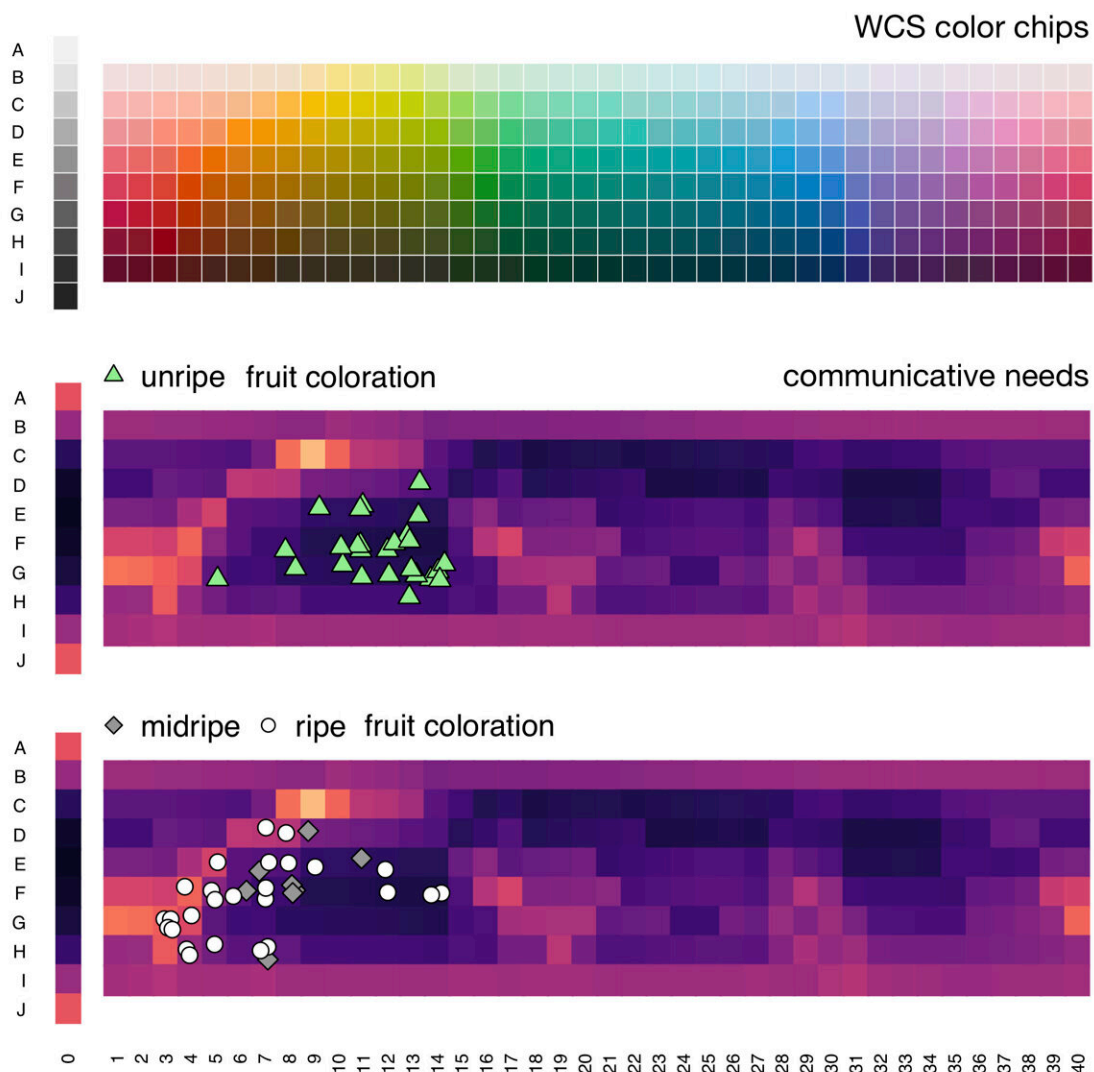


Researchers develop algorithm to map words to colors across languages

September 23 2021, by Katherine Unger Baillie



Across languages, the study found that colors with high communicative need values matched with the colors of ripe fruits in primate diets. Credit: Colin Twomey

No language has words for all the blues of a wind-churned sea or the greens and golds of a wildflower meadow in late summer. Globally, different languages have divvied up the world of color using their own set of labels, from just a few to dozens.

The question of how humans have done this—ascribe a finite vocabulary to the multitude of perceivable colors—has been long studied, and consistent patterns have emerged, even across wildly divergent languages and cultures. Yet slight differences among languages persist, and what is less understood is how the differing communicative needs of local cultures drive those differences. Do some cultures need to talk about certain colors more than others, and how does that shape their language?

In a new study, researchers led by Colin Twomey, a postdoc in Penn's MindCORE program, and Joshua Plotkin, a professor in the School of Arts & Sciences' Biology Department, address these questions, developing an algorithm capable of inferring a culture's communicative needs—the imperative to talk about certain colors—using previously collected data from 130 diverse languages.

Their findings underscore that indeed, cultures across the globe differ in their need to communicate about certain colors. Linking almost all languages, however, is an emphasis on communicating about warm colors—reds and yellows—that are known to draw the human eye and that correspond with the colors of ripe fruits in primate diets.

The work, a collaboration that included Penn linguist Gareth Roberts and

psychologist David Brainard, is published in *Proceedings of the National Academy of Sciences*.

"The fact that color vocabularies could be an efficient representation of the communicative needs of colors is an idea that's been around for 20 years," says Twomey. "It struck me that, OK, if this is our idea about how color vocabularies are formed, then we could go in reverse and ask, 'Well, what would have been the communicative needs that would have been necessary for this vocabulary to arrive at its present form?' It's a hard problem, but I had an intuition that it was a solvable one."

"The color-word problem is a classical one: How do you map the infinitude of colors to a discrete number of words?" says Plotkin. "Colin noticed an evolutionary interpretation of the problem. It's as if the different terms are competing for what colors they will be used to represent. That was a key mathematical insight that allows us to infer the communicative needs of colors in each of these 130 languages."

The study relied on a robust dataset known as the World Color Survey, collected more than 50 years ago by anthropologist Brent Berlin and linguist Paul Kay. Traveling to 130 linguistic communities worldwide, Berlin and Kay presented [native speakers](#) with the same 330 color chips. They found that even completely [different languages](#) tended to group colors in roughly the same way. What's more, when they asked speakers to identify the focal color of a particular named color—the "reddest red" or "greenest green"—speakers' choices were highly similar across languages.

"Their results were so astonishing," Plotkin says. "They demanded explanation."

Substantial research followed, some of which suggested that one major reason for the remarkable similarities between languages' color

vocabularies came down to physiology.

"Languages differ, cultures differ, but our eyes are the same," says Plotkin.

But another reason for the overarching similarities could be that humans, regardless of what language they speak, are more interested in talking about certain colors than others.

The Penn team used data from the World Color Survey on focal colors to work backwards, going from speakers' observations of the reddest red or greenest green to infer the communicative need associated with each of the 330 colors in the survey.

"What was really surprising was that we could use just those best-example colors to say what those communicative needs would have been," says Twomey.

The researchers were able to use the second part of the World Color Survey data, on how languages divided color, to validate that their inference algorithm could predict the communicative needs of different languages.

Their analysis underscores findings from earlier research, that warm-hued colors have a higher communicative need. "On average across languages, the reds and yellows have 30-fold greater demand than other colors," Plotkin says.

"No one really cares about brownish greens, and pastels aren't super well represented in communicative needs," Twomey adds.

The researchers also looked at existing data on fruit-eating primates with color vision systems like our own. These primates tend to eat ripe fruit

with colors that line up almost precisely with the places in the color spectrum with high communicative need. "Fruits are a way for a plant to spread its seeds, hitching a ride with the animals that eat them. Fruit-producing plants likely evolved to stand out to these animals. The relationship with the colors of ripe fruit tells us that communicative needs are likely related to the colors that stand out to us the most," says Twomey. "To be clear, this doesn't say that we have the communicative needs we have because we need to communicate about fruit specifically."

The team's algorithm could predict not only the similarities but also the differences between languages. While an emphasis on reds and yellows was universal, certain languages also had high communicative needs for blues, while greens turned up as important in other languages. The research team found that some of these differences were associated with biogeography and distance. Cultures that shared similar ecoregions were more similar in their communicative needs around colors, perhaps owing to plants or animals in that region that were important for food or other uses.

This approach to the study of communicative needs opens up many other areas for study. "This is something that could be carried to other systems where there is a need to divide up some cognitive space," says Twomey, "whether it's sound, weight, temperature, or something else."

And using this framework also opens up opportunities for pursuing various questions around [color](#) and [language](#).

"Now that we have inferred how often people want to talk about certain colors today, we can take a phylogeny of languages and try to infer what people were talking about 500 or 1,000 years ago. What historical events coincide with changes in our needs to talk about colors?" Plotkin says. "There is tons of work still to be done here."

Such questions will demand unique collaborations like the one undergirded by MindCORE, a campus hub for study of human intelligence and behavior which enabled this work. "Inherently interdisciplinary questions like the ones we tackle in our paper together can be challenging to work on precisely because it takes a team of experts from different fields to answer them," Twomey says. "So I feel very fortunate to have had MindCORE's support here at Penn to assemble exactly the right team for this problem."

More information: What we talk about when we talk about colors, *PNAS* (2021). [DOI: 10.1073/pnas.2109237118](https://doi.org/10.1073/pnas.2109237118) , www.pnas.org/content/118/39/e2109237118

Provided by University of Pennsylvania

Citation: Researchers develop algorithm to map words to colors across languages (2021, September 23) retrieved 20 March 2024 from <https://phys.org/news/2021-09-algorithm-words-languages.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--