Biological origins of color categorization

Alice E. Skelton[a], Gemma Catchpole[a], Joshua T. Abbott[b], Jenny M. Bosten[c], and Anna Franklin[a,1]

[a]The Sussex Color Group, School of Psychology, University of Sussex, Sussex BN1 9RH, United Kingdom; [b]Computational Cognitive Science Laboratory, Department of Psychology, University of California, Berkeley, CA 94720; and [c]School of Psychology, University of Sussex, Sussex BN1 9RH, United Kingdom

The biological basis of the commonality in color lexicons across languages has been hotly debated for decades. Prior evidence that infants categorize color could provide support for the hypothesis that color categorization systems are not purely constructed by communication and culture. Here, we investigate the relationship between infants' categorization of color and the commonality across color lexicons, and the potential biological origin of infant color categories. We systematically mapped infants' categorical recognition memory for hue onto a stimulus array used previously to document the color lexicons of 110 nonindustrialized languages. Following familiarization to a given hue, infants' response to a novel hue indicated that their recognition memory parses the hue continuum into red, yellow, green, blue, and purple categories. Infants' categorical distinctions aligned with common distinctions in color lexicons and are organized around hues that are commonly central to lexical categories across languages. The boundaries between infants' categorical distinctions also aligned, relative to the adaptation point, with the cardinal axes that describe the early stages of color representation in retinogeniculate pathways, indicating that infant color categorization may be partly organized by biological mechanisms of color vision. The findings suggest that color categorization in language and thought is partially biologically constrained and have implications for broader debate on how biology, culture, and communication interact in human cognition.

color lexicons | infant | categorization | color perception | vision

The extent to which cognition is biologically "hardwired" has been hotly debated (1–4). Color cognition has provided a fertile testing ground for such debate. One key question has been the origin of color terms and their categories. Although the spectrum of color is continuous, humans typically refer to colors with a number of discrete terms (e.g., red, green, blue). Some have argued that how terms categorize the continuum of color and how color lexicons evolve is biologically constrained (5, 6); others have argued that color terms and their categories are culturally and linguistically constructed (7). Cognitive scientists from a broad range of disciplines (e.g., linguistics, neuroscience, vision science, anthropology, developmental science) have been working for decades to understand how color terms and their categories form. These efforts have established that although the color lexicons of the world's languages vary in the number of color terms and in how they parse the continuum of color, there is also striking commonality across languages and evidence for "universal" constraints (8–12). For example, in the World Color Survey (WCS), speakers of 110 nonindustrialized languages named 320 colors (13), and analyses have shown that the centers of the categories denoted by these languages' color terms cluster around particular hues (8). These particular hue regions also appear to be central to the color categories of industrialized languages (e.g., English), and are commonly the location of the "focal" best examples of color terms (10). The organizing principles for this common categorization structure have been sought, and computational models have suggested a number of sources, such as the chromatic structure of natural scene statistics (14), chromatic discrimination thresholds (15), or "near-optimal partitioning" mechanisms based on basic principles of categorization (9). There has also been a

hunt for neurons that encode color categorically in regions of the visual cortex and early ventral stream (16–18).

Although we do not yet have solid evidence for the neural basis of commonalities and universal constraints in color naming, further impetus for the idea that color terms and categories have a biological basis has come from studies with infants (19–27). Converging evidence suggests that prelinguistic infants as young as 4-mo-old respond categorically to color. Many of the infant studies have relied on the "novelty preference" method that has been used to demonstrate that categorization is a domain-general and fundamental aspect of infant cognition (28). Infants are familiarized to a given hue through repeated presentation (until the infant's looking at the hue wanes) and a novel hue is then presented during a test phase. If infants look longer at a novel than familiar hue at test (a novelty preference) then infants are deemed to distinguish the two hues in their recognition memory. Studies have shown that infant recognition memory appears to distinguish hues that are differentiated by certain lexical distinctions (e.g., blue–green), and infant recognition memory treats hues within these lexical categories as if they are equivalent (e.g., no novelty preference). This "same-category" equivalence in recognition memory has been found even when hues within a lexical category are well above infants' chromatic discrimination thresholds when measured with simple detection tasks (24, 25), and when hue differences are maximized (24). Infants' responses therefore fit the classic definition of categorization: "responding in an equivalent manner to discriminably different stimuli" (29). Evidence for a categorical response in infant recognition memory for hue has also been provided using neuroimaging methods, such as event-related potentials (22) and near infra-red spectroscopy (27).

Infants' apparent categorization of color suggests that color categorization may have a biological origin. Of course, lexical color categories cannot be completely biologically determined because color lexicons vary across languages, both in the number of color terms and in the location of lexical color boundaries. Communication needs and cultural and environmental forces are

Significance

Humans parse the continuum of color into discrete categories (e.g., "red" and "blue"), and the origin of these categories has been debated for many decades. Here, we provide evidence that infants have color categories for red, yellow, green, blue, and purple. We show that infants' categorical distinctions align strikingly with those that are commonly made in the world's different color lexicons. We also find that infants' categorical distinctions relate to the activities of the two neural subsystems responsible for the early stages of color representation. These findings suggest that color categorization is partly organized and constrained by the biological mechanisms of color vision and not arbitrarily constructed by language.

inevitably valuable in explaining the evolution of a color lexicon within a culture (29). However, it is possible that lexical color categories are partly rooted in the underlying mechanisms of the early visual system that code for color. This partial constraint could potentially explain the commonality in categorization structure across the world's languages, such as the clustering of categories around particular regions of color (8) or the common category motifs that are seen across languages (10).

Although it is theoretically possible that color categories have biological roots, the current evidence of a categorical response to color in infancy is insufficient for a full endorsement of this theory. First, the majority of the evidence for a categorical response to color in infants comes from testing a few color categories that are defined by their lexical distinction in English (e.g., blue–green and blue–purple). The full continuum of hue has not been tested and so categorical distinctions may have been missed. This means that, although there is converging evidence that infants respond categorically to color, the number and location of infant color-category boundaries are not currently known.

A more complete characterization of infant color categorization is needed to clarify the relationship between infants' categorical response and lexical color categories. One possibility is that the way in which infant categories divide up the hue continuum is highly similar to the structure of color lexicons. An alternative possibility is that infants' categorical response is a quirk of a few limited regions of color space, and that it has little resemblance to more comprehensive categorization systems seen in language. Infant color categories could align with those of specific languages, for example, lexicons of industrialized languages that have a greater number of basic terms than lexicons of nonindustrialized languages. However, we consider it more likely that infant color categories would align not with any one language in particular, but rather with the categorization structure that is common across the world's languages, because it is this commonality that potentially suggests some form of biological constraint. For example, one hypothesis is that infant color categories are organized around the hues that Kay and Regier (8) have revealed to be commonly at the centers of the categories of the WCS.

A second reason why there is currently insufficient evidence for the theory that color categorization has biological origins is that the underlying mechanisms of infant color categorization have not been systematically investigated. Infant categories, at least at 4 mo, are unlikely to have communication or cultural origins, but they are not necessarily rooted in the biological mechanisms of the visual system. It is often assumed that the presence of infant categorization is evidence of a biologically determined innateness, yet infants also have a remarkable ability to learn categories by tuning into the statistical regularities present in stimulus exemplars (30). Environmental origins of color categories have been proposed (e.g., ref. 13), and it is at least theoretically possible that infants are able to tune into the statistical regularities and structure of their chromatic environment to extract a categorization structure (22). Another possibility is that infants' color categorization is based on near-optimal partitioning of the color spectrum, as has been argued for lexical categorization (9), with infants also applying basic principles of categorization (31) to an uneven perceptual color space.

When Berlin and Kay (5) first discovered the regularity in the evolution of color terms across color lexicons, it was proposed that color categories had a biological basis. However, subsequent investigation of the retinogeniculate cone-opponent pathways of the visual system that underpin the early encoding of color has revealed that these pathways do not encode the "basic" white, black, red, green, blue, and yellow categories (and the perceptually pure "unique" hues of these categories) as originally proposed (32). What are commonly known as the "red–green" and "blue–yellow" cardinal cone-opponent mechanisms are actually better described as "cherry–teal" and "chartreuse–violet" in terms of the appearance of the colors they encode. The idea that the cardinal mechanisms can explain the commonality in color categorization across languages, such as good examples (focals) and perceptually pure examples (unique hues) of color,

has therefore been largely dismissed. Nevertheless, there has been some recent tentative evidence for a link between the early color mechanisms and color categorization from a study that concluded that the red–green cone-opponent mechanism accounted for the common "warm–cool" category distinction that arises from analysis of WCS naming data (33). It is theoretically possible that infants are able to draw on cone-opponent mechanisms in a similar way to categorize colors. Evidence that infant color categorization is related to the cardinal axes of color vision would be strong support for the theory that color categorization has biological origins.

The current investigation had three aims: first, to establish the hue categories that infants have; second, to establish the relationship between infant color categories and adult color lexicons; and third, to identify the underlying mechanisms of infant color categorization and whether infant categories are related to the cardinal cone-opponent mechanisms that underlie early coding of color. Together, the findings aim to shed light on whether color categorization has biological origins that partially constrain the formation of color categories within a language.

To address these aims, we mapped infant categories onto the hue circle using the stimulus array from the WCS. We systematically sampled colors from a row of the WCS stimulus grid. Within this row, hues span the hue circle, are at constant lightness, and at varying chroma (similar to saturation or colorfulness). We sampled colors at regular hue intervals that we predicted, on the basis of infant chromatic thresholds would be large enough to be discriminated at 4–6 mo (34). We confirmed in an additional experiment that the colors are discriminable at 4–6 mo and that the findings of the current experiment are not related simply to the perceptual similarity of the colors (see *SI Discrimination and Novelty Preference* and Fig. S1). In the current experiment, we used the novelty preference method to look for hue pairs that are distinguished in infant recognition memory, and for hue pairs for which there is no novelty preference despite being discriminable in other contexts. As a second test, we also tested three larger hue pairs that straddled two or more smaller hue pairs to confirm whether or not hues in that region were distinguished in infant recognition memory when chromatic differences were larger. Regions where infants appear to treat hues as if equivalent are identified as being from one hue category, and regions where infants distinguish hues are identified as being categorically distinct.

Systematically mapping infants' hue distinctions, rather than just testing a few English categories as in previous research, avoids a priori assumptions that infant color categories align with lexical categories from any one language. Using the WCS stimulus grid allows direct comparison between the hues distinguished in infant recognition memory and the lexical distinctions and commonalities present in the world's color lexicons, enabling the relationship between infant and lexical color categories to be clarified. For example, infant color categorization is compared with color lexicons from the WCS, to identify any correspondence between the categorical distinctions infants make and those in specific languages. In addition, we conducted an analysis across all 110 languages in the WCS to see whether infant color categories are structured to capture hues that are commonly central to lexical categories. We then investigated the underlying mechanisms of infant color categorization, and plotted the stimuli and infants' responses in a color space that has axes that correspond to the cardinal mechanisms of color vision and represent activation along the retinogeniculate pathways.

Results

Novelty Preference and Infant Color Categories. Analysis of infant looking times during the familiarization phase confirmed that infants familiarized to each hue (*SI Familiarization and Novelty Preference*). Novelty preference scores were calculated as: (time looked at novel color during test phase/total time looked at novel + familiar color during test phase) × 100 (see Fig. S2 and Table S1 for raw data). Novelty preferences are tested using Bayesian analysis (35–38). Bayes factors are reported that calculate the ratio of how probable the data are given one model (e.g., the null) relative to a second model (e.g., the alternative). A *B* of 3 or

greater indicates substantial evidence for the alternative hypothesis (H1) over the null (H0), often equivalent to $P < 0.05$ (36). A B of 0.33 or below indicates substantial evidence for H0 over H1 and a B between 3 and 0.33 indicates data-insensitivity for distinguishing between either hypothesis. Bayesian analysis is more appropriate for our data than null hypothesis significance testing (NHST) as: (*i*) we require our analysis to enable us to make statements on whether the null hypothesis can be accepted (whether infants treat colors equivalently), whereas NHST provides no measure of credibility in favor of the null and nonsignificant results do not enable a definite conclusion (39); (*ii*) we needed a statistical approach that guarantees sensitivity with a minimum number of participants (we have a between-subjects design with 16 conditions) and Bayes combined with an optional stopping rule (test until B is sensitive in either direction) enables this, as B retains its exact meaning as the evidence in favor of H1 over H0 with additional data collection (40); and (*iii*) Bayes factors should not be adjusted for multiple testing (we have 16 tests), as false-alarm rates are dealt with through information in the data with no reference to how many other tests are conducted (41). In addition to Bayes factors, we report associated P values from NHST, although these are for reference only because they are affected by the optional stopping rule and multiple testing. We interpret all effects with respect to Bayes factors only.

Stimuli were drawn from row G of the WCS Munsell array (Munsell value = 4, Y = 12 cd/m^2) in steps of 3 Munsell hues insofar as possible, yielding 14 stimulus hues. The Munsell coordinates of the stimulus hues, preceded by the column number in Fig. 1, are (2, 5R), (3, 7.5R), (6, 5YR), (9, 2.5Y), (12, 10Y), (15, 7.5GY), (18, 5G), (21, 2.5BG), (24, 10BG), (27, 7.5B), (30, 5BP), (33, 2.5P), (36, 10P), and (39, 7.5RP). Starting from column number 3 in Fig. 1, adjacent sampled hues were paired (e.g., 3&6, 6&9, 9&12......39&2) and novelty preferences for each pair were recorded (*Methods*). One-sample t tests were conducted (against 50%) on novelty preference scores for each pair. Here, B refers to a Bayes factor in which the predictions of H1 were modeled as a normal distribution with an SD of 20. All pairs had a sensitive B with 10 participants, apart from three pairs (7.5R–5YR, 7.5GY–5G, and 5PB–2.5P) that required 17, 20, and 12 participants, respectively. Bayes factors revealed support for H1 for four hue pairs: green–yellow: (7.5GY–10Y), $t(9) = 4.19$, $P = 0.002$, $B = 686.05$; blue–purple (5PB–2.5P), $t(11) = 2.98$, $P = 0.001$, $B = 7.43$; blue–green (2.5BG–10BG), $t(9) = 2.81$, $B = 5.05$; red–yellow (7.5R–5YR), $t(16) = 2.67$, $P = 0.02$, $B = 3.51$; and purple–red (10P–7.5RP), $t(9) = 7.31$, $P = 0.001$, $B= 4.9E+09$ (columns 3–6, 12–15, 30–33, 21–24, and 36–39 in Fig. 1*A*). The other nine pairs showed firm support for H0 (all $B < 1/3$, weakest

probability of H0 was for pair 7.5GY–5G, where $t = 1.88$, $P = 0.07$, $B = 0.19$). Infants were also tested on three more widely separated stimulus pairs that spanned several of the original pairs (three larger pairs within green, blue, red–yellow). There was firm support for H1 for the large red–yellow pair (7.5RP–10Y, columns 38–12 in Fig. 1*A*), $t(9) = 2.84$, $P = 0.02$, $B = 6.30E+00$ and for H0 for the larger hue differences within green (7.5GY–2.5BG, columns 24–30 in Fig. 1*A*) or blue (10BG–5PB, columns 15–21 in Fig. 1*A*), (largest $t = 1.34$, smallest $P = 0.21$, largest $B = 0.17$). We identify stimulus 10Y, 2.5Y, and 7.5GY as yellow here because of Munsell hue notation, although at the Munsell value sampled, these hues are darker than prototypical yellow. This issue is returned to in *Discussion*.

Relationship to Lexical Color Categories. Fig. 1*B* gives examples of color-naming systems for the sampled stimulus row for a selection of languages from the WCS that illustrate agreement between infants' response and adult lexical color categories (Wobé, Ivory Coast; Jicaque, Honduras; and Huave, Mexico) with three, five, and seven basic color terms (BCTs), respectively (9). Fig. 1*B* also gives English naming data from Witzel et al. (42) and the green–blue–purple naming motif (GBPm) that was identified from Lindsey and Brown's (12) cluster analysis of WCS naming data, for row G stimuli. Correspondence between the distinctions made in infant color memory and those made by color terms and the color-naming motif can be seen by comparing Fig. 1 *A* and *B*. The five categorical distinctions that infants make align with the location of four of the distinctions made in the English color lexicon, and with lexical distinctions in color lexicons with fewer basic terms than English. For example, Huave is a color lexicon with seven basic color terms and four of five of the categorical distinctions in Huave for row G stimuli are in the same hue region as the distinctions that infants make in their recognition memory. We show correspondence between infant and lexical color categories for a selection of three WCS languages, yet inspection of naming data from the other WCS languages reveals correspondences for many other languages as well. We also find that infants' categorical distinctions align with three of the categorical distinctions in the GBPm (12).

Fig. 1*C* plots the frequency of category centroids in WCS languages for all hues in row G of the WCS stimulus grid (8). The plot shows that the centers of the categories of 110 nonindustrialized color lexicons peak at particular hues and have minima at particular hues (low centroid counts are likely to indicate category boundaries). Infant novelty preferences are indicated underneath the plot by gaps in the solid black horizontal bars. The gaps align qualitatively well with the low points in the bar plot, corresponding to few WCS centroids. Distinctions
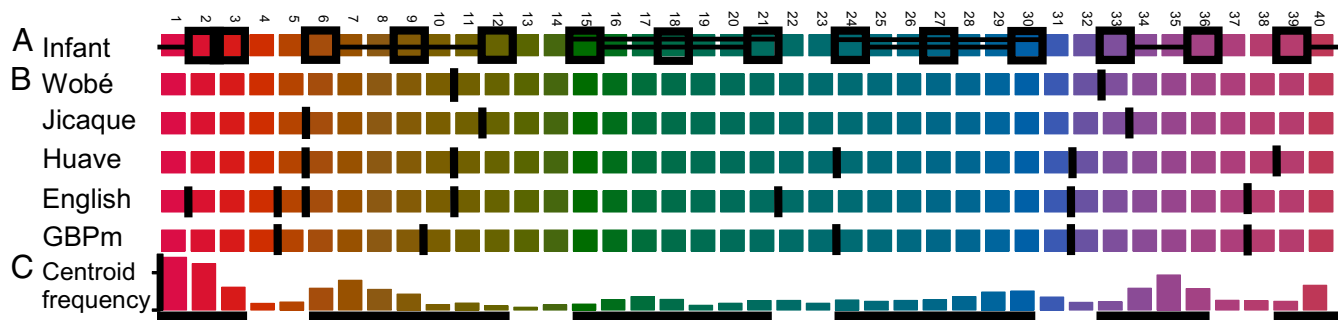

Fig. 1. Infant color categorization and the relationship to lexical color categorization. (*A*) Novelty preferences suggest infant recognition memory distinguishes red–yellow, green–yellow, blue–green, and blue–purple hues but not hues within these categories. Sampled stimuli are outlined in black, horizontal lines joining stimuli indicate color pairs that did not elicit novelty preference, gaps indicate novelty preference. Numbers are from the WCS stimulus grid and indicate the different hues. (*B*) Color naming systems for row G stimuli based on: WCS data for Wobé (3 BCTs), Jicaque (5 BCTs), and Huave (7 BCTs); English color naming (11 BCTs) (42); and the GBPm for row G stimuli (12). Vertical thick black lines indicate category boundaries between stimuli given the same most frequent term within a language. Correspondence between the category boundaries in language and infant novelty preferences can be seen. (*C*) Frequency of category centroids from the WCS for each hue in row G (8). The gaps in the thick black horizontal bars at the bottom of *C* indicate hues which were straddled by color pairs which elicited novelty preference, as also shown by the gaps in the black horizontal lines in *A*. Category centroid frequencies tend to peak in regions which are not distinguished by infant recognition memory and are generally lowest in regions which are distinguished.

infants make between green–yellow, blue–purple, purple–red, and red–yellow hues appear to provide fault lines that separate the centroid peaks from each other (Fig. S3). Infants' distinction between blue–green hues does not fit so well with this pattern, because although it separates the centroid peaks at blue and green, it also spans a region of high centroid counts from WCS languages that have composite blue–green "grue" terms. Analysis of the number of category centroids for hues that were straddled by each pair identified that the particular combination of five pairs for which infants had novelty preference spanned hues with fewer centroids than 4.27% of any other combination of five pairs from the pairs tested (*SI Analysis of WCS Centroids*). This finding suggests that hue pairs that are categorically different for infants are in regions that are infrequently at the center of lexical categories and that infant color categories are optimally organized around hues that are commonly central to lexical color categories: fewer than 5% of other five-pair combinations are better organized.

Underlying Mechanisms of Infants' Response. To test the hypothesis that infant color categorization is related to the cardinal mechanisms of color vision, we plotted—using reflectance spectra taken from the University of Joensuu Color Group database (https://www.uef.fi/en/web/spectral/-spectral-database), the Stockman and Sharpe 2° cone fundamentals (43), and a D65 illuminant—the stimuli and infants' novelty response in a version of the MacLeod–Boynton chromaticity diagram (44). In this color diagram the axes L/(L+M) and S/(L+M) represent the cardinal mechanisms of color vision that correspond to the two main retinogeniculate color pathways. The results are shown in Fig. 2.

A regression analysis found that the Euclidean distances in our version of the MacLeod–Boynton chromaticity diagram did not predict infants' novelty preferences ($R^2 = 0.01$, $P = 0.71$, $B = 2.71E-17$; see also an equivalent analysis in CIELAB color space in *SI Discrimination and Novelty Preference* and Fig. S4). Inspection of stimuli and novelty preferences plotted in the MacLeod–Boynton chromaticity diagram suggests a relationship with the cardinal color mechanisms. Four of the pairs for which there were novelty preferences straddle the vertical and horizontal axes originating from the background chromaticity, Munsell N5, on which our stimuli were presented. This finding indicates a novelty preference in infants when stimulus pairs are at different polarities (relative to the background) either in S/(L+M) or in L/(L+M). Further analyses ruled out the hypothesis that variation in novelty preference across the pairs was driven by a priori preferences (*SI A Priori Preference*).
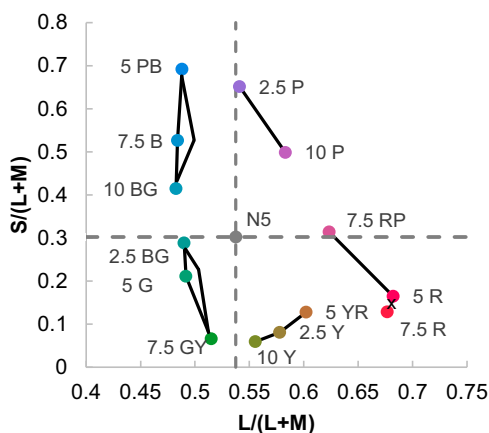

Fig. 2. Stimuli plotted a version of the MacLeod–Boynton chromaticity diagram with L/(L+M) and (S/L+M) cardinal axes of color vision that correspond to the retinogeniculate pathways. The dashed vertical and horizontal lines indicate the background (Munsell N5) to which infants were adapted. The Munsell hue codes for stimuli are given and black lines connecting stimuli indicate no novelty preference for that pair. The cross between 7.5R and 5R indicates a pair that was not tested.

Discussion

Converging evidence from prior research suggests that infants respond categorically to color. However, because the full hue circle has not been previously tested, the number and location of categorical hue distinctions were unknown. Therefore, the relationship between infant color categories and those of the world's color lexicons, and the underlying mechanism of infant color categories have been unclear. To address these issues, we systematically mapped infants' novelty preference responses onto a row of the WCS stimulus grid. Infants successfully familiarized to a given hue, and when presented subsequently with a novel hue, infants had a novelty preference for five pairs of hues in red–yellow, green–yellow, blue–green, blue–purple, and purple–red regions of color space. There was also firm evidence for a lack of novelty response within the hue regions of blue, green, purple, yellow, and red. The lack of a novelty response within purple, blue, and green regions was found even when the largest hue differences spanning pairs that lacked a novelty response were tested.

As noted earlier, we use the term "yellow" to denote stimuli in the yellowish region of the hue continuum following the Munsell hue notation. However, stimuli were sampled at constant lightness close to the lightness of the prototypes of other hues, but darker than prototypical yellow. The dark and nonprototypical nature of the stimuli in the yellow region may account for why there was much more individual variation for the less widely separated red–yellow and green–yellow pairs than most other pairs (more infants had to be tested on these pairs for a sensitive Bayes factor), and the novelty preference effect was weakest for the red–yellow pair. Testing these hues at prototypical lightness levels would confirm this. Nevertheless, whatever the appropriate gloss for the yellowish hues (and this would obviously vary across languages), the present study identifies a categorical distinction for infants in this region.

An additional experiment and further analysis ruled out the hypothesis that novelty preferences were simply related to perceptual similarity and confirmed that colors can be discriminated in other contexts (*SI Discrimination and Novelty Preference*). The findings therefore suggest that infants' recognition memory is categorical: some colors are treated as if they are equivalent in infants' recognition memory, yet others that are not always more perceptually dissimilar are treated as if they are different. Overall, our findings suggest that infant memory parses the hue continuum at the lightness level tested into five categories: red, yellow, green, blue, and purple. Prior research has provided evidence that infants' recognition memory parses color into blue, green (19, 20, 22, 23), and purple (23) categories, with a suggestion of separate red and yellow categories as well (19). Unlike prior studies, we systematically sampled at regular intervals around the complete hue circle, providing evidence that infant memory also distinguishes purple and red.

Relationship to Lexical Color Categories. Mapping infant categories onto the WCS stimulus grid clarifies the relationship between infant color categories and lexical color categorization. Correspondences between the location of infants' categorical distinctions and those in color lexicons can be seen. In particular, the common GBPm that was revealed by Lindsey and Brown's (12) cluster analysis of WCS data has categorical distinctions in four of the same hue regions as those of infants. This correspondence between the distinctions that infants make and those in color-naming systems is also highlighted by our comparison of infant color categories and the locations of the WCS category centroids from Kay and Regier (8). Infants have separate red, yellow, green, blue, and purple categories, and there are also separate clusters of category centroids from WCS languages in these hue regions. Infants' categorical distinctions isolate the peaks in category centroids from one another. An analysis revealed that hues at infant color category boundaries are not commonly at the center of lexical color categories. This finding suggests that hues within infant color categories (hues straddled by pairs where we find no novelty preference) are commonly at the center of lexical color categories. Infants' blue–green categorical distinction did not fit the pattern so well, as it fell in a region where there are a high number of lexical category centers because of the

high incidence of WCS languages with a composite green–blue "grue" term. Nevertheless, that fewer than 5% of other five-pair combinations were better organized around hues that are commonly central to lexical color categories does indicate a striking similarity between prelinguistic and lexical categorization.

Underlying Mechanisms. We found no evidence that infants' novelty preference is driven by how perceptually similar hues are. However, when infants' novelty response was plotted in a color space defined by the cardinal color subsystems that correspond to the retinogeniculate pathways underlying color vision, we see that four of the categorical distinctions that infants make are separated by axes in this color space that pass through the chromaticity of the background, which could be considered the "adaptation point." We propose, therefore, that the null points of the two cardinal subsystems of color vision provide boundaries that infants may use to parse the color continuum into categories in their recognition memory. This of course cannot account for infants' categorical distinction between red and yellow, which must derive from alternative mechanisms.

Our proposal that the cardinal color mechanisms provide fault lines for infant color categorization is related to similar arguments made about adult color naming. For example, one study of adult color naming and hue settings provides evidence that the cardinal axes align with adults' blue–green and yellow–green category boundaries (45). Another study suggests that whether or not L-M cone-contrast between the color and background is positive (reddish) or negative (greenish) aligns with a common categorical distinction in WCS languages between warm and cool colors (33). Similarly, we also find that whether or not colors are "redder" or "greener" than the background [values higher and lower than the background on the L/(L+M) axis] provides a fault line for infants' green–yellow and blue–purple categorical distinctions, but we also find that S/(L+M) provides a fault line for infants' red–purple and blue–green categorical distinctions. A link between low-level mechanisms of color vision and categorization is also implied by a computational simulation, which shows that universal color categorization can be accounted for by human's just noticeable difference function (15; see also ref. 46). It has been suggested that categorical clustering of neurons at V1 can account for the warm–cool categorization in languages (33). However, an association between the cardinal mechanisms and categorization does not necessarily indicate that neurons code categorically at early stages of the visual system. The cardinal mechanisms may simply provide perceptual inequalities, which provide a basis for postperceptual categorization in temporal and frontal regions of the brain (16, 18, 47, 48).

Although we point to an association between the cardinal mechanisms of color vision and infant color categorization here, it is clear that in addition to such biological forces, culture, environment, and communication are likely also to determine both how many color terms there are in a lexicon and the categorical distinctions that are needed. It seems likely that languages would override categorical distinctions that are important in infant color memory if they are not relevant for a given culture or environment (2). One important question for further research is how we transition from a prelinguistic color categorization based largely on biological mechanisms to a lexical color categorization that may make additional or fewer distinctions. The lack of one-to-one mapping between prelinguistic and lexical color categorization may well partly explain the difficulty that children have in learning the words for colors (49–52). Further research should also examine whether the distinctions in infant color memory are universal across different cultures and environments. Given the link to the cardinal mechanisms of color vision, we expect them to be largely consistent across cultures, but cultural variation in prelinguistic categorization is at least theoretically possible.

To conclude, we find that infant recognition memory parses the hue circle into blue, green, purple, yellow, and red categories, and we find similarity in the structure of prelinguistic and lexical color categorization. The retinogeniculate mechanisms of color vision appear to provide fault lines for infants' categorical distinctions between hues. Our findings provide further evidence that color categorization has biological origins. Our results also relate to broader debate on the biological origins of aspects of cognition, such as knowledge about the physical world, mathematical ability, and spatial cognition (2). Although certain cognitive processes may seem to a large extent to be culturally or linguistically constructed because of apparent linguistic and cultural diversity, by looking at commonalities across cultures and languages, and by investigating these cognitive processes in infancy, biological origins of cognitive processes can also be revealed.

Methods

Participants. A total of 295 4- to 6-mo-old infants took part in the study, with 116 infants excluded from the final sample for the following reasons: infant fussiness or lack of looking ($n = 78$); family history of color vision deficiency ($n = 3$); failure to familiarize during the familiarization phase, as defined by a nonnegative slope of looking times across familiarization trials ($n = 17$); failure to look during the test phase ($n = 12$); equipment or experimenter error ($n = 3$); prematurity ($n = 3$). The final sample of 179 infants (89 males) had a mean age of 21.3 wk (SD 2.42). All infants had a birth weight greater than 2,500 g and no known visual or neurological conditions. Ethical approval for this study was obtained from the Sciences and Technology cross schools ethical committee at University of Sussex, and the European Research Council Executive Agency ethics committee. Written informed consent was obtained from the parents of the participants.

Apparatus. Stimuli were presented to infants in two square 12 × 12-cm windows of a wooden booth painted with gray Munsell N5 paint (Y = 19.77 cd/m², x = 0.312, y = 0.325). The two stimulus windows were 3.5 cm to the left and right from central fixation, and the infant was sat in a car seat at eye-level to the horizontal center of the windows at a distance of 50 cm. A 4 × 5-cm digital display was embedded into another window at the central fixation point to centrally fixate infants in between trials, and a circular hole above this (diameter = 2 cm) had a webcam positioned behind it focused on the infants' face. A pulley system around the back of the booth allowed the left and right stimulus windows to either display stimuli during trials, or two squares of Munsell N5 gray and the central digital display during intertrial intervals. The webcam fed into a computer (Dell Precision 390), which recorded the infants' faces in QuickTime. A Matlab program indicated the onset and offset of trials and allowed coding of infant looking. Stimuli were viewed under a D65 illuminant [X-Rite: Judge II (6500K) 24 inch], with a D65 bulb embedded in a hood at the top of the booth and two D65 spot lights angled directly onto the two stimulus windows from behind the infant to ensure the appropriate amount of light was reflected from the stimuli. There was no other light source in the room and the room had black walls and no windows.

Stimuli were sampled from the WCS stimulus array: an array of 320 colors from the Munsell system that vary in Munsell value (lightness) and hue and are at maximum chroma (similar to saturation or colorfulness) for each given stimulus. Stimuli were sampled from row G of the array (Munsell value = 4, Y = 12 cd/m²) in steps of 3 Munsell hue, giving 14 hues in total (5R, 7.5R, 5YR, 2.5Y, 10Y, 7.5GY, 5G, 2.5BG, 10BG, 7.5B, 5PB, 2.5P, 10P, 7.5RP). Stimuli were presented as squares (12 cm for infants) of the reflective Munsell card.

Design and Procedure. Infants were tested with a novelty preference procedure, where there was a familiarization phase with one hue repeatedly shown, and then a test phase where the familiar hue was paired with a novel hue across four trials. The time spent looking at the novel hue relative to the test hue during the test phase was calculated (novelty preference). Each infant saw one hue pair, and there was a minimum of 10 infants per pair (infants randomly allocated), with the hue that was familiar or novel counterbalanced for each pair. Hue pairs were defined first by pairing adjacent stimuli separated by two Munsell hue units. Where there was no novelty preference for two or more adjacent hue pairs then the larger hue difference spanning the adjacent pairs formed another hue pair to be tested. During the familiarization phase the same hue was presented in left and right windows for eight 8-s trials, and in the test phase one familiar and one novel hue were presented for four 5-s trials, with the left/right location of the novel hue counterbalanced and randomized. Intertrial intervals were a minimum of 1 s and the trial then commenced once infants were centrally fixated, with a minimum 2-s interval between the familiarization phase and the test phase. An experimenter sat behind the infant testing booth and worked the pulley system to reveal the stimuli during each trial and to reveal, during the intertrial intervals, the digital display that played a black and white looming and contracting bullseye central attention getter. A second experimenter viewing the webcam output coded infant

looking online, while blind to the condition (colors tested) and location of the novel color. A subset (21%) of the data were coded twice, by an independent experimenter blind to condition and stimulus location, giving an interrater reliability of Pearson's $r = 0.91$.

1. Carey S (2009) *The Origin of Concepts* (Oxford Univ Press, New York).
2. Hespos SJ, Spelke ES (2004) Conceptual precursors to language. *Nature* 430:453–456.
3. Wynn K (1992) Addition and subtraction by human infants. *Nature* 358:749–750.
4. Cohen LB, Chaput HH, Cashon CH (2002) A constructivist model of infant cognition. *Cogn Dev* 17:1323–1343.
5. Berlin B, Kay P (1969) *Basic Colour Terms* (Univ of California Press, Berkeley, CA).
6. Kay P (2005) Color categories are not arbitrary. *Cross-Cultural Res* 39:39–55.
7. Roberson D, Davies I, Davidoff J (2000) Color categories are not universal: Replications and new evidence from a stone-age culture. *J Exp Psychol Gen* 129:369–398.
8. Kay P, Regier T (2003) Resolving the question of color naming universals. *Proc Natl Acad Sci USA* 100:9085–9089.
9. Regier T, Kay P, Khetarpal N (2007) Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci USA* 104:1436–1441.
10. Regier T, Kay P, Cook RS (2005) Focal colors are universal after all. *Proc Natl Acad Sci USA* 102:8386–8391.
11. Lindsey DT, Brown AM (2006) Universality of color names. *Proc Natl Acad Sci USA* 103:16608–16613.
12. Lindsey DT, Brown AM (2009) World Color Survey color naming reveals universal motifs and their within-language diversity. *Proc Natl Acad Sci USA* 106:19785–19790.
13. Kay P, Berlin B, Maffi L, Merrifield W, Cook R (2009) *The World Colour Survey* (CSLI Publications, Stanford).
14. Yendrikhovskij SN (2001) Computing color categories from statistics of natural images. *J Imaging Sci Technol* 45:409–417.
15. Baronchelli A, Gong T, Puglisi A, Loreto V (2010) Modeling the emergence of universality in color naming patterns. *Proc Natl Acad Sci USA* 107:2403–2407.
16. Brouwer GJ, Heeger DJ (2013) Categorical clustering of the neural representation of color. *J Neurosci* 33:15454–15465.
17. Bird CM, Berens SC, Horner AJ, Franklin A (2014) Categorical encoding of color in the brain. *Proc Natl Acad Sci USA* 111:4590–4595.
18. Persichetti AS, Thompson-Schill SL, Butt OH, Brainard DH, Aguirre GK (2015) Functional magnetic resonance imaging adaptation reveals a noncategorical representation of hue in early visual cortex. *J Vis* 15:18.
19. Bornstein MH, Kessen W, Weiskopf S (1976) Color vision and hue categorization in young human infants. *J Exp Psychol Hum Percept Perform* 2:115–129.
20. Catherwood D, Crassini B, Freiberg K (1987) The nature of infant memory for hue. *Br J Dev Psychol* 5:385–394.
21. Catherwood D, Crassini B, Freiberg K (1990) The course of infant memory for hue. *Aust J Psychol* 42:277–285.
22. Clifford A, Franklin A, Davies IRL, Holmes A (2009) Electrophysiological markers of categorical perception of color in 7-month old infants. *Brain Cogn* 71:165–172.
23. Franklin A, Davies IRLL (2004) New evidence for infant colour categories. *Br J Dev Psychol* 22:349–377.
24. Franklin A, et al. (2008) Categorical perception of color is lateralized to the right hemisphere in infants, but to the left hemisphere in adults. *Proc Natl Acad Sci USA* 105:3221–3225.
25. Franklin A, Pilling M, Davies I (2005) The nature of infant color categorization: Evidence from eye movements on a target detection task. *J Exp Child Psychol* 91:227–248.
26. Ozturk O, Shayan S, Liszkowski U, Majid A (2013) Language is not necessary for color categories. *Dev Sci* 16:111–115.
27. Yang J, Kanazawa S, Yamaguchi MK, Kuriki I (2016) Cortical response to categorical color perception in infants investigated by near-infrared spectroscopy. *Proc Natl Acad Sci USA* 113:2370–2375.
28. Mareschal D, Quinn PC (2001) Categorization in infancy. *Trends Cogn Sci* 5:443–450.
29. Quinn P (2003) Concepts are not just for objects: Categorisation of spatial relation information by infants. *Early Category and Concept Development: Making Sense of the Blooming, Buzzing Confusion*, eds Rakison DH, Oakes LM (Oxford Univ Press, New York), pp 50–76.
30. Fiser J, Aslin RN (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn* 28:458–467.
31. Garner WR (1974) *The Processing of Information and Structure* (L. Erlbaum Associates, New York).
32. Jameson K, D'andrade RG (1997) It's not really red, green, yellow, blue: An inquiry into perceptual color space. *Color Categories in Thought and Language*, eds Hardin J, Maffi L (Cambridge Univ Press, Cambridge, UK), p 295.
33. Xiao Y, Kavanau C, Bertin L, Kaplan E (2011) The biological basis of a universal constraint on color naming: Cone contrasts and the two-way categorization of colors. *PLoS One* 6:e24994.
34. Knoblauch K, Vital-Durand F, Barbur JL (2001) Variation of chromatic sensitivity across the life span. *Vision Res* 41:23–36.
35. Jeffreys H (1961) *Theory of Probability* (Oxford Univ Press, Oxford), 3rd Ed.
36. Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773–795.
37. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16:225–237.
38. Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HL (2011) Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *J Pers Soc Psychol* 100:426–432.
39. Dienes Z (2014) Using Bayes to get the most out of non-significant results. *Front Psychol* 5:781.
40. Rouder JN (2014) Optional stopping: No problem for Bayesians. *Psychon Bull Rev* 21:301–308.
41. Dienes Z (2016) How Bayes factors change scientific practice. *J Math Psychol* 72:78–89.
42. Witzel C, Sanchez-Walker E, Franklin A (2013) The development of categorical colour constancy. *Perception ECVP Abstract* 42:19.
43. Stockman A, Sharpe LT (2000) The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Res* 40:1711–1737.
44. MacLeod DIA, Boynton RM (1979) Chromaticity diagram showing cone excitation by stimuli of equal luminance. *J Opt Soc Am* 69:1183–1186.
45. Malkoc G, Kay P, Webster MA (2005) Variations in normal color vision. IV. Binary hues and hue scaling. *J Opt Soc Am A Opt Image Sci Vis* 22:2154–2168.
46. Mullen KT, Kulikowski JJ (1990) Wavelength discrimination at detection threshold. *J Opt Soc Am A* 7:733–742.
47. He X, Witzel C, Forder L, Clifford A, Franklin A (2014) Color categories only affect post-perceptual processes when same- and different-category colors are equally discriminable. *J Opt Soc Am A Opt Image Sci Vis* 31:A322–A331.
48. Koida K, Komatsu H (2007) Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nat Neurosci* 10:108–116.
49. Wagner K, Dobkins K, Barner D (2013) Slow mapping: Color word learning as a gradual inductive process. *Cognition* 127:307–317.
50. Franklin A (2006) Constraints on children's color term acquisition. *J Exp Child Psychol* 94:322–327.
51. Johnson EK, McQueen JM, Huettig F (2011) Toddlers' language-mediated visual search: They need not have the words for it. *Q J Exp Psychol (Hove)* 64:1672–1682.
52. Pitchford NJ, Mullen KT (2003) The development of conceptual colour categories in pre-school children: Influence of perceptual categorization. *Vis Cogn* 10:51–77.
53. Wyszecki G, Stiles W (2000) *Color Science: Concepts and Methods, Quantitative Data and Formulae* (Wiley and Sons, New York), 2nd Ed.
54. Pereverzeva M, Hui-Lin Chien S, Palmer J, Teller DY (2002) Infant photometry: Are mean adult isoluminance values a sufficient approximation to individual infant values? *Vision Res* 42:1639–1649.
55. Jones PR, Kalwarowsky S, Atkinson J, Braddick OJ, Nardini M (2014) Automated measurement of resolution acuity in infants using remote eye-tracking. *Invest Ophthalmol Vis Sci* 55:8102–8110.
56. Franklin A, et al. (2008) Salience of primary and secondary colours in infancy. *Br J Dev Psychol* 26:471–483.
57. Franklin A, Bevis L, Ling Y, Hurlbert A (2010) Biological components of colour preference in infancy. *Dev Sci* 13:346–354.
58. Zemach IK, Teller DY (2007) Infant color vision: Infants' spontaneous color preferences are well behaved. *Vision Res* 47:1362–1367.
59. Brown AM, Lindsey DT (2013) Infant color vision and color preferences: A tribute to Davida Teller. *Vis Neurosci* 30:243–250.
60. Franklin A (2009) Pre-linguistic categorical perception of colour cannot be explained by colour preference: response to Roberson and Hanley. *Trends Cogn Sci* 13:501–502.
61. Bornstein MH (1975) Qualities of color vision in infancy. *J Exp Child Psychol* 19:401–419.