# Abstract

One popular theory to explain the existence of color categories is to enable efficient inter-person communication via language; reducing the essentially infinite range of perceivable colors down to a subset of categories allows us communicate that information quickly and efficiently to another person via spoken language. Other theories exist. We recently found that a non-linguistic animal (macaques) appear to naturally use color categories. This strikes a blow to the idea that color categories exist *solely* for linguistic inter-personal communication.

# Introduction

Color categorization appears to be a universal human phenomenon. Widespread variability in culture-specific color terminology suggests that color categories are linguistic in nature; however, fundamental similarities in color naming and color categorization across languages suggests that there may be some underlying structure which is universally inherent to human cognition and neurophysiology (Berlin and Kay 1991; Gibson et al. 2017).

Studying categories in non-linguistic animals allows us to pick apart the relative contributions of language and innate factors. It has been previously shown that non-linguistic animals can be trained to perform tasks that involve the categorization of color (Sandell, Gross, and Bornstein 1979; Fagot et al. 2006). What has been unclear is whether these categories are used without explicit externally-motivated training.

We developed a method to test for evidence of categorical behavior without explicit categorical training, based on a task which has been extensively used in the investigation of working memory. In this task, the participant is shown a colored circle on screen, which they remember the color of, and then after a delay, they select a circle of matching color from a set of differently colored circles.

This task has traditionally been used in working memory experiments because color was seen as a simple continuous scale with well-defined perceptual uniformity across the scale. Unfortunately, for working memory researchers, these assumptions have been shown to be ill-founded. (Bae et al. 2015) found that certain colors were remembered more accurately than others and that responses for certain colors were biased towards other colors. The pattern in responses could be accounted for by a model that encoded a memorized color in two distinct ways simultaneously - as a point on a continuous scale and *also* as a member of one of a number of categories that carved up colorspace. For humans they found that a four-category model performed as well as models with higher numbers of categories (though it is possible this was due to a noise floor in their data).

In seeking to explain this result, they showed that these categories correspond relatively well with categories identified in a separate experiment (with separate observers) to the categories recovered from people’s linguistic categories for color. As far as we are aware there is no data that shows correspondence at an individual level between linguistic category definitions and data from an experiment such as this, but the (Bae et al. 2015) data show us that at least at the population level there is an apparent correspondence between the location of color category centers and the biases in response.

(Panichello et al. 2019) extended this line of research in several key ways. Firstly, they extended the logic to account for the fact that when the memory period is extended, participants seemed to draw more heavily on their categorical representation (you can test this yourself - how well can you recall the color of an item you saw fleetingly yesterday? You can probably only confidently report the color category). They did so by casting the category centers as ‘attractor points’ in the space - over time noise would be added to the continuous representation and this noise would be biased by a drift function which would cause the memorized color to drift towards the closest attractor point over time. This provides a computational rationale for why such a mechanism would have value - it places an upper bound on the amount of degradation that can occur as a result of noise acting upon a memory; it can only drift to the nearest attractor point, after which it gets ‘stuck’.

As well as collecting data from a large number of humans on this task, (Panichello et al. 2019) also collected data from 2 macaques on a related task. These animals both showed behavior that was consistent with a model that included attractor dynamics, but the results of each animal didn’t show clear correspondence with the other, or with the human data. It is unclear whether the comparison with the human data is valid, considering a number of differences between the task presented to the humans and the task presented to the macaques.

neural

Here we present the results of a task where 3 macaques performed a task which could also be administered in humans.

[] Summary of things to add

# Methods

## Experimental model and subject details

## Behavioral tasks

#### Stimuli.

Stimuli were discs presented on a XXX screen. The color of stimuli varied only in hue, and were sampled from 64 equally spaced points on a circle in CIELUV space (figure X), with a white point of XXX (xy), a radius of 37, and a luminance of XXX. These values were chosen to maximize gamut while maintaining a fixed saturation and luminance. The background was XXX. CIELUV was used, in contrast to previous work which has used CIELAB, because CIELUV has the benefit of an associated chromaticity diagram. We also noted that nominally equi-saturated stimuli defined in CIELAB tended to have significant variation in apparent saturation, whereas the same in CIELUV were much closer to visually equi-saturated. Luminance noise was added by XXX to the extent of YYY.

The experiment was controlled by multiple computers running ‘Kofiko’ (a MATLAB/Psychtoolbox (ref) based software for working with monkeys).

#### 4-Alternative Forced Choice (4-AFC): Non-human primates.

Non-human primates were trained on a color-matching task. Trials begin with fixation (50 ms) on a white cross in the center of the screen. A cue stimulus (colored disc) is shown to one side of the fixation cross (750 ms). The position of the cue is invariant throughout a daily session. Following cue presentation, the monkey must maintain fixation (600-900 ms) before the choice stimuli appear on the screen alongside the fixation cross (500-1000 ms). The choices are positioned at constant eccentricity and with equal spacing in the hemifield opposite the cue stimulus, with the exact positions of the stimuli varying randomly trial-to-trial. One choice is always a direct match to the cue, and the other three are randomly sampled. Upon offset of the fixation cross, the animal makes a selection by saccade, and is rewarded for selecting the choice that is identical to the cue. Animals were head-fixed at a distance of XXX from the screen. Stimuli had a radius of XXX degrees of visual angle, at an eccentricity of XXX/degrees from a central fixation.

#### 4-Alternative Forced Choice: Human participants.

Human participants were recruited via Amazon Mechanical Turk to perform an analogous version of the non-human primate 4-AFC task. Participants click on an initial fixation cross to request a trial, after which a cue is shown to one side of the fixation cross (750 ms). After cue offset, a fixation cross is shown and the cursor is hidden to de-incentivize mouse movement (1500 ms). Four choices are then shown, and participants make their selection by clicking.

#### Pseudo-continuous color matching task: Human participants.

All 64 stimuli were displayed in a ring at XXX eccentricity.

4-Alternative Forced Choice (4-AFC) Paradigm.

4-Alternative Forced Choice (4-AFC) Paradigm.

#### Color-naming task

[DANNY - get info from Sihan?]

## Data Analysis

To measure color categorization behavior, we analyzed the monkey’s errors in the behavioral task. For each cue, we plotted the monkey’s choices on incorrect trials by the angular error between the cue and the choice. These values were normalized by the number of times each choice was presented as a choice option for all completed trials of the given cue.

The peak of the Gaussian was used as a measure of bias in the monkeys’ reports.

We then fit a Gaussian curve to the error distribution for each cue. This fit was weighted by the number of times each choice color was an option for the given cue across all completed trials. Bias was given in degrees by the difference between the cue (at 0 degrees) and the peak of the Gaussian. Bias was plotted for each cue. Category centers correspond to where the bias curve crosses the zero line on a downward slope.

To find the 95 percent confidence intervals for the locations of the category centers, we performed 1000 bootstraps on all completed trials and found

# Results

Bias as a function of hue, for Conway monkeys

Bias as a function of hue, for Conway monkeys

Bias as a function of hue, for Panichello monkeys

Bias as a function of hue, for Panichello monkeys

# Conclusion

# Data availability statement

Data is available at: zenodo(?)...

# Code availability statement

Code as used in this paper: doi... Latest version of code: github...

# Acknowledgments

# Author contributions

# Competing interests

# References

# Supplementary information

Bae, Gi-Yeul, Maria Olkkonen, Sarah R. Allred, and Jonathan I. Flombaum. 2015. “Why Some Colors Appear More Memorable Than Others: A Model Combining Categories and Particulars in Color Working Memory.” *Journal of Experimental Psychology: General* 144 (4): 744–63. <https://doi.org/10.1037/xge0000076>.

Berlin, Brent, and Paul Kay. 1991. *Basic Color Terms: Their Universality and Evolution*. University of California Press.

Fagot, Joël, Julie Goldstein, Jules Davidoff, and Alan Pickering. 2006. “Cross-Species Differences in Color Categorization.” *Psychonomic Bulletin & Review* 13 (2): 275–80. <https://doi.org/10.3758/BF03193843>.

Gibson, Edward, Richard Futrell, Julian Jara-Ettinger, Kyle Mahowald, Leon Bergen, Sivalogeswaran Ratnasingam, Mitchell Gibson, Steven T. Piantadosi, and Bevil R. Conway. 2017. “Color Naming Across Languages Reflects Color Use.” *Proceedings of the National Academy of Sciences* 114 (40): 10785–90. <https://doi.org/10.1073/pnas.1619666114>.

Panichello, Matthew F., Brian DePasquale, Jonathan W. Pillow, and Timothy J. Buschman. 2019. “Error-Correcting Dynamics in Visual Working Memory.” *Nature Communications* 10 (1): 3366. <https://doi.org/10.1038/s41467-019-11298-3>.

Sandell, Julia H., Charles G. Gross, and Marc H. Bornstein. 1979. “Color Categories in Macaques.” *Journal of Comparative and Physiological Psychology* 93 (4): 626–35. <https://doi.org/10.1037/h0077594>.