrect grouping-rule state. Experiment 3 showed a tripartite structure: most participants used strategy-based regimes, a small group found the relational state, and a minority exhibited persistent switching or entered an absorbing, non-cognitive-noise state. Three cognitive profiles appeared across experiments: (1) stable heuristic users, (2) unstable switchers, and (3) rare relational-rule users. These profiles were consistent across experiments.

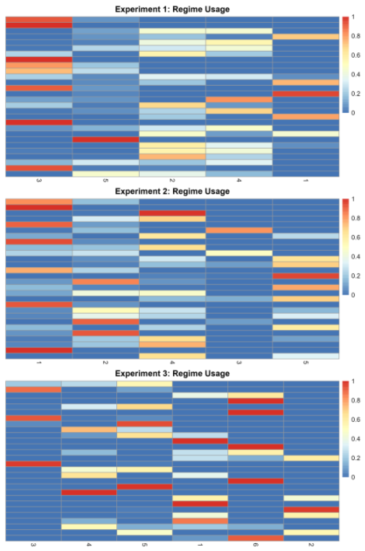

Fig. 7. Regime usage across participants. Each row corresponds to a participant, and each column corresponds to an HMM state. Cell color indicates the proportion of trials assigned to each state. The three panels show regime usage across Experiments 1–3. Stable heuristic users show strong concentration in heuristic states across all experiments. Unstable switchers show high occupancy of unstable states. A small subset of participants transitions to the relational-rule regime in Experiment 3 (State 2). Several participants fell into an absorbing state in Experiment 3 (State 6). Together, the heatmaps reveal three robust cognitive profiles: stable heuristic users, unstable switchers, and rare relational-rule users.

3.8 Cross-experiment Transitions

To track regime transitions across the tasks, we used a Sankey diagram (Figure 8). Most participants started in perceptual regimes in Experiment 1, then moved to various strategy-based regimes in Experiment 2. In Experiment 3, some found the relational state, others remained in strategy-based or unstable regimes, and a few entered an absorbing, non-meaningful state. This progression—from perceptual to strategy-based to (rarely) relational reasoning—corresponds to the cognitive profiles seen in regime usage and confirms individual consistency among tasks.

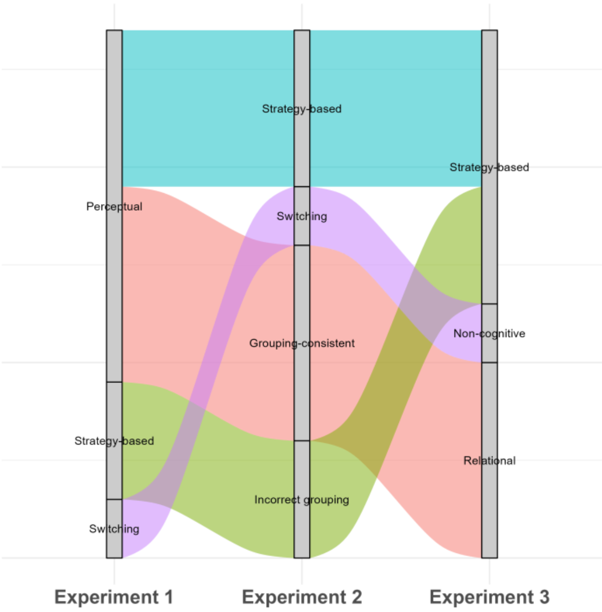

Fig. 8. Cross-experiment transitions between behavioral regimes. The Sankey diagram shows how participants move from perceptual, strategy-based, and switching regimes in Experiment 1 (left) to grouping-consistent, grouping-inconsistent, strategy-based, or switching regimes in Experiment 2 (middle), and to relational, strategy-based, switching, or non-cognitive regimes in Experiment 3 (right). Each ribbon represents the number of participants transitioning among dominant regimes. Participants who adopt a stable strategy in one experiment tend to remain stable in subsequent tasks, whereas switching participants show greater variability across experiments.

4 Discussion

This study investigated how participants classified visual patterns as the rules became more complex, using a Hidden Markov Model to analyze their behavior. Both accuracy and response time reflected task difficulty, with the most complex rule (Experiment 3) producing the lowest accuracy and greatest variability. Results showed that participants used the intended rules in Experiments 1 and 2, whereas the relational rule in Experiment 3 did not show a clear strength–performance link.

A key feature of the design was the use of 196 unique, non-repeated stimuli. This choice prevented exemplar memorization and minimized stable perceptual clusters, reducing the chance that HMM-identified states reflected stimulus-driven segmentation. Moreover, category boundaries were uncertain due to monotonic feature progressions and random presentation order, requiring participants to estimate boundaries using feedback and noisy evidence—especially in the relational task, where the relevant dimension was structural.

Consistent with this, low-level and spatial features did not differ across categories or HMM states in Experiments 2 and 3. In Experiment 1, the only feature that showed differences across states was entropy, which is directly related to the intended perceptual dimension. Thus, latent regimes primarily reflected internal processing modes rather than stimulus properties.

The HMM analysis showed distinct behavioral regimes: in Experiments 1 and 2, accuracy and response time covaried, reflecting differences in how efficiently participants used the relevant rule; in Experiment 3, accuracy and response time decoupled, pointing to a shift in strategy for the more complex relational task. State stability also differed: Experiment 1 was dominated by a single perceptual regime; Experiment 2 had multiple stable regimes (including grouping-based and strategy-driven states); and Experiment 3 included one stable relational regime, several short-lived strategy-based states, and a persistent non-cognitive state.

Individual differences were most pronounced in Experiment 3, with some participants consistently using the relational regime and others relying on heuristics or unstable strategies. The progression from perceptual to mixed to divergent relational/heuristic strategies across experiments reflects increasing cognitive demands and emerging individual differences in relational reasoning.

These results show the value of HMM segmentation for uncovering latent cognitive structures in behavioral time-series data. Unlike standard analyses, HMMs extract internal regimes, assess their stability, and track individual shifts throughout tasks. In line with previous work [12], our findings show that these regimes reflect internal processing modes rather than stimulus features.

To summarize, participants did not use low-level visual features to solve the tasks, and relational reasoning elicited distinct and more variable behavioral regimes than perceptual or structural classification. Using HMM segmentation provides a robust foundation for examining how people adapt strategies under uncertainty and flexibly solve complex rule-learning problems.

5 Limitations

This study has several limitations. First, while the HMM identifies latent regimes from behavioral data, their interpretation relies on the available behavioral measures; adding modalities such as eye-tracking or neural data may provide extra constraints. Second, the number of states was chosen based on fit and understandability, but other choices may reveal finer distinctions in strategy. Third, the relational task used a specific stimulus set, so the generalizability of the relational regime to other relational tasks is unknown. Finally, although sessions were separated by at least a week, cross-task influences cannot be ruled out. These points suggest directions for future research but do not diminish the main finding: HMM segmentation robustly distinguishes strategy use across classification tasks of increasing complexity.

6 Conclusions

This study shows that categorization behavior can be parsed into latent regimes differing in stability, efficiency, and cognitive complexity. By combining controlled stimulus design and HMM segmentation, we demonstrate that perceptual and structural classification rely on stable, efficient regimes, whereas relational reasoning leads to more variable strategies. Cross-experiment trajectories reveal robust individual differences in cognitive plasticity, with only a few participants discovering the relational rule. Overall, our results show that people don't rely on a single rule when learning categories in uncertain situations. Instead, they flexibly switch between model-based and strategy-based regimes, updating their approach to the tasks. These shifts were not random—they were ways for the mind to cope with confusion or complexity. These outcomes emphasize the value of latent-state modeling for understanding how people adjust their strategies as task demands increase and for studying cognitive flexibility inside complex domains.

Acknowledgments. This work was partially supported by the Bulgarian Science Fund (project "Modeling post-perceptual stages of mental processing and conscious representations of visual information," No KP-06-N52/6, 12.11.2021). We thank Tsvetalin Totev for developing the experimental software.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ashby, F. G., & Maddox, W. T.: Human category learning. Annual Review of Psychology, **56**, 149–178 https://doi.org/10.1146/annurev.psych.56.091103.070217 (2005)
2. Ashby, F. G, Paul, E. J, Maddox, W. T.: COVIS. In: Pothos, E. M., Wills, A. J. (eds). Formal Approaches in Categorization. Cambridge University Press, New York, 65-87 (2011)
3. Badre, D., Kayser, A. S., & D'Esposito, M.: Frontal cortex and the discovery of abstract action rules. Neuron, 66(2), 315–326. https://doi.org/10.1016/j.neuron.2010.03.025 (2010)
4. Collina, J.S., Erdil, G., Xia, M. et al.: Individual-specific strategies inform category learning. Scientific Reports, **15**, 2984 (2025) https://doi.org/10.1038/s41598-024-82219-8
5. Collins, A. G. E., & Frank, M. J.: Cognitive control over learning: Creating, clustering, and generalizing task-set structure. Psychological Review, 120(1), 190–229 https://doi.org/10.1037/a0030852 (2013)
6. Hahsler, M., Piekenbrock, M., & Doran, D.: dbscan: Fast density-based clustering with R. Journal of Statistical Software, 91(1), 1–30 https://doi.org/10.18637/jss.v091.i01(2019)
7. Halford, G. S., Wilson, W. H., & Phillips, S.: Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. Behavioral and Brain Sciences, 21(6), 803–864 ttps://doi.org/10.1017/s0140525x98001769 (1998)

8. Kucharský, Š., Tran, NH., Veldkamp, K. et al.: Hidden Markov Models of Evidence Accumulation in Speeded Decision Tasks. Comput Brain Behav 4, 416–441. https://doi.org/10.1007/s42113-021-00115-0 (2021)
9. Nosofsky, R. M. Attention, similarity, and the identification–categorization relationship. Journal of Experimental Psychology: General, 115(1), 39–57. https://doi.org/10.1037/0096-3445.115.1.39 (1986)
10. R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
11. Rabiner, L. R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, **77**(2), 257–286 https://doi.org/10.1109/5.18626 (1989)
12. Raijmakers, M. E., Schmittmann, V. D., & Visser, I.: Costs and benefits of automatization in category learning of ill-defined rules. Cognitive psychology, 69, 1–24. https://doi.org/10.1016/j.cogpsych.2013.12.002 (2014)
13. Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. Science, 237(4820), 1317–1323 (1987)
14. Villarreal, M., & Lee, M. D. (2024). A Coupled Hidden Markov Model framework for measuring the dynamics of categorization. Journal of Mathematical Psychology, 123, 1–11. https://doi.org/10.1016/j.jmp.2024.102884
15. Visser, I., & Speekenbrink, M. depmixS4: An R package for hidden Markov models. Journal of Statistical Software, **36**(7), 1–21 (2010)