Color Categories in Macaque monkeys

**Authors:** Daniel Garside1\*, Audrey Chang1, Hannah Selwyn1, Shriya Awasthi1, Bevil R. Conway1,2\*

**Affiliations:**

1 Laboratory of Sensorimotor Research, National Eye Institute and National Institute of Mental Health; Bethesda MD 20892.

2 National Institute of Mental Health, Bethesda MD 20892

\*Correspondence:. Email: [Daniel.garside@nih.gov](mailto:Daniel.garside@nih.gov); [bevil@nih.gov](mailto:bevil@nih.gov)

Address: Building 49, Room 2A72, 49 Convent Drive, Bethesda, MD-20892-4435

**Abstract**

Previous investigations concerning whether macaques use color categories have produced limited and conflicting results. The extent to which they do exhibit this behavior sets the extent to which neurophysiological results in macaque can be applied as a model for higher level human visual function, and also allows us to disentangle linguistic / cultural / top-down from innate / biological / earthly / terrestrial theories / factors / drivers of / for why we use color categories. Using a delayed forced-choice/match-to-sample paradigm, and building on prior work which established the correspondence between color categories and memory bias, we find that the macaques we studied all show evidence of using two shared categories: warm (orange-ish), and cool (blue-ish). We find some variability between our animals, with one showing evidence for the use of an additional third category. The biases do not appear to align with post-receptoral cardinal mechanisms, though learning rates on the task appear do appear to uncover these mechanisms [to be confirmed].

**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.

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

**Results**

Figure 1 shows…

Figure 2 shows…

Figure 3 shows…

Animals were trained to identify colors in a task (Figure 1A), and in this task they showed errors (Figure 1B)…

In [[fig:BiasCurves]](#fig:BiasCurves) it can be seen that the confidence intervals deviate substantially from the dashed zero line, with the attractor points being found where the black line crosses the zero line from positive to negative (going counter-clockwise) and the repeller points are found where the line crosses from negative to positive.

*Shared color categories across monkeys*

We see that all tested monkeys share two common attractor points, which we interpret as evidence of two shared color categories: a warm/orange-ish category (between 0 and 45), and a cool/blue-ish category (between 180 and 225). In the absence of language, we can infer that these shared categories arise either due to innate biological factors, environmental factors such as the distribution of colors in the terrestrial environment, or a combination of the two. These categories align well with the daylight locus, and also the object/category distinction previously identified.

*Individual differences between monkeys*

In one animal we see evidence of additional categories: strong evidence for a greenish category and weak evidence for a purple category.

*Controls for non-uniformity of colorspace*

We [hopefully will] disambiguate biases that arise from residual non-uniformity in the colorspace, and those that arise from memory biases.

It is plausible that there are non-uniformities in CIELUV that may result in our nominally iso-saturated colors actually appearing to have variable saturation. This would be a concern, as it would be a reasonable prediction that higher saturation colors would be more salient, and thus more likely to be selected as responses. In a control experiment we see no (or very little) bias towards higher saturation colors.

*Learning rates/DKL*

*Limitations of colorspace*

*Comparison with humans*

*Independent verification of found categories*

**Discussion**

**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 ([[fig:StimuliAndParadigm]](#fig:StimuliAndParadigm) ), with a white point of XXX (xy), a radius of 37, and a luminance of XXX (L\* = 76.0693). 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.

*Color-naming task*

**Data Analysis**

*Computation of bias*

To assess the bias in responses for each cue, we computed the distribution of responses on trials where the monkey made an incorrect choice. For each completed trial, we calculated the error as the angular difference between the correct option and the chosen option. For each cue, we computed the number of times the monkey selected each incorrect choice, normalized by the number of times each choice color was available as a choice option for all completed trials of the given cue (this was approximately uniformly distributed).

We then fit a Gaussian with a variable floor () 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 (as before, this was approximately uniformly distributed). Bias was taken as the difference between the cue and the peak of the corresponding Gaussian, for each cue color. These values were then smoothed (with a circular moving average filter of 5 cues) since our primary interest was in the broader structure of the bias distribution, and this is shown as the black lines in [[fig:BiasCurves]](#fig:BiasCurves).

Attractor points (thought to indicate color category centers) occur where the bias curve crosses the zero line from positive to negative (going counter-clockwise). At these points, there is zero bias, and hues on either side provoke choices that are biased inwards towards this point. Correspondingly, repeller points occur where the bias curve crosses the zero line from *negative to positive* (again, going counter-clockwise). At these points there is also zero bias, but hues either side of this point are biased *away* from this point.

*Confidence intervals*

To find the 95% confidence intervals for the locations of the category centers, we performed 1000 bootstraps on all completed trials. For each bootstrapped sample, we found the bias values for each cue color, smoothed the bias curve, and found the category center locations for each bootstrapped dataset. To find the category center locations for the full data set and their confidence interval, we found the category boundaries and segmented all bootstrapped category centers that fell between two consecutive category boundaries. For each of these segments, we found the circular mean and circular standard deviation of all category crossings that fell within these boundaries.

*Learning Rates*

**Data availability statement**

Data is available at: zenodo(?).

**Code availability statement**

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

**References**

# 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>.