p-ISSN: 2184-2205 e-ISSN: 2184-3414 ISBN: 978-989-35728-4-9 © 2025
https://doi.org/10.36315/2025inpact144

CLASSIFICATION OF VISUAL SCENES BY OVERALL COLORFULNESS

Bilyana Genova, Nadejda Bocheva, & Tsvetalin Totev
Department of Sensory Neurobiology, Institute of Neurobiology, Bulgarian Academy of Sciences, Sofia (Bulgaria)

Abstract

Classifying objects or events is vital for survival and daily life. Categorization learning varies in its structure, stimulus-response associations, and feedback methods. In typical experiments, observers classify objects and receive feedback after each response, gradually associating stimuli with correct responses through trial and error. Determining which features of objects are relevant for classification can be complex. Color is a common visual characteristic used in this process, though colors are spread among multiple objects in natural scenes. This study explored observers' ability to classify visual scenes based on color dominance when the number of objects varied. The stimuli included 21 red and yellow squares within a 10x10 cm black square, with proportions of red and yellow ranging from 6/15 to 15/6. A total of 16 different stimuli were generated based on these color ratios, whereas the groupings of squares varied from 2 to 18 clusters. The classification was based on the rule that stimuli with more red squares were one category, while those with more yellow squares constituted the other. Each stimulus was shown five times in random order, totaling 80 presentations. Thirty-five observers (23 females and 12 males) aged 20 to 69 (average age 45) participated, learning to classify the stimuli through trial and error with feedback provided via sound signals. All observers were unaware of the classification rule. The cumulative sum of responses was formed and normalized by the number of presentations to reflect the observers' alignment with the classification rule. A robust regression method and a generalized mixed model regression analyzed factors influencing response accuracy and time. Results indicated: an improvement in response accuracy and a reduction in response time among observers; accuracy plateauing at about 75% in the final experimental block; just over half of the observers recognized the classification rule at various times during the experiment, observers who did not identify the classification rule changed their strategies more frequently, the percentage of correct responses increased with the color ratio more for dispersed stimuli; response times decreased with distance from the decision boundary between categories. The challenges in quickly classifying stimuli based on color suggest that the spatial characteristics of the objects may be the dominant feature, and separating the scenes into objects may hinder the ability to perceive individual object colors and the image's overall colorfulness accurately.

Keywords: *Vision, category learning, color and spatial characteristics, colorfulness.*

1. Introduction

Classifying objects or events is a complex and multifaceted process vital for survival and daily life. It allows inference of object characteristics based on category membership and predictions for novel events and conditions. Categorization learning varies in its structure, stimulus-response associations, and feedback methods. In typical experiments, observers classify objects and receive feedback after each response, gradually associating stimuli with correct responses through trial and error. Various theories try to describe the cognitive processes involved in classification performance like the exemplar model (e.g., Nosofsky et al., 2022), the prototype model (e.g., Posner & Keele, 1968), the decision boundary theory for categorization (Ashby & Maddox, 1994), and their hybrid versions. The results from the classification studies show that the frequency of stimulus appearance, the feature distribution, the probability of correct response, and the distance from the classification boundary affect the classification performance.

Determining which stimulus characteristics are relevant for classification can be complex. Color is a common visual characteristic used in this process, though colors are spread among multiple objects in natural scenes.

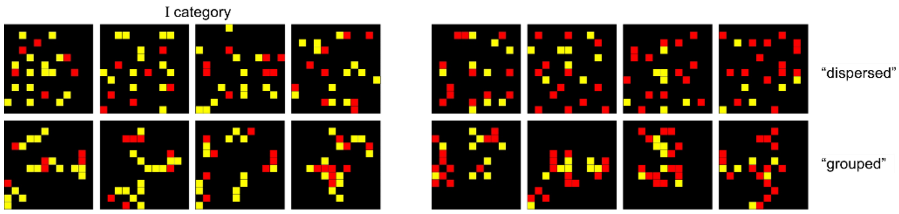
2. Objectives

This study explored observers' ability to classify visual scenes based on color dominance when the number of objects varied. Therefore, we tested whether the feature determining the classification—"colorfulness," a characteristic that is typically viewed as continuous—determines the proportion of correct responses and response time and whether the spatial stimulus characteristics, like element compactness, affect categorization.

3. Methods and design

We used 16 stimuli as each of them consisted of a total of 21 red and yellow squares interspersed within a 10x10 cm black square (Figure 1) with proportions of red and yellow squares 6/15, 7/14, 8/13, 9/12, 12/9, 13/8, 14/7, and 15/6. For each color ratio, two stimuli were generated: "dispersed" – a stimulus with relatively many groupings of squares, and "grouped" – a stimulus with few groupings of squares. The groupings of squares varied from 2 to 18 clusters. The rule for separation into two categories, unbeknownst to the observers, was based on color ratio: stimuli with more red squares were one category, while those with more yellow squares constituted the other. Each stimulus was shown five times in random order, totaling 80 presentations.

Figure 1. Illustration of the set of stimuli.



The stimuli were presented in the middle of a computer screen on a gray background for 3 sec. The observer had to classify the stimulus into one of two categories by pressing the left or right button on a joystick. After giving a response, observers received feedback via a sound signal - two successive high tones for a correct response and a low tone for an incorrect response. Thus, at the beginning of the experiment, the observers did not know the stimulus belonged to any of the two categories. However, they could learn to categorize the stimuli by following the feedback through trial and error.

The stimuli were binocularly viewed from 57 cm and presented on the computer screen (20-inch NEC SpectraView 2090, 1600 × 1200 pixels, 60 Hz refresh rate). A custom program developed in Visual C++ and OpenGL controlled the experiments.

3.1. Participants

Thirty-five naive observers (23 females and 12 males) aged between 20 and 69 (average age 42) participated in the experiment. They gave written informed consent for the study. The experiment was approved by the Ethical Board of the Institute of Neurobiology, Bulgarian Academy of Sciences, and complied with the requirements of the Declaration from Helsinki.

3.2. Statistical analyses

We used Bayesian hierarchical modeling of the response time and the correct response proportion to describe group performance. In separate analyses, we evaluated the contribution of the sequential blocks of trials, the effect of the number of red squares, and the effect of the distance of the patterns from the classification boundary. In all analyses, we included the compactness of the patterns (grouped or dispersed) as a fixed factor, and we considered the participants to be random factors. In the analyses of the experimental effects on performance accuracy, we used the Bernoulli distribution as a likelihood function, whereas for the analyses of response time, we used a shifted lognormal distribution. We used weakly informative priors for the model parameters and evaluated the correctness of the models by posterior predictive check. All models converged (as suggested by trace plots and Gelman-Rubin Rhat value). All these analyses were performed in R (R Core Team, 2020) using the brms package (Bürkner, 2017).

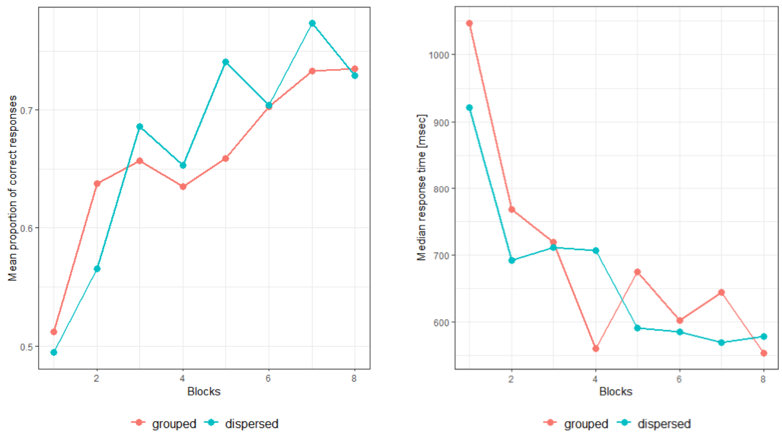
To characterize the process of strategy switching, we formed the cumulative sum of observers' responses and normalized it by the number of stimulus presentations. This representation of the data allows the evaluation of the changes in the classification strategy and its correspondence to the classification rule.

If an observer always gives correct responses, the maximum value of the normalized cumulative sum will be 1.0, and a linear function with a slope of 1.0 can describe it. If the observer answers randomly, the maximum value of the normalized cumulative sum will be 0.5. For all other cases, the normalized cumulative sum will have a slope different from the ideal performance. If the observers use the same strategy, the dependence of the cumulative sum on the series of stimulus presentations will be linear but with random fluctuations due to peculiarities in the series of stimuli or to inattention, fatigue, and others. We applied a robust regression method (repeated median regression method, Siegel (1982)) to reduce the effect of such random fluctuations. We approximated the effect of the trials on the normalized cumulative sum by a different number of regression lines, checking the correspondence between the model and the data after each approximation. This approach allows for the assessment of strategy switches.

4. Results

The data imply an improvement in classification performance with sequential blocks. Figure 2 presents the change in the average proportion of correct responses and the median response time for all observers when dividing the samples into eight blocks. The results show an increase in response accuracy and a decrease in the response time for the group of observers. The accuracy is not high – it is only about 75% in the last experimental block. The figure also implies that the compactness of the images affects the learning process.

Figure 2. Left: Mean proportion of correct responses in sequential blocks of experiment. Right: Median of response time in sequential blocks.



The analysis of the effect of the sequential blocks on the accuracy and the response time confirmed this observation. The marginal effect for the average slope is 0.0362 [0.0286, 0.0435] for the accuracy -18.2 [-28.5, -7.7], the values in brackets show the 2.5% and 97.5% confidence intervals. The hypothesis testing shows a tendency for a higher accuracy for the more dispersed patterns than for the more compact ones (Evid. Ratio = 4.32; probability = .81). The results provide strong evidence for a shorter response time for the dispersed patterns than for the more compact ones (Evid. Ratio = 18.25; probability = 0.95).

Figure 3 shows the average proportion of correct responses as a function of the number of red elements ("redness"). The number of red squares describes the data better than the distance from the boundary between the two categories. The dispersed stimuli tend to lead to higher accuracy (Evid. Ratio = 2.83, 74% probability).

Figure 4 shows the effect of the distance from the boundary on the accuracy and the response time. For both characteristics, the classification performance is better (i.e., a higher proportion of correct responses and shorter response times) with the distance to the boundary between the categories. The overall effect of pattern compactness shows an unequal influence of grouping on the performance (Evid. Ratio = 10.36; probability = 0.91 for the accuracy and Evid. Ratio = 15.07; probability = 0.94 - for the response time). The interaction between the element compactness and the distance from the boundary on accuracy suggests a more substantial effect for the grouped than for the dispersed patterns.

Figure 3. Mean proportion of correct responses on the number of red elements.

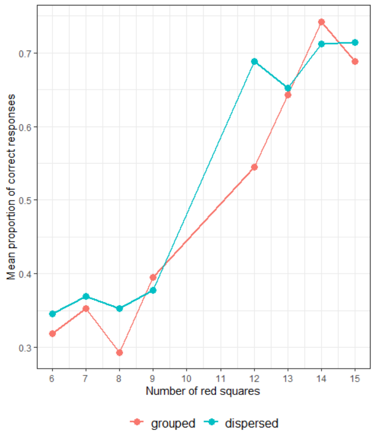
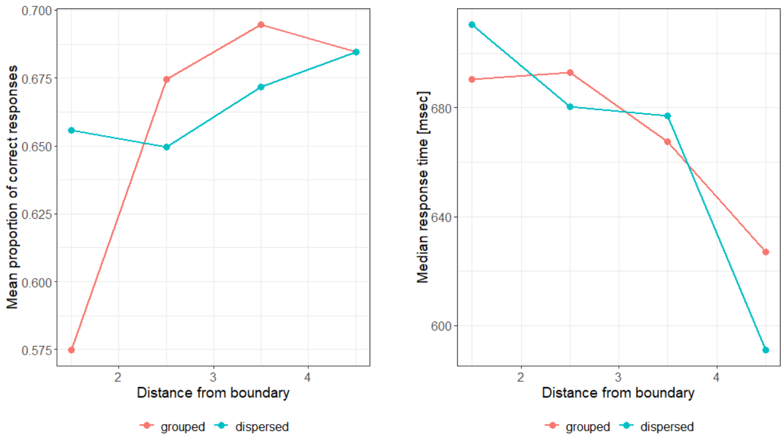
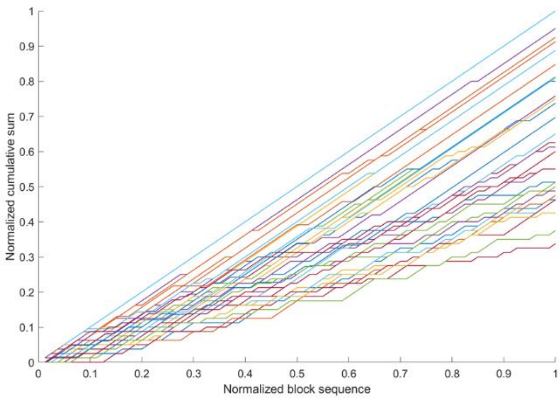


Figure 4. Dependence of the mean proportion of correct responses (left) and the median of response time (right) on the distance from the boundary between the categories.



Looking at the individual data, we obtained that only two observers did not change their strategy; seven changed it once, another seven changed their strategy two times, four participants changed their strategy three times, eight participants changed strategy four times, and seven – five times. Figure 5 represents the obtained dependencies of the cumulative sum of responses from the sequence of presentations. It shows that slightly more than half of the observers succeeded in discovering the classification rule. For them, the slope of the dependency becomes 1.0, although at a different point in the experiment. The figure also shows that the observers who fail to discover the classification rule change their strategy much more often.

Figure 5. Cumulative response sums depending on the sequence of presentations.
A different color represents each observer.



5. Discussion

The study's results implied a significant effect of the compactness of the patterns on the classification performance when the rule for classification is defined by their colorfulness (i.e., the dominance of one of the two colors). The interference of the spatial characteristics on the classification might reflect that the coloring could not exist without a shape or object, i.e., color and shape form unitary wholes. Our results suggest that image classification in distinct categories based on color dominance is easier when the image contains more objects than when its elements are grouped in larger units. A potential explanation for this finding might be that the image objects are predominantly unicolored for the dispersed stimuli, whereas the more grouped elements often contain two colors. The presence of two colors introduces edges in the shape of the objects, destroying their wholeness and affecting the integration of the color information. Edge-based information (shape, lines) is more important than surface-based (e.g., color, texture) information in object recognition and incidental category learning (Zhou et al., 2020) and also in shape bias (Smith, 2000). Another possibility for the deficient learning and categorization performance for the grouped stimuli might be the dominance of shape over color as in natural conditions, shadows or reflections from nearby objects, as well as changes in illumination, often change the color appearance and the visual system might give less weight to color variations in an object.

Our data also show that classifying objects based on color dominance is complex. The overall performance accuracy is not very high, and many participants failed to find the classification rule, even though they were actively seeking it. Whereas we systematically changed only two stimulus dimensions – color dominance and grouping, we inevitably introduced changes in the position of the stimulus elements, their relative disposition, and orientation. Nearly half of observers fail to learn the classification rule, probably due to the natural inclination of observers to integrate the separate stimulus features (Ashby & Gott, 1988) and the dominance of geometric features. The complex characteristics of the stimulation might hinder the relevant features for stimulus representation and categorization. The individual data and the shifts in strategy provide an opportunity to discover the link between the stimulus characteristics used by the participants for category determination. This prospect is planned for a future examination.

Acknowledgments

This study was funded by the Bulgarian National Science Fund, grant No KP-06-N52/6 from 12.11.2021.

References

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 33-53.

Ashby, F. G., & Maddox, W. T. (1994). A response time theory of separability and integrality in speeded classification. *Journal of Mathematical Psychology, 38*(4), 423-466

Bürkner, P. C. (2017). brms: An R package for bayesian multilevel models using Stan. Austria: *R Journal of Statistical Software, 80*(1), 1-28. https://www.jstatsoft.org/article/view/v080i01

Nosofsky, R. M., Meagher, B. J., & Kumar, P. (2022). Contrasting exemplar and prototype models in a natural-science category domain. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(12), 1970-1999. https://doi.org/10.1037/xlm0001069

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of experimental psychology, 77*(3), 353-363. https://doi.org/10.1037/h0025953

R Studio Team. (2020). *RStudio: Integrated Development for R*; RStudio, PBC: Boston, MA, USA, 2020. Retrieved from http://www.rstudio.com

Siegel, A. F. (1982). Robust regression using repeated medians. *Biometrika*, 69(1), 242-244.

Smith, L. B. (2000) Learning How to Learn Words: An Associative Crane. In R. M. Golinkoff, & K. Hirsh-Pasek (Eds), *Becoming a Word Learner: A Debate on Lexical Acquisition, Counterpoints: Cognition, Memory, and Language* (pp. 51-80). New York: Oxford University Press.

Zhou, X., Fu, Q., & Rose, M. (2020). The Role of Edge-Based and Surface-Based Information in Incidental Category Learning: Evidence From Behavior and Event-Related Potentials. *Frontiers in Integrative Neuroscience, 14,* 36. https://doi.org/10.3389/fnint.2020.00036