Color naming across languages reflects color use

Edward Gibson[a,2], Richard Futrell[a], Julian Jara-Ettinger[a,1], Kyle Mahowald[a], Leon Bergen[a], Sivalogeswaran Ratnasingam[b], Mitchell Gibson[a], Steven T. Piantadosi[c], and Bevil R. Conway[b,2]

[a]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; [b]Laboratory of Sensorimotor Research, National Eye Institute and National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892; and [c]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627

Edited by Dale Purves, Duke University, Durham, NC, and approved August 15, 2017 (received for review November 30, 2016)

What determines how languages categorize colors? We analyzed results of the World Color Survey (WCS) of 110 languages to show that despite gross differences across languages, communication of chromatic chips is always better for warm colors (yellows/reds) than cool colors (blues/greens). We present an analysis of color statistics in a large databank of natural images curated by human observers for salient objects and show that objects tend to have warm rather than cool colors. These results suggest that the cross-linguistic similarity in color-naming efficiency reflects colors of universal usefulness and provide an account of a principle (color use) that governs how color categories come about. We show that potential methodological issues with the WCS do not corrupt information-theoretic analyses, by collecting original data using two extreme versions of the color-naming task, in three groups: the Tsimane', a remote Amazonian hunter-gatherer isolate; Bolivian-Spanish speakers; and English speakers. These data also enabled us to test another prediction of the color-usefulness hypothesis: that differences in color categorization between languages are caused by differences in overall usefulness of color to a culture. In support, we found that color naming among Tsimane' had relatively low communicative efficiency, and the Tsimane' were less likely to use color terms when describing familiar objects. Color-naming among Tsimane' was boosted when naming artificially colored objects compared with natural objects, suggesting that industrialization promotes color usefulness.

color categorization | information theory | color cognition | Whorfian hypothesis | basic color terms

The question of color-naming systems has long been caught in the cross-fire between universality and cultural relativism. Cross-cultural studies of color naming appear to indicate that color categories are universal (1–3). However, the variability in color category boundaries among languages (4), and the lack of consensus of the forces that drive purported universal color categories (5, 6), promotes the idea that color categories are not universal, but shaped by culture (7). Here, we focus on two color categories, WARM and COOL, which are not part of the canonical set of "basic" categories proposed by Berlin and Kay but which nonetheless may be fundamental (8, 9), and which might relate to the basic white/black categories occupying the first stage in the Berlin/Kay hierarchy. Are WARM and COOL universal categories, and if so, why? Here we address these questions by collecting an extensive original dataset in three cultures and leveraging an information-theoretic analysis that has been useful in uncovering the structure of communication systems (10, 11).

The World Color Survey (WCS) provides extensive data on color naming by 110 language groups (www1.icsi.berkeley.edu/wcs/data.html). However, the WCS instructions are complex, and different WCS researchers likely adopted different methods (*SI Appendix*, section 1 and, Figs. S1 and S2), possibly undermining the validity of the WCS data (12, 13). To evaluate this possibility, we obtained extensive color-labeling data using two extreme versions of how the WCS instructions might have been implemented: a "free-choice" paradigm that placed no restrictions on how participants could name colors, and a "fixed-choice" paradigm on separate participants, where participants were constrained to only say the most common terms we obtained in the free-choice paradigm. We conducted experiments in three groups: the Tsimane' people, an indigenous nonindustrialized Amazonian group consisting of about 6,000 people from lowland Bolivia who live by farming, hunting, and foraging for subsistence (14); English speakers in the United States; and Bolivian-Spanish speakers in Bolivia, neighboring the Tsimane'. Tsimane' is one of two languages in an isolate language family (Mosetenan). Although there is trade with local Bolivian towns, most of the Tsimane' participants have limited knowledge of Spanish. To the extent that the communities have organized schooling, education is conducted in Tsimane'.

We analyzed our data using information theory, building on work that suggests that color naming can be better understood by considering informativeness rather than opponent-process theory (3, 11, 15). This analysis can be understood in terms of a communication game (16–18). Imagine that a speaker has a particular color chip c in mind and uses a word w to indicate it. The listener has to correctly guess c, given w. On each trial, the listener guesses that c is among a set of the chips; the listener can pick a set of any size and is told "yes" or "no." The average number of guesses an optimal listener would take to home in on the exact color chip provides a measure of the listener's average surprisal (S, measured in bits; Eq. **1**), a quantitative metric of communication efficiency. The surprisal score for each color c is computed by summing together a score for each word w that might have been used to label c, which is calculated by multiplying $P(w|c)$ by $-\log(P(c|w)$, the listener's surprisal that w

Significance

The number of color terms varies drastically across languages. Yet despite these differences, certain terms (e.g., red) are prevalent, which has been attributed to perceptual salience. This work provides evidence for an alternative hypothesis: The use of color terms depends on communicative needs. Across languages, from the hunter-gatherer Tsimane' people of the Amazon to students in Boston, warm colors are communicated more efficiently than cool colors. This cross-linguistic pattern reflects the color statistics of the world: Objects (what we talk about) are typically warm-colored, and backgrounds are cool-colored. Communicative needs also explain why the number of color terms varies across languages: Cultures vary in how useful color is. Industrialization, which creates objects distinguishable solely based on color, increases color usefulness.

Author contributions: E.G., S.T.P., and B.R.C. designed research; E.G., R.F., J.J.-E., S.R., M.G., and B.R.C. performed research; E.G., R.F., J.J.-E., K.M., L.B., S.R., and B.R.C. analyzed data; and E.G. and B.R.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]Present address: Department of Psychology, Yale University, New Haven, CT 06520.

[2]To whom correspondence may be addressed. Email: bevil@nih.gov or egibson@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619666114/-/DCSupplemental.

would label c. We estimate $P(c|w)$ via Bayes Theorem assuming a uniform prior on $P(c)$.

$$S(c)=\sum_{w} P(w|c)\log\frac{1}{P(c|w)}. \qquad \mathbf{[1]}$$

The novelty of our analysis is to define communicative efficiency about a particular color chip, relative to the set of color chips, so that we have an information-theoretic measure—average surprisal—that allows us to rank the colors for their relative communicative efficiency within a language. To compare color communication efficiency across languages, we follow others (3, 11) in estimating the overall informativeness of the color system of each language by averaging the average surprisal across the chips (Eq. **2**; see *SI Appendix*, section 3 for a worked-out example):

$$\sum_{c} P(c)S(c). \qquad \mathbf{[2]}$$

Results

Color Naming. In our first analysis, we determined the most frequently used (modal) term for each color chip. Identifying modal terms within a language provides an objective way to relate one language to another, but modal terms are not a good measure of the sophistication of a language because they do not account for the variability in term use within a population. Compared with Bolivian-Spanish and English speakers, the Tsimane' speakers showed much greater variability in what color terms were used for all chromatic chips except red. Participants also showed high consensus in labeling black and white chips; we did not solicit responses for other achromatic Munsell chips, although some of the 80 colored chips were labeled using terms glossed as "black" or "white." Diamond sizes in Fig. 1 show the fraction of participants who reported the modal term for the chip at each location in the color array in the free-choice task; diamonds are smaller, on average, in Tsimane' compared with English and Bolivian-Spanish, indicating higher variability in color-term use among the Tsimane' (individual-participant results are shown in *SI Appendix*, Figs. S3–S5; *SI Appendix*, Table S3 shows the key which links the color of the diamonds to the modal color terms). Across the 80 chips, the Tsimane' had eight modal color terms whereas English had 10 and Bolivian-Spanish had 11 (*SI Appendix*, Table S3). Despite differences in variability of term use among the languages, modal term assignments resulted in a generally similar partitioning of the color space across all three languages we studied (similar results were obtained using the fixed-choice version of the task; *SI Appendix*, Fig. S6). These results show that, at the population level, all three languages have a comparable representation of color space. The results for the Tsimane' people mirror observations made in the Hadzane of Tanzania, another indigenous community previously thought to have limited color knowledge, but now known to possess a rich color lexicon distributed across the population (3).

Objects from Memory. We were concerned that the higher variability among the Tsimane' might reflect unfamiliarity with the Munsell cards. To test this possibility, we asked the participants to tell us what color word they associated with familiar objects (a test of memory color). Given each object (O), we determined the uncertainty (H, simple entropy) over color words (W) that were used to refer to that object (Eq. **3**).

$$H(W|O)=-\sum_{i} P(W_i|O)\log(P(W_i|O)). \qquad \mathbf{[3]}$$

Eq. **3** quantifies across people in the culture the consistency of the color label for a given object. Although the objects were selected because each has a characteristic color and is familiar to Tsimane' speakers, Tsimane' on average had higher uncertainty over the color words they associated with the objects (1.06 bits); uncertainty was comparable between English (0.33 bits) and Bolivian-Spanish (0.30 bits) [in paired t tests between the entropy scores per object, the Tsimane'-English comparison is $t(15) = -2.88$, $P = 0.01$; the Tsimane'-Spanish comparison is $t(15) = -3.16$, $P = 0.006$; the English-Spanish comparison is $t(16) = -0.16$, $P = 0.9$]. This difference was driven predominantly by objects whose color names have generally low agreement among Tsimane' as assessed in the color-chip naming experiment (yellow, orange, green, blue, brown), but high agreement among English and Bolivian-Spanish speakers (Fig. 2). Objects such as blood, hair, and teeth whose colors fall in high-consensus color categories in Tsimane' (glossed red, black, and white) were associated with low uncertainty (all Tsimane' speakers have black hair). The relatively large variability in memory colors associated with familiar objects among the Tsimane' corroborates the conclusions drawn from the naming of color chips. Additional control experiments assessing reaction times

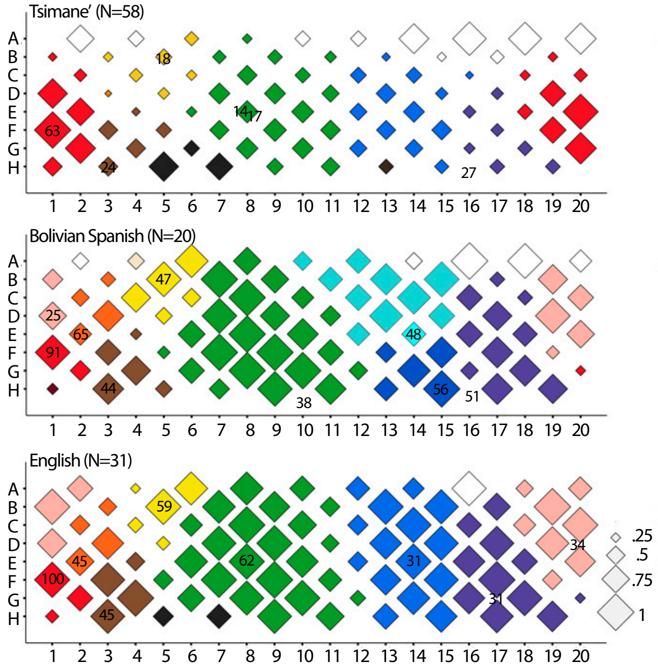

Fig. 1. The Amazonian Tsimane' people show large individual differences in color naming, but at the population level, similar color categories to those observed among Bolivian-Spanish and English speakers. Color naming of 80 chips evenly sampling the Munsell array, presented singly in random sequence under controlled lighting, in Tsimane', Spanish-speakers in neighboring San Borja (Bolivia), and English-speaking students near Boston (see *SI Appendix*, Table S2 for the key relating the axes to Munsell chip designations). Color of each diamond corresponds to the modal color for the chip (see *SI Appendix*, Table S3 for the key matching the color with the terms in each language). Diamond size shows the fraction of people who gave the modal response. All participants showed 100% consistency for black and white chips: negro, blanco (Bolivian-Spanish); tsincus, jaibas (Tsimane'). The location of the numbers overlying the plot indicate the color chips in the 160-chip Munsell array that were most frequently selected as the best example of the subset of modal color terms queried (*SI Appendix*, Table S4). The numbers are the percentage of respondents who made the given selection. Note that two modal color terms in Tsimane', *yụshñus* and *shandyes*, correspond to the same chip (E8). English speakers were asked about red, green, yellow, blue, orange, brown, purple, and pink. Bolivian-Spanish speakers were asked about rojo, verde, amarillo, azul, celeste, anaranjado, morado, cafe, and rosa. Tsimane' were asked about jäinäs (~red), yụshñus (~blue), shandyes (~green), itsidyeisi (~purple), cafedyeisi (~brown), and chames (~yellow). Data are from the free-choice version of the task (n = 58 Tsimane', 20 Bolivian-Spanish, 31 English); data from the fixed-choice version of the task, conducted in separate participants (n = 41 Tsimane', 25 Bolivian-Spanish, 29 English), yielded similar results (*SI Appendix*, Fig. S6).

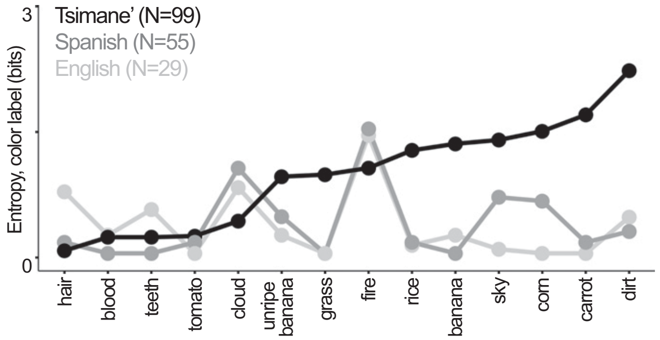

Fig. 2. Variability of color labels (entropy, Eq. **3**) for familiar objects, ordered by Tsimane' results. On average, Tsimane' has higher entropy over color words for a particular object (1.06 bits, compared with English, 0.33 bits, and Bolivian-Spanish, 0.30 bits).

further supports the conclusion that participants were fully engaged in the various tasks (*SI Appendix*, section 2 and Fig. S7).

Similar Results for Free-Choice vs. Fixed-Choice Versions of Color-Naming Task. Returning to the analyses of the color-chip naming tasks, the higher color-naming variability among the Tsimane' speakers, compared with English and Bolivian-Spanish speakers, is reflected in higher average surprisal for all color chips, using either open (*SI Appendix*, section 3 and Fig. S8) or fixed versions of the task. Averaging across color chips (Eq. **2**), the Tsimane' color system (4.88 bits) was shown to be less informative than that of English (3.80 bits) or Bolivian-Spanish (3.86 bits) (free-choice task). Unsurprisingly, the average number of color words produced in each population during the fixed-choice task was lower than in the free-choice version (Fig. 3*A*). Remarkably, the overall informativeness of a language was very similar for the two versions of the task (Fig. 3*A* and *SI Appendix*, section 4 and Figs. S9–S11; the tasks were performed in separate sessions, with different people, about a year apart). Furthermore, Spearman correlations of the rank-ordered sequence of color chips (ordered by increasing average surprisal) for each version of the task were high (Tsimane': $\rho = 0.71$; Bolivian-Spanish: $\rho = 0.90$; English: $\rho = 0.92$). These results come as a relief, allaying widespread concerns about the methodology of the WCS, and licensing further information-theoretic analyses of the WCS data (12, 13).

Focal Color Analyses and Unique Hues. In separate experiments, for several frequently used color terms in each language, subjects were asked to indicate which color chip was the best representative of the color term. Such "focal" colors can be reasonably predicted by statistical models that identify the most representative color chip given each speaker's color-naming data (19). The contours in Fig. S8 show the probability density of color samples chosen for red, green, blue, yellow (English); rojo, verde, azul, amarillo (Bolivian-Spanish); jainas, shandyes, yushñus, chamus (Tsimane') (*SI Appendix*, section 5 and Table S4 provides the chip designations for the most frequently chosen chip for each focal color). The contours tend to cover a broader area of the array for Tsimane' speakers compared with English or Bolivian-Spanish speakers, consistent with the results of the other color-naming experiments showing that the Tsimane' are more variable in the color terms they use (Fig. 1). The contours in *SI Appendix*, Fig. S8 are for colors that correspond to the "unique hues," which have long been postulated to be psychological primary colors (20): The unique hues are considered to be "irreducible" primaries which cannot be described using any more primitive labels (unlike "orange," which some consider yellowish red). Given their purported primary status, one might have hypothesized that these colors would be associated with relatively low average surprisal. Contrary to this prediction, we found that only the red and yellow focal colors had low surprisal across all three languages. The relatively low surprisal of red and yellow, compared with the higher surprisal of blue and green, recalls the smaller individual differences in unique hue settings for red and yellow compared with blue and green (21). These results add to a growing body of research suggesting that the unique hues might not be as special as widely thought (6, 22, 23). Instead, the results suggest that warm colors (reds, yellows) are associated with higher communicative efficiency compared with cool colors (blues, greens).

Analysis of Average Surprisal Values Within Each Language. The pattern of average surprisal values (the variations in gray shown in *SI Appendix*, Fig. S8) was consistent across the three languages. Consistent with this impression, the rank-ordered sequence of color chips was similar across the three languages (Fig. 3*B*; Spearman rank correlation, between Bolivian-Spanish and English $\rho = 0.87$; between Bolivian-Spanish and Tsimane' $\rho = 0.51$; between English and Tsimane' $\rho = 0.53$; Table S5; *SI Appendix*, section 6 and Fig. S12). The ordering forms a striking pattern that is not determined by the unique hues or the focal colors. Warm-colored chips (red, pink, orange, yellow, brown) across Tsimane', English, and Bolivian-Spanish showed relatively low average surprisal, whereas cooler colors (blues, greens) showed higher average surprisal. The rank ordering is also not explained by Berlin and Kay's proposed order of acquisition (1), which has blue and green arising before pink, orange, and brown. Our results suggest that despite overall gross differences in the communication efficiency across languages, among chromatic chips, warm colors are always the easiest to communicate precisely. Remarkably, we found that this relationship was true across the entire WCS of 110 languages (Fig. 4 and *SI Appendix*, section 7 and Fig. S13). These results provide an explanation for the universal distinction between warm and cool colors: Warm colors are always associated with higher communicative efficiency compared with cool colors. Together the results suggest two complementary conclusions: All languages, even those with very few consensus color terms, have a comprehensive

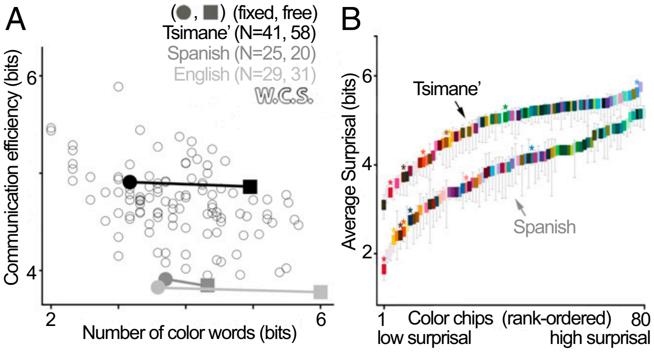

Fig. 3. Communication efficiency of color naming, across languages and among color chips. (*A*) Communication efficiency for each language of the WCS (open symbols), Tsimane' (black symbols), Bolivian-Spanish (dark gray symbols), and English (light gray symbols), as a function of number of unique color words used by the population of participants tested in each language. The two data sets collected in Tsimane', Bolivian-Spanish, and Tsimane' show that variability in experimental methods have little impact on assessments of communicative efficiency of color naming, licensing the use of the WCS data for further analysis. Circles show data from experiments in which participants were constrained to use a fixed vocabulary of basic color terms; squares show data where participants were free to use any term. Number of participants stated as (N=fixed choice, free choice). Communicative efficiency for each language was computed using Eq. **2**. (*B*) Color chips rank-ordered by their average surprisal (computed using Eq. **1**) for Tsimane' and Bolivian-Spanish (pattern for English overlaps Spanish, omitted for clarity). *SI Appendix*, Table S5 provides the chip identity in rank order. The asterisks represent focal colors determined as described in Fig. 2. The sequences of colors in each population are highly correlated (Spearman rank correlation between Bolivian-Spanish and English, $\rho = 0.87$; between Bolivian-Spanish and Tsimane', $\rho = 0.51$; and between English and Tsimane', $\rho = 0.53$).

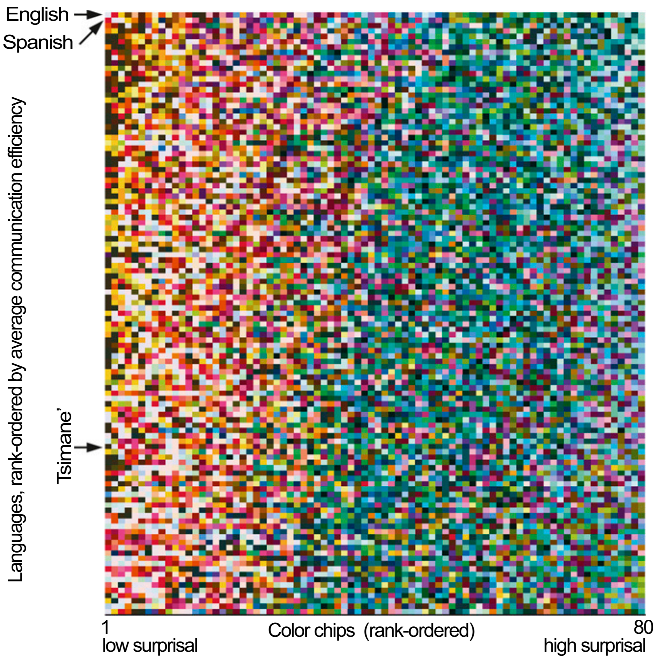

Fig. 4. Color chips rank-ordered by their average surprisal (computed using Eq. **1**) for all languages in the WCS, and the three languages tested here. Each row shows data for a given language, and the languages are ordered according to their overall communication efficiency (Eq. **2**).

color lexicon distributed across the population; however, all languages, even those with a very sophisticated color language, prioritize the same set of (warm) colors.

The colors of objects versus backgrounds. The discovery that warm colors are more precisely communicated compared with cool colors is a finding that emerges from the information-theoretic analysis. However, what determines this universal asymmetry in communicative efficiency between warm and cool colors? We hypothesized that the ordering of chips by average surprisal arises because of the color statistics associated with salient objects, in contrast to their indistinct backgrounds. Natural scenes typically do not show an equal representation of colors. Instead, warm colors (yellow/orange/red) and cool colors (blue/green) are overrepresented (24–26), regardless of season or ecosystem (27), and primary visual cortex is adapted to these statistics (28, 29). Attempts have been made to relate the chromatic statistics of a small sample of natural images to color categories (30), but it is not clear whether objects are representative of natural images in general. To fill this gap in knowledge, we analyzed the colors of objects identified by independent observers in a dataset of 20,000 photographs; this dataset was curated by Microsoft from over 200,000 photographs for the purpose of depicting salient objects (31). We discovered that pixel colors for the objects were more often within the red/yellow/orange ("warm") range, compared with backgrounds, which were typically blue/green ("cool"). Moreover, the likelihood that a color would be found in an object was negatively correlated with its average surprisal in the three languages we studied (Fig. 5 and *SI Appendix*, section 8 and Fig. S14) and the 110 languages of the WCS. These results suggest that what determines the universal patterns across the diversity of languages is the consistent link between warm colors and behaviorally relevant items—salient objects—in the environment. We confirmed these conclusions in an analysis of spectral measurements obtained from objects with and without behavioral relevance to trichromatic primates (32). We found that behaviorally relevant objects (such as fruit eaten by the animals) tended to have colors associated with lower average surprisal (*SI Appendix* and Fig. S15).

These results support the hypothesis that usefulness is the reason why languages acquire a color name. The relatively low communicative efficiency of color naming among Tsimane' suggests that color is simply not that useful for this population. The Tsimane' have little exposure to artificial (industrialized) objects. Industrialization has greatly expanded the gamut and color stability of objects. One idea is that exposure to artificially colored objects promotes the usefulness of color for object identification, which is hypothesized to promote greater precision in color language (33, 34). To test this idea, we performed a contrastive-labeling task (35) (*SI Appendix*, section 9). Eight pairs of objects, familiar to Tsimane' and English, were used in the experiment, four pairs of natural objects and four pairs of artificial objects. Participants were first presented with one object and were asked to name it. Then they were presented with the second object of the same type and asked to name it (Fig. 6*A*). Tsimane' were much less likely to use a color term (Fig. 6*B*). But a mixed-effect logistic regression showed a main effect of object class among the Tsimane': They were more likely to use a color word when naming artificial compared with natural objects ($\beta = 3.59$, $z = 4.00$, $P < 0.0001$), which is consistent with the idea that industrialization promotes color-naming efficiency.

Discussion

The debate on the origins of color categories pits the hypothesis that color-naming systems emerge from universal underlying principles determined innately against the view that culture determines color categories; it is often implied that only one or the other of these theories is correct. Our results favor a reconciliation of these ideas through the the efficient-communication hypothesis (11), which states that categories reflect a tradeoff between informativeness of the terms and their number (10). Cultures across the globe show common patterns in color naming, and even languages with few high consensus color terms appear to have a complete color lexicon distributed across the population, as shown by Lindsey et al. (3) for the Hadza of

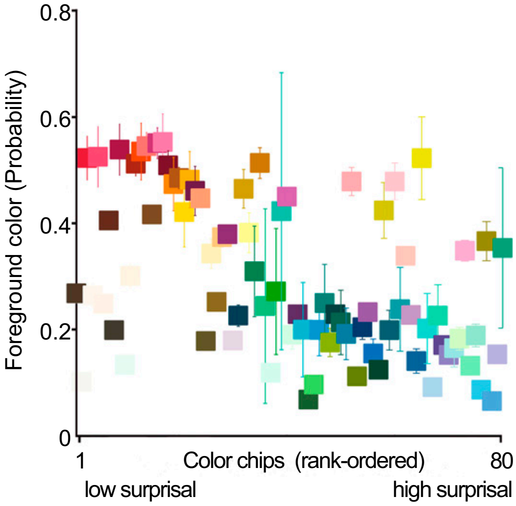

Fig. 5. The color statistics of objects predict the average surprisal of colors. Objects in the Microsoft Research Asia database of 20,000 photographs were identified by human observers who were blind to the purpose of our study (31). The colors of the pixels in the images were binned into the 80 colors defined by the Munsell chips used in the behavioral experiments (across the images there were 9.2×10^8 object pixels and 1.54×10^9 background pixels). The *y* axis shows the [(number of pixels of given color in objects)/(number of pixels of given color in objects + number of pixels of given color in backgrounds)]; the color chip ranking is that obtained for the Tsimane'. Error bars are SE. The three languages were not significantly different from each other (English: slope = −0.0064, $\rho = -0.57$, *P* value = 3×10^{-8}; Bolivian-Spanish: slope = −0.0049, $\rho = -0.44$, *P* value = 5×10^{-4}; Tsimane': slope = −0.0054, $\rho = -0.49$, *P* value = 5×10^{-6}).

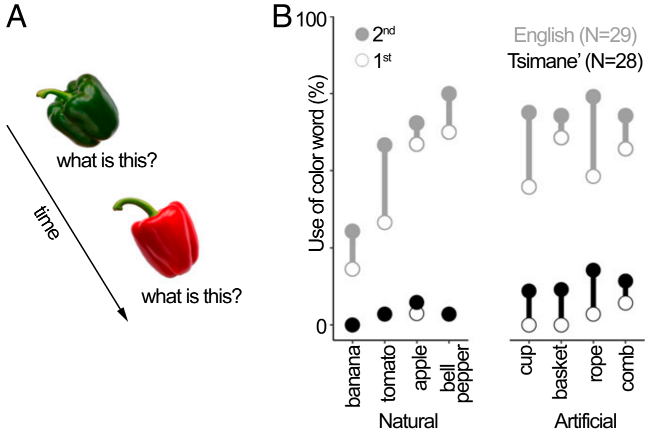

Fig. 6. The Tsimane' use color terms less frequently than English speakers. (*A*) Contrastive-labeling paradigm, adapted to assess use of color terms in normal communication. (*B*) Percent of trials in which participants used a color word to describe objects presented in sequential pairs. Members of each pair were identical except for color (e.g., green banana/yellow banana). Tsimane' speakers were less likely to use a color word (mixed effect logistic regression, $\beta = -5.23$, $z = -5.48$, $P < 0.0001$). Among Tsimane', a mixed-effect logistic regression shows a main effect of artificiality ($\beta = 3.59$, $z = 4.00$, $P < 0.0001$) and presentation order ($\beta = 1.57$, $z = 3.09$, $P < 0.01$) with no interaction ($\beta = 0.91$, $z = 1.19$, $P = 0.23$). Among English, we find a main effect of presentation order ($\beta = 1.53$, $z = 4.00$, $P < 0.001$).

Africa and by us in the Tsimane' of South America. These common patterns across cultures suggest some universal constraint on color naming, but the variability in communicative efficiency about color terms across cultures suggests that culture plays a role too. According to the communication-efficiency hypothesis, if a culture has little need for many high-consensus color categories, it is simpler in that communication system not to have them. We show that all cultures around the world favor communication about warm colors over cool colors, and that this phenomenon reflects a universal feature of natural scenes: Objects defined by human observers tend to be warm colored while backgrounds tend to be cool colored. These results provide evidence that usefulness is the reason for the addition of color terms (36, 37). For example, there simply are not that many natural blue objects, which may explain why many languages acquire the term "blue" relatively late (this left Homer scrambling to come up with an alternative description for the sea: "wine-dark" instead of "blue"; ref. 34). That many if not all "basic" color terms derive, historically, from the names of objects we care about (or cared about) provides yet another clue that usefulness is the principal force that drives color categorization. Consider "orange." Our results suggest that the color statistics of natural objects establish the relative salience of different colors and the informativeness of the associated terms. But we recognize that it is possible that the causal relationship is the inverse: that important natural objects acquired warm coloring to exploit the salience of these colors to trichromatic primates, for example to attract primates to assist in seed dispersal (32).

Although all languages appear to possess a fundamental distinction between warm and cool colors, the large variance of average surprisal values across languages suggests that the average usefulness of color varies among language groups. Our results on object/color associations support this hypothesis by showing that many objects that are common in both Tsimane' and US cultures have a diagnostic color term in English but not in Tsimane'. These results support the idea that color is not as useful for Tsimane' as it is for English and Bolivian-Spanish, consistent with findings in other non-Western groups (36). The Tsimane' have an extensive botanical vocabulary (14), which might obviate the need for color terms in their culture, which is heavily dependent on natural objects. Our results in a contrastive-naming task (Fig. 6) provide direct support for the idea that the predominance of artificially colored objects in Western cultures promotes the usefulness of color and, consequently, increases color-naming efficiency. The number of color terms used by Tsimane' individuals and the efficiency of color-term use increased with more exposure to Bolivian-Spanish (*SI Appendix*, section 10, Fig. S16, and Table S6), suggesting a mechanism for cultural transmission.

The present results confirm, for the Tsimane', prior work showing that language groups with relatively few consensus color categories nonetheless possess a large repertoire of color categories distributed across the population (3). The forces that give rise to the partitioning of color space into color categories more refined than warm/cool remain unclear, but our work promotes the idea the main driving force is the extent to which color categories are behaviorally relevant. Contributions to behavioral relevance may depend on stimulus saturation (38) and reflect efficient partitioning of the irregularities in perceptual color space (2, 15). Relative lightness must also be important in establishing behavioral relevance. Black was communicated among the Tsimane' with high efficiency, which replicates prior work showing that black and white are named reliably across all languages. The efficient communication of black is consistent with our overall hypothesis, that color categories reflect usefulness: Blacks are prevalent in natural images, and retinal processing favors darks over lights (39).

Color processing depends upon an extensive network of brain regions that process retinal signals (40), culminating in the highest levels of processing, in frontal cortex (41). The present report leverages color language as perhaps the best readout of this machinery as it pertains to behavior to uncover the forces behind the most fundamental color categorization, warm versus cool. Finally, we wonder to what extent the fundamental asymmetry in usefulness associated with warm colors versus cool colors underlies their emotional valence (42), as indicated by the warm/cool terminology itself.

Methods

All experimental procedures were approved by Massachusetts Institute of Technology's Committee on the Use of Humans as Experimental Subjects. Informed consent was obtained from all participants, as required by the Committee.

Color-Naming Munsell Chips. Participants were presented with each of 80 colored chips evenly sampling the standard Munsell array of colors (*SI Appendix*, Table S2) in a different random order for each participant under controlled lighting conditions using a light box (nine phosphor broadband D50 color-viewing system, model PDV-e, GTI Graphic Technology, Inc.). Each participant initially took a test of normal color vision (43). All participants that failed this task (~5% of participants) were excluded from further study. The task was performed indoors. For Tsimane' speakers, we assessed their knowledge of Spanish using a short questionnaire of very common words. Most participants did not know all of these words, suggesting a poor knowledge of Spanish for most (see *SI Appendix*, section 1 for more information).

Free-Choice Version. The instructions for this task were as follows: We want to know the words for colors in English/Spanish/Tsimane'. So we want you to tell us the colors of these cards. Tell us what other English/Spanish/Tsimane' speakers would typically call these cards. (Fixed-choice version of the task: There are 11 choices: black, white, red, green, blue, purple, brown, yellow, orange, pink, gray. Choose the closest color word.) See *SI Appendix*, section 1 for the Tsimane' and Spanish translations, with color terms from Spanish/Tsimane' for the fixed-choice version in *SI Appendix*, Table S3. Fifty-eight Tsimane'-speaking adults (mean age: 33.2 y; SD: 12.8 y; range 16–78; 38 females); 20 Spanish-speaking adults (mean age: 29.0 y; 9.1 SD: years; range 18–55; 11 females); and 31 English-speaking adults (mean age: 37.1 y; 11.6 SD: years; range 21–58; 10 females) completed this task. From the complete list of terms used in the population, we determined for each chip the term that was used most often (the modal term). Across the chips, we tallied the set of unique modal terms, and removed from the list any modal terms that were only used for one chip, thus omitting maracayeisi in Tsimane' (a color chip on which jäinäs, glossed red, was a close second), and fuschia and piel ("skin color") in Bolivian Spanish. This set of terms provides an estimate of the basic color terms in the population (*SI Appendix*, Table S3).

Fixed-Choice Version. Forty-one Tsimane' adults (mean age: 38.9 y; SD: 17.6 y; range 18–74; 24 females); 25 Spanish adults (mean age: 25.7 y; 9.1 SD: years;

range 18–55; 13 females); and 29 English adults (mean age: 26.0 y; 8.9 SD: years; range 18–55; 14 females) took part, where participants were given a fixed set of color labels to choose from for each color chip: the modal terms discovered in the free-choice version. We also included black/negro, white/blanco/gray/gris in the English/Spanish tasks, because they are regarded as basic color terms.

Focal Colors. Following the Munsell-chip color naming experiment, each participant ($n = 99$ Tsimane'; 55 Spanish; 29 English) was then presented with a standard 160-chip Munsell array of colors (illuminated by the lightbox) and was asked to point out the best example of several color words ("focal" colors). English speakers were asked about red, green, yellow, blue, orange, brown, purple, and pink. Spanish speakers were asked about rojo, verde, amarillo, azul, celeste, anaranjado, morado, cafe, and rosa. For Tsimane', in the free-choice version of the task, we asked about the colors that the participant produced. For the fixed-choice version, we asked about jäinäs (~red), yushñus (~blue), shandyes (~green), itsidyeisi (~purple), cafedyeisi (~brown), and chames (~yellow). The chips most often selected as focal colors for all of the terms probed are given in *SI Appendix*, Table S4. To show the population results and evaluate the possible privilege of the unique hues, we computed the probability density function for each of the four unique hues over the grid space. The contours in Fig. 2 show the probability that a given color word was used for each color chip on the basis of our empirical data.

Color-Naming Objects from Memory. Following the preceding two tasks, each participant ($n = 99$ Tsimane'; 55 Spanish; 29 English) was read a list of items that have typical colors, and was asked what color each item was in their experience. Each of the items had a conventional color in Tsimane' culture, usually the same as that in North American culture: a cloud (white/gray), dirt (brown), grass (green), hair (black), teeth (white), rice (white), an unripe banana (green), a ripe banana (yellow), the sky (blue), corn (yellow), yucca for eating (white), the outer husk of yucca (brown), blood (red), fire (orange/yellow), a carrot (orange), and a ripe tomato (red).

Spontaneous Use of Color in a Contrastive-Labeling Task. A subset of the participants who participated in the other experiments also took part in a contrastive-labeling experiment; some people took part only in the contrastive-labeling experiment and not in the Munsell-chip color-naming experiment. $n = 28$ Tsimane' adults (mean age: 30.9 y; SD: 17.8 y; range 18–90; 23 females) and 29 English participants (mean age: 35.5 y; SD: 11.0 y; range 21–58; nine females). Eight pairs of objects were obtained for naming, including four pairs of fruits and vegetables: a ripe (yellow) banana and an unripe (green) banana; a ripe (red) tomato and an unripe (green) tomato; a red apple and a green apple; a red bell pepper and a green bell pepper; and four pairs of artifact objects: a red and a yellow cup; a red and a blue comb; a red and a yellow piece of rope; and a red and a green small basket. All of the pairs of objects were identical except for their colors. Our method was an adaptation of the method used by Sedivy (35). Participants were first presented with one object and were asked to name the object. Then the second object of the same type was presented for naming. Each participant named all eight pairs of objects consecutively in this fashion. There were four different random orders of presentation. The experimenter/translator transcribed what was said.

ACKNOWLEDGMENTS. We thank Ricardo Godoy and Tomas Huanca for logistical help; Dino Nate Añez and Salomon Hiza Nate for help translating and running the task; Rashida Khudiyeva and Isabel Rayas for helping with the timing of the reaction times for the color and object naming; Serena Eastman and Theodros Haile who helped segmenting images for the natural-scene analysis; and Evelina Fedorenko, Alexander Rehding, Roger Levy, Caroline Jones, Joshua Green, and Neir Eshel for comments on the manuscript. The work was supported by NIH Grant R01EY023322 (to B.R.C.), NSF IOS Program Award 1353571 (to B.R.C.) and Linguistics Program Award 1534318 (to E.G.), and the Intramural Research Program of the NIH (National Eye Institute/National Institute of Mental Health). Additional support was provided by University of Rochester start-up funding (to S.T.P.) and the Wellesley College Undergraduate Research Summer Program.

1. Berlin B, Kay P (1969) *Basic Color Terms: Their Universality and Evolution* (Univ of California Press, Berkeley, CA).
2. Regier T, Kay P, Khetarpal N (2007) Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci USA* 104:1436–1441.
3. Lindsey DT, Brown AM, Brainard DH, Apicella CL (2015) Hunter-gatherer color naming provides new insight into the evolution of color terms. *Curr Biol* 25:2441–2446.
4. Roberson D, Davidoff J, Davies IR, Shapiro LR (2005) Color categories: Evidence for the cultural relativity hypothesis. *Cognit Psychol* 50:378–411.
5. Valberg A (2001) Unique hues: An old problem for a new generation. *Vision Res* 41:1645–1657.
6. Bohon KS, Hermann KL, Hansen T, Conway BR (2016) Representation of perceptual color space in macaque posterior inferior temporal cortex (the V4 complex). *eNeuro* 3:1–28.
7. Roberson D, Hanley JR (2007) Color vision: Color categories vary with language after all. *Curr Biol* 17:R605–R607.
8. Lindsey DT, Brown AM (2006) Universality of color names. *Proc Natl Acad Sci USA* 103:16608–16613.
9. Holmes KJ, Regier T (2017) Categorical perception beyond the basic level: The case of warm and cool colors. *Cogn Sci* 41:1135–1147.
10. Piantadosi ST, Tily H, Gibson E (2011) Word lengths are optimized for efficient communication. *Proc Natl Acad Sci USA* 108:3526–3529.
11. Regier T, Kemp C, Kay P (2015) Word meanings across languages support efficient communication. *The Handbook of Language Emergence*, eds MacWhinney B, O'Grady W (Wiley-Blackwell, Hoboken, NJ), pp 237–263.
12. Lucy JA (1997) The linguistics of "color". *Color Categories in Thought and Language*, eds Hardin CL, Maffi L (Cambridge Univ Press, Cambridge, UK), pp 320–346.
13. Saunders BA, van Brakel J (1997) Are there nontrivial constraints on colour categorization? *Behav Brain Sci* 20:167–179, discussion 179–228.
14. Leonard WR, et al. (2015) The Tsimane' Amazonian Panel Study (TAPS): Nine years (2002-2010) of annual data available to the public. *Econ Hum Biol* 19:51–61.
15. Jameson K, D'Andrade RG (1997) It's not really red, green, yellow, blue: An inquiry into perceptual color space. *Hardin and Maffi Color Categories in Thought and Language* (Univ of Cambridge Press, Cambridge, UK), pp 295–319.
16. Lantz D, Stefflre V (1964) Language and cognition revisited. *J Abnorm Psychol* 69:472–481.
17. Steels L, Belpaeme T (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behav Brain Sci* 28:469–489, discussion 489–529.
18. Baddeley R, Attewell D (2009) The relationship between language and the environment: Information theory shows why we have only three lightness terms. *Psychol Sci* 20:1100–1107.
19. Abbott JT, Griffiths TL, Regier T (2016) Focal colors across languages are representative members of color categories. *Proc Natl Acad Sci USA* 113:11178–11183.
20. Hering E *Outlines of a Theory of the Light Sense*, trans Hurvich LM, Jameson D (1964) (Harvard Univ Press, Cambridge, MA).
21. Webster MA, Miyahara E, Malkoc G, Raker VE (2000) Variations in normal color vision. II. Unique hues. *J Opt Soc Am A Opt Image Sci Vis* 17:1545–1555.
22. Bosten JM, Lawrance-Owen AJ (2014) No difference in variability of unique hue selections and binary hue selections. *J Opt Soc Am A Opt Image Sci Vis* 31:A357–A364.
23. Wool LE, et al. (2015) Salience of unique hues and implications for color theory. *J Vis* 15:10.
24. Vrhel MJ, Gerson R, Iwan LS (1994) Measurement and analysis of object reflectance spectra. *Color Res Appl* 19:4–9.
25. Nascimento SM, Ferreira FP, Foster DH (2002) Statistics of spatial cone-excitation ratios in natural scenes. *J Opt Soc Am A Opt Image Sci Vis* 19:1484–1490.
26. Webster MA, Mollon JD (1997) Adaptation and the color statistics of natural images. *Vision Res* 37:3283–3298.
27. Webster MA, Mizokami Y, Webster SM (2007) Seasonal variations in the color statistics of natural images. *Network* 18:213–233.
28. Conway BR (2001) Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *J Neurosci* 21:2768–2783.
29. Conway BR (2014) Color signals through dorsal and ventral visual pathways. *Vis Neurosci* 31:197–209.
30. Yendrikhovskij SN (2001) Computing color categories from statistics of natural images. *J Imaging Sci Technol* 45:409–441.
31. Liu T, et al. (2011) Learning to detect a salient object. *IEEE Trans Pattern Anal Mach Intell* 33:353–367.
32. Regan BC, et al. (2001) Fruits, foliage and the evolution of primate colour vision. *Philos Trans R Soc Lond B Biol Sci* 356:229–283.
33. Kay P, Maffi L (1999) Color appearance and the emergence and evolution of basic color lexicons. *Am Anthropol* 101:743–760.
34. Deutscher G (2010) *Through the Language Glass: Why the World Looks Different in Other Languages* (Metropolitan Books, New York).
35. Sedivy JC (2003) Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *J Psycholinguist Res* 32:3–23.
36. Kuschel R, Monberg T (1974) 'We don't talk much about colour here': A study of colour semantics on Bellona Island. *Man (Lond)* 9:213–242.
37. Levinson SC (2000) Yélî Dnye and the theory of basic color terms. *J Linguist Anthropol* 10:3–55.
38. Witzel C (2016) New insights into the evolution of color terms or an effect of saturation? *Iperception* 7:2041669516662040.
39. Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V (2010) Retina is structured to process an excess of darkness in natural scenes. *Proc Natl Acad Sci USA* 107:17368–17373.
40. Lafer-Sousa R, Conway BR (2013) Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex. *Nat Neurosci* 16:1870–1878.
41. Bird CM, Berens SC, Horner AJ, Franklin A (2014) Categorical encoding of color in the brain. *Proc Natl Acad Sci USA* 111:4590–4595.
42. Palmer SE, Schloss KB, Sammartino J (2013) Visual aesthetics and human preference. *Annu Rev Psychol* 64:77–107.
43. Neitz M, Neitz J (2001) A new mass screening test for color-vision deficiencies in children. *Color Res Appl* 26:S239–S249.