Shared understanding of color among congenitally blind and sighted adults

Authors: Judy Sein Kim[1]*, Brianna Aheimer[1], Verónica Montané Manrara[1], Marina Bedny

Affiliations: [1]Department of Psychological and Brain Sciences, Johns Hopkins University

***Correspondence to**: judyseinkim@gmail.com

Abstract

Empiricist philosophers such as Locke famously argued that people born blind could only acquire shallow, fragmented facts about color. Contrary to this intuition, we report that blind and sighted people share an in-depth understanding of color, despite disagreeing about arbitrary color facts. Relative to the sighted, blind individuals are less likely to generate 'yellow' for banana and 'red' for stop-sign. However, blind and sighted adults are equally likely to infer that two bananas (natural kinds) and two stop-signs (artifacts with functional colors) are more likely to have the same color than two cars (artifacts with non-functional colors), make similar inferences about novel objects' colors, and provide similar causal explanations. We argue that people develop inferentially-rich and intuitive "theories" of color regardless of visual experience. Linguistic communication is more effective at aligning people's theories than their knowledge of verbal facts.

Introduction

Humans acquire knowledge through a variety of means—through direct sensory experience, communication with others, and by thinking (1,2). The question of where knowledge comes from is at the heart of many cognitive theories, but disentangling the relative contributions of these sources is challenging (e.g., 3-8). Color knowledge in blindness provides a wedge into this puzzle by isolating the contribution of first-person sensory experience (9-11). British Empiricists like Locke and Hume argued that color ideas are inaccessible to people born blind since sensory experience is the foundation of knowledge. More recently, Frank Jackson's influential thought experiment featuring Mary, a color scientist who lives in a black-and-white room, inspired many debates about knowledge gained from first-person experience (e.g., 11-15). Only a handful of empirical studies have examined what people born blind actually know about color. Landau & Gleitman (16) showed that Kelli, a congenitally blind 4-year-old, applied color words to concrete objects but not mental entities (e.g., ideas) and understood that color could only be perceived visually, unlike texture or size. Blind and sighted adults share knowledge of similarities between colors (e.g. green and blue are similar but different from orange and red) (17-19). Several studies have also identified differences in blind and sighted people's color knowledge. A recent study found that unlike sighted adults, adults born blind are unlikely to agree on the colors of common animals (20). Even for objects for which blind and sighted people generate the same color labels, blind individuals are less likely to use color as a dimension during semantic similarity judgments (21). One interpretation of this evidence is that color knowledge acquired through verbal communication consists largely of associative verbal facts and is thus inferentially shallow (21-23). An individual born blind might know that the word 'yellow' is used to describe 'bananas' without having in-depth understanding of color.

An alternative possibility is that blind individuals share with those who are sighted inferentially-rich understanding of color. In fact, it is possible that in the absence of first-person sensory experience of color, inferentially-rich knowledge is more preserved than knowledge of associative color facts. Sighted children's inferentially-rich, causal understanding of color enables them to make inferences about real and novel object colors. Children expect an object's relationship with color to differ depending on whether it is a natural kind (e.g. animal, plant, gem) or an artifact (e.g. machine, tool). When asked, "Could something still be a Glick even if it was a different color?" 5 year-old children are more likely to say yes for an artifact than for an animal (24). Such intuitions about color are part of broader frameworks, often referred to as 'intuitive theories', about physical objects (e.g., 25, 26). Children view natural kinds as having intrinsic essences emanating from nature, but link properties of artifacts to human intentions (e.g., 27-29). In response to "Why is this object yellow?" children prefer explanations that appeal to biological mechanisms for natural kinds but human intentions for artifacts (30). Such knowledge about color can be distinguished from that of other object properties: for example, while different instances of a natural kind (e.g., bananas) are more likely to have the same color than instances of an artifact (e.g., cars), natural kinds and artifacts are similarly likely to have consistent shapes (24). The role of first-person sensory experience in acquiring such causal-explanatory and inferentially-rich color knowledge is not known.

In the current study, we probed sighted and congenitally blind people's associative and causal-explanatory knowledge of color. Experiment 1 first queried associative memory for real objects' colors by asking participants to generate "a common color of X" (Figure 1). We next asked participants to judge how likely two instances of the same object are to have the same color, for natural kinds (e.g. two bananas) and artifacts (e.g. two cars). We reasoned that if people share intuitive theories about the relationship between color and object kind, blind and sighted people would make similar inferences about color consistency, even while disagreeing on associative facts (i.e., the particular colors of objects). For example, blind individuals might share with the sighted the intuition that polar bears but not cars have a consistent color, despite showing low agreement that polar bears are white. Alternatively, the experience of seeing that most polar bears are white but that cars come in many different colors might be required to learn the different importance of color for natural kinds and artifacts.

To test nuanced intuitions about the causal mechanisms that lead artifacts to have the colors they do, we included a third object category: artifacts with functional colors (e.g. stop-sign, paper, coin). Artifacts vary according to how much and in what way color relates to their function. For some (e.g., mugs), color is not related to function (holding liquid), while for others (e.g., stop signs) it is integral (e.g. stop signs are consistently red for visibility and recognizability). If sighted and blind people appreciate how color relates to artifact function, they should judge stop signs to have more consistent colors than mugs.

The ability to support generalization to novel instances is a key test of whether knowledge is inferentially-rich (e.g., 28, 31). In Experiment 2, we thus asked participants to make inferences about color consistency for novel objects (natural kinds, artifacts with function-relevant color, and artifacts with function-irrelevant color) in an imaginary island scenario (Fig. 1). If abstract knowledge about the origins and causes of color is shared, then blind and sighted participants should be able to make systematic judgments about color consistency on the basis of object category (e.g., kind of fruit, creature, gem, or household item, gadget, coin) alone. Finally, in Experiment 3, we elicited open-ended explanations for why objects have their colors (e.g., “Why is a carrot orange?”). This allowed us to probe the specific nature of blind and sighted people’s knowledge of the causal mechanisms that give rise to object colors.

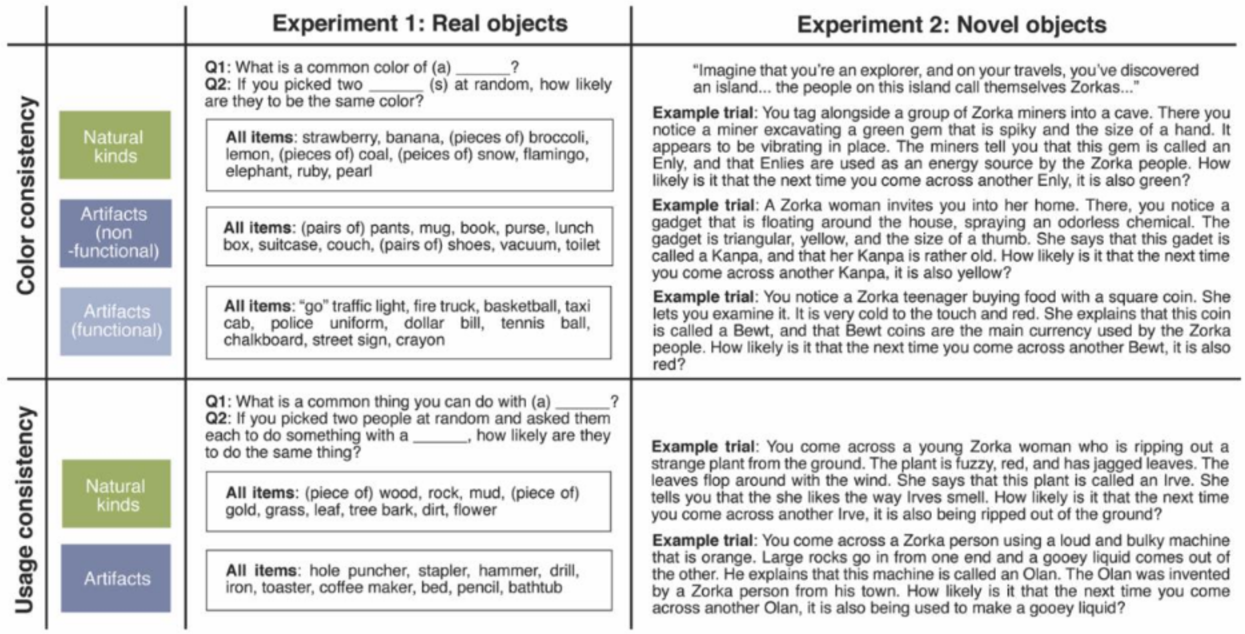

Fig1. Experimental conditions and trials for color consistency inference. Participants were asked about color and usage consistency for real (Experiment 1) and novel (Experiment 2) objects. In both experiments, color trials asked about natural kinds, artifacts with non-functional colors, and artifacts with functional colors, while usage trials asked about natural kinds and artifacts. Different items were used in every trial. For Experiment 1, all items used are listed, and for Experiment 2, one sample trial (an Appendix with full list of trials can be found in Supplementary Materials).

Results

Knowledge of specific object colors among sighted and blind participants

Blind and sighted participants were asked to name a common color of 54 real objects (Experiment 1, 30; Experiment 3, 24, collapsed for the current analysis) (Figure 2A). Objects were chosen from three larger types: natural kinds (NK) (e.g. lemons), artifacts with non-functional colors (A-NFC) (e.g. cars) and artifacts with functional colors (A-FC) (e.g. stop-signs). Color naming agreement was quantified for each object using Simpson's Diversity Index (SDI) (32), and log-transformed SDIs were modeled using linear mixed effects regression (see Methods for details). For both sighted and blind groups, color naming agreement was higher for natural kinds (e.g. lemon) than for artifacts with non-functional colors (e.g. car), but similar to artifacts with functional colors (e.g. stop signs) (Figure 2B; SDI for sighted NK: M=0.86, SD=0.2; A-NFC: M=0.49, SD=0.29; A-FC: M=0.74, SD=0.29; blind NK: M=0.5, SD=0.25; A-NFC: M=0.24, SD=0.09; A-FC: M=0.48, SD=0.26). Naming agreement was substantially higher for sighted compared to blind participants across all object types, and there was no group-by-object kind interaction (result of regression, effect of group: $\chi^2(1) = 71.11$, p<0.0001, $\omega_p^2 = 0.57$; effect of object type: $\chi^2(2) = 20$, p<0.0001, $\omega_p^2 = 0.25$; object-type-by-group interaction: $\chi^2(2) = 1.49$, p=0.5).

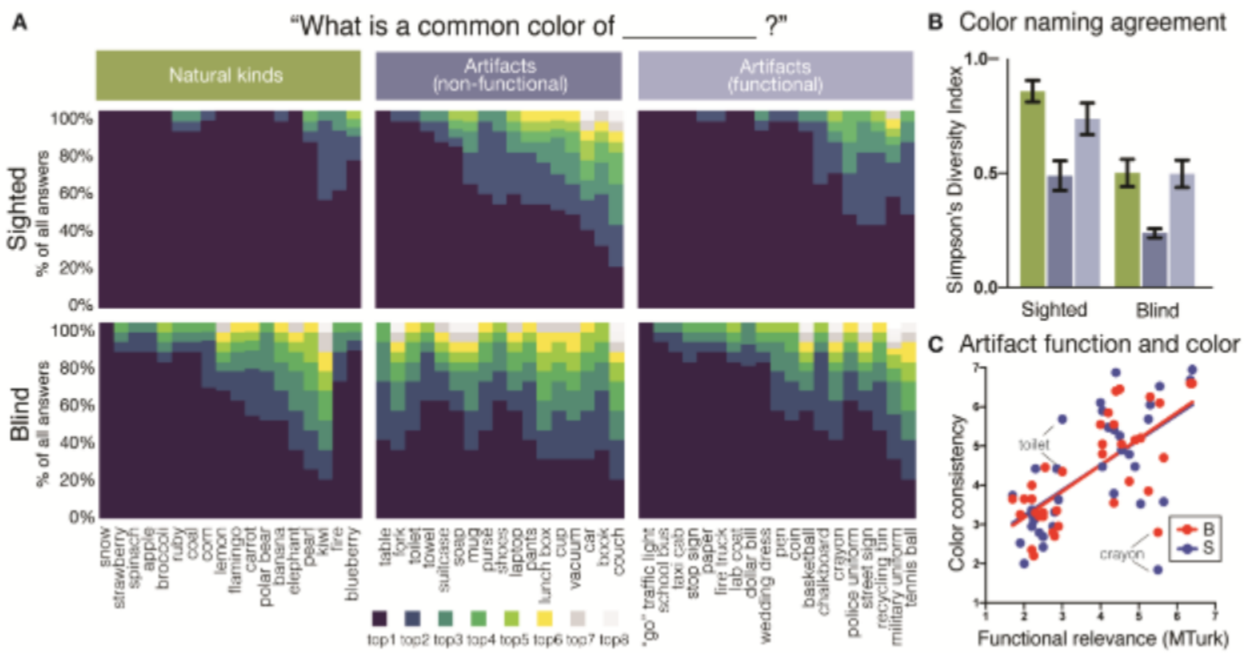

Fig 2. Object color naming agreement. Blind and sighted participants were asked to name common colors of real objects (Experiments 1 and 3). (A) Stacked bars show the frequency of the 8 most frequent colors provided for each object. Frequency for each unique color word is shown as a proportion of all words provided for an object. (B) Bar graph showing naming agreement (Simpson's Diversity Index calculated for individual objects). Mean +/- SEM (across objects). (C) Correlation of ratings from Amazon Mechanical Turk participants (n=20) for artifact colors' relevance to function with blind and sighted participants' color consistency judgments.

Color consistency inferences in blind and sighted individuals: real objects

Sighted and blind participants judged the likelihood that two objects (e.g. two lemons), randomly chosen from the same object category, would have the same color for 10 natural kinds (NK e.g. lemon), 10 artifacts with non-functional colors (A-NFC e.g. car) and 10 artifacts with functional colors (A-FC e.g. stop sign) (henceforth color consistency judgment). Participants rated consistency likelihood on a scale of 1 to 7 (1: not likely, 7: very likely). As a control, participants judged the likelihood that two people chosen at random would do the same thing with an object (e.g. a leaf vs. a car) (henceforth usage consistency judgment). Usage consistency was tested for 10 natural kinds (NK) and 10 artifacts.

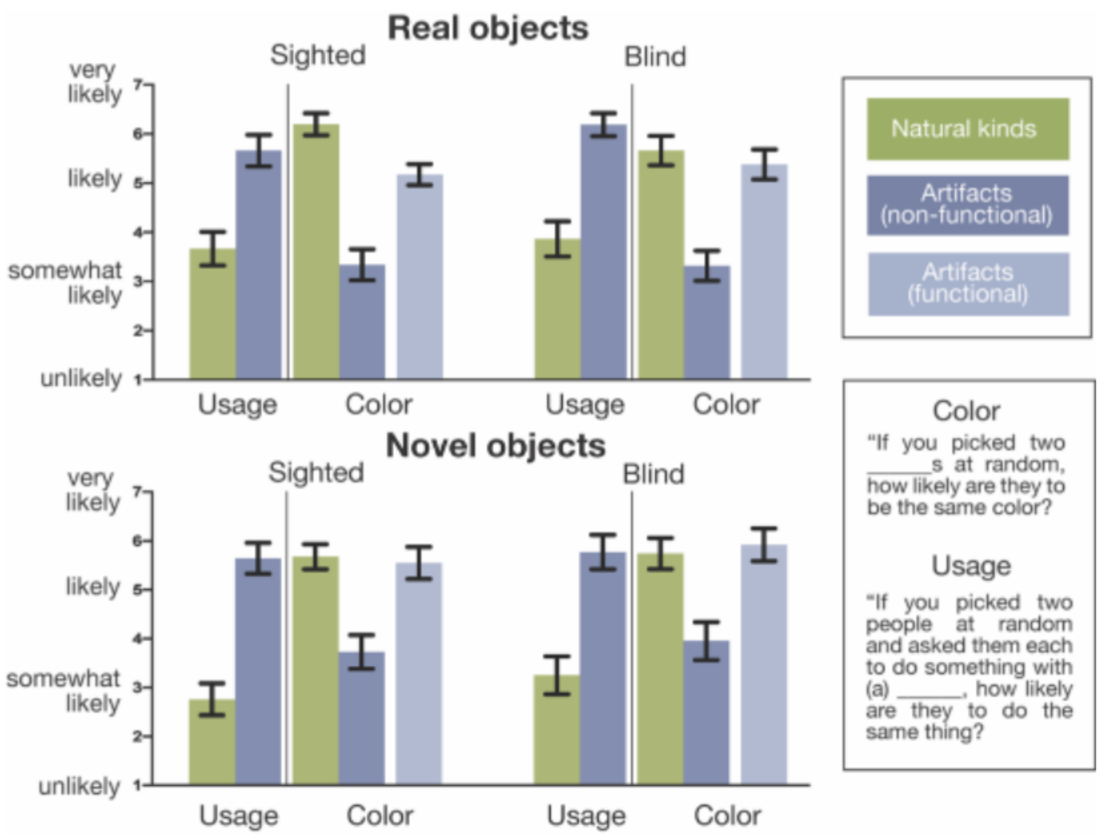

Fig 3. Inferences about color and usage consistency across instances of an object. Consistency judgments for real (Experiment 1) and novel (Experiment 2) objects. Bars are mean +/- SEM.

Sighted participants judged natural kinds (e.g. lemons) to have lower usage consistency but higher color consistency, relative to artifacts (with non-functional colors, e.g. cars) (Fig. 3; sighted usage NK: M=3.67, SD=1.59; usage A: M=5.66, SD=1.49, Wilcoxon matched-pairs signed rank test for usage NK vs. A, two-tailed: z=-3.82, p=0.0001, r=0.88; color NK: M=6.2, SD=1.12, color A-NFC: M=3.34; SD=1.72; color NK vs. A-NFC, z=3.78, p=0.0002, r=0.87). Sighted participants' color consistency ratings for artifacts with functional colors (e.g. stop-signs) were higher than those for artifacts with non-functional colors and lower than those of natural kinds (color A-FC: M=5.17, SD=1.81, comparing A-FC vs. A-NFC: z=3.8, p=0.0002, r=0.87; A-FC vs. NK: z=-3.66, p=0.0003, r=0.84). For all artifacts, we obtained ratings of an object color's relevance to the its function from a separate group of sighted Amazon Mechanical Turk participants. These function relevance judgments for artifacts were positively correlated with sighted participants' color consistency judgments (Spearman's rank correlation: rho=0.61, p<0.0001; Fig. 2C).

The same effect of object type on color and usage consistency judgments was observed in the blind group. Blind participants again judged natural kinds to have lower usage consistency but higher color consistency, compared to artifacts (blind usage NK: M=3.87, SD=1.61; A: M=6.19, SD=1.15; Wilcoxon matched-pairs test for usage NK vs. A: z=-3.92, p<0.0001, r=0.88; color NK: M=5.66, SD=1.52; A-NFC: M=3.32, SD=1.47; NK vs. A-NFC: z=3.92, p<0.0001, r=0.88). Artifacts with functional colors were judged to have higher color consistency than artifacts with non-functional colors, but lower than natural kinds (color A-FC: M=5.38, SD=1.81; comparing A-FC vs. A-NFC: z=3.92, p<0.0001, r=0.88; A-FC vs. NK: z=-1.98, p=0.049, r=0.44). Blind participants' consistency judgments for artifacts were positively also correlated with MTurk participants' ratings of color's relevance to object function (Spearman's rho=0.61, p<0.0001; Fig. 2c).

When groups were compared directly to each other, object kind and trial type did not interact with group (mixed ordinal logistic regression, group (blind vs. sighted) x trial type (color vs. usage) x object kind (NK vs. A-NFC), with sighted group, usage trial, and A-NFC treatment coded as baselines, no three-way interaction (β=-0.24, SE=0.38, z=-0.63, p=0.53). We also analyzed color judgments separately (group (blind vs. sighted) x object kind (NK vs. A-NFC vs. A-FC), with sighted and A-NFC as baseline). There was no significant interaction between group and object kind when comparing artifacts with functional color to artifacts with non-functional color (β=0.45, SE=0.27, z=1.65, p=0.099), although the interaction was significant when comparing natural kinds to artifacts with non-functional color (β=-1.02, SE=0.27, z=-3.78, p=0.0002).

Color consistency inferences in blind and sighted individuals: Novel objects

For real familiar objects, blind and sighted individuals could have made color consistency judgments based on knowledge of their actual color frequencies (e.g., learned from seeing or hearing that bananas are often yellow but that cars can be red, blue, black, etc.). Alternatively, people might have a more general understanding of the relationship between object kind (e.g. natural kind vs. artifact) and color. To distinguish between these possibilities, we collected color consistency judgments for novel objects, for which neither blind nor sighted participants could have directly experienced their color. Participants were presented with "explorer on an island" scenario and judged the consistency of color and usage for novel natural kinds (e.g. gem, plant) (5 objects), novel artifacts with non-function-relevant colors (e.g. cleaning-gadget, speaking device) (5 objects), and novel artifacts with function-relevant colors (e.g. coin, ceremonial clothing) (5 objects).

As with real objects, both groups judged artifacts to be more likely to have consistent usage than natural kinds (sighted usage NK: M=2.76, SD=1.42; A: M=5.64, SD=1.38; Wilcoxon matched-pairs signed rank test for NK vs. A: z=-3.82, p=0.0001, r=0.88; blind usage NK: M=3.25, SD=1.73; A: M=5.77, SD=1.58; NK vs. A: z=-3.92, p<0.0001, r=0.88). For color trials, consistency was again judged to be higher for natural kinds than for artifacts with non-functional color by both groups (sighted color NK: M=5.67, SD=1.12; A-NFC: M=3.73, SD=1.5; NK vs. A-NFC: z=3.81, p=0.0001, r=0.84; blind color NK: M=5.74, SD=1.42; A-NFC: M=3.95, SD=1.74; NK vs. A-NFC, z

= 3.72, $p<0.0001$, $r=0.85$). For both groups, artifacts with functional colors were judged as likely to have consistent colors as the natural kinds but more likely compared to artifacts with non-functional colors (for sighted color A-FC: $M=5.55$, $SD=1.43$; NK vs. A-FC: $z=-0.65$, $p=0.52$; A-NFC vs. A-FC: $z=3.78$, $p=0.0002$, $r=0.87$; for blind color A-FC: $M=5.92$, $SD=1.50$; NK vs. A-FC: $z=1.03$, $p=0.3$; A-NFC vs. A-FC: $z=3.9$, $p<0.0001$, $r=0.87$).

The interaction between group, question type, and object kind was non-significant (mixed ordinal logistic regression, three-way interaction: $\beta=-0.15$, $SE=0.93$, $z=-0.16$, $p=0.88$). The group-by-condition interaction for color trials only were also not significant (for NK vs. A-NFC: $\beta=-0.07$, $SE=0.36$, $z=-0.18$, $p=0.86$, for A-FC vs. A-NFC: $\beta=0.52$, $SE=0.39$, $z=1.39$, $p=0.17$). In sum, blind and sighted people have similar understanding of the relationship between object kind and color.

Blind and sighted people’s causal explanations of object color (Experiment 3)

In Experiment 3, blind and sighted participants were asked to explain why each object has its particular color. The explanations were coded according to what type of information they appealed to: process, depends on…, just is that way, material, social, maker of the object, visibility and cultural convention (see Supplementary Materials for coding details). Both blind and sighted participants provided rich and coherent explanations of the cause of object color (Figure 4). Both groups tended to provide different explanations for natural kinds, artifacts with non-functional colors, and artifacts with functional colors. For natural kinds, both groups most often said “it just is that way” (sighted: 33%, blind: 36%) or appealed to a process that give the object its color (sighted: 32%, blind: 31%). For example, participants often described how the process of photosynthesis makes plants green. By contrast for artifacts with non-functional colors (e.g. cars) both blind and sighted participants appealed to people’s social and esthetic preferences (sighted: 64%, blind: 44%), and referred to the material of which the object was made (sighted: 18%, blind: 13%). For example, people frequently stated “personal preference” as a cause for cars, and for cup, mentioned that they could be different colors depending on whether they are made of plastic, porcelain, or metal. For artifacts with functional colors, participants most often appealed to cultural convention (sighted: 57%, blind: 51%) and visibility (sighted: 24%, blind: 23%). For example, for school bus, participants frequently mentioned tradition and history, and for stop sign, that the color makes it easy to see.

We examined how similar explanations were across groups by computing Spearman's correlation across groups within object kind. The frequencies of explanations by type were highly correlated across groups for all three kinds of objects (natural kind: rho=0.99, p<0.0001; artifacts with non-functional color: rho=0.72, p=0.03; artifacts with functional color: rho=0.97, p<0.0001). Correlations across object kinds within each group were comparatively much lower (within sighted group: natural vs. A-NFC: rho=-0.31, p=0.4; natural vs. A-FC: rho=-0.27, p=0.5; A-NFC vs. A-FC: rho=0.78, p=0.01; within blind group: natural vs. A-NFC: rho=-0.02, p=1; natural vs. A-FC: rho=-0.37, p=0.3; A-NFC vs. A-FC: rho=0.28, p=0.5).

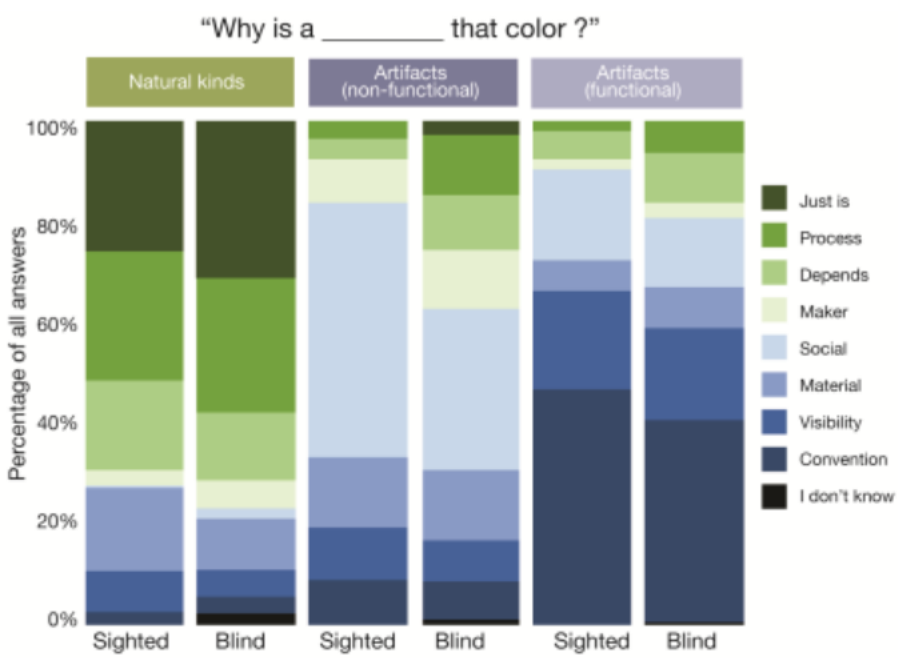

Fig 4. Explanations about object color. Explanation types were coded by 5 different coders who were blind to group and object. Stacked bar shows the frequency of each explanation type

as a proportion of all explanations provided for an object (within object type) across participants (within a group). A detailed key of explanation types can be found in Supplementary Materials.

Discussion

A straightforward idea is that we acquire color knowledge through seeing. Consistent with this, we find that people who have never seen are less likely to agree with each other and with sighted people about associative color facts: although 100% of blind participants generate the label ‘white’ for snow, only 50% say ‘yellow’ for bananas (see also 20, 21). This observation suggests that when it comes to learning associative color facts, direct visual access is more effective than linguistic communication.

By contrast, we find that inferentially rich color knowledge is shared among blind and sighted individuals—blind and sighted participants alike judge that two instances of a natural kind (e.g. two bananas or two gems) are more likely to have the same color than two instances of an artifact (e.g. two cars or two mugs). Blind and sighted people also provide similar explanations of why real objects have the colors that they do, and these explanations vary systematically across natural kinds and artifacts. For natural kinds, both blind and sighted appeal to an objects’ intrinsic nature (e.g., “that’s just how it is”, “that’s nature”) or describe processes such as photosynthesis, growth, or evolution. For artifacts, participants consistently cite individual or groups of people’s needs and intentions (e.g., culture, aesthetic preference, visibility). Blind individuals produce coherent explanations for object color even when they do not agree with the sighted about the typical color of that particular object type. For example, while both groups’ explanations for the color of polar bears mention their arctic habitat, almost all sighted participants explain that their white fur allows camouflage in the snow while some blind participants explain that they are black to absorb heat in the cold. (Interestingly, polar bears indeed have black skin underneath their transparent fur, and these features are thought to have evolved for both camouflage and heat absorption) (33). Such cases provide an illustration of causal understanding of color that is independent of knowing object-color associations.

Blind and sighted people's intuitions about the relationship between kind and color go beyond the natural kind/artifact distinction (34). Among artifacts, ratings of how important color is to an artifact's function are highly correlated with blind and sighted participants' ratings of color consistency. Explanations produced by sighted and blind adults also vary systematically by artifact type. For household and personal items such as mugs and cars, participants appeal to aesthetic preferences. For institution-related objects like police uniforms and dollar bills, participants cite social need for recognition. For stop signs, participants appeal to visibility (e.g., "red because red jumps out and warns people to stop"). Across artifacts, sighted and blind alike appeal to a range of causes such as camouflage, recognizability, cultural convention, symbolism, history, and aesthetic preference.

Finally, sighted and blind people make similar color consistency inferences for novel objects with which neither group has visual or linguistic experience. For example, both blind and sighted participants judge that two instances of a novel gem (natural kind) would be more likely to have the same color than two instances of a novel household gadget (artifact). Blind and sighted people also make distinctions within novel artifacts, intuiting which are most likely to have functionally relevant and therefore consistent colors (e.g. coins, toxic waste containers). Together, this evidence suggests that people living in the same culture, regardless of their visual experience, develop similar intuitive theories of color and use these theories to make inferences that go beyond the data.

While the present evidence suggests that blind and sighted people alike have a coherent and causal understanding of color, this understanding is likely to differ in substantial ways from formal scientific color theories (35, 36, 12). Participants' explanations of object colors did sometimes cite scientifically studied processes (e.g., photosynthesis), but more commonly consisted of informal justifications lacking mechanistic detail (e.g., "that's just how it grows", "it's nature”, “God made it that way”, “manufacturer decided to paint it that way”, “the material it’s made of”). When more specific causes and processes are mentioned, they are often social and historical, and unlikely to be taught through formal education (e.g., both blind and sighted participants mentioned personality of the owner for cars and “the patriarchy” for the color of wedding dresses). During development, sighted children’s beliefs about color depart systematically from scientific knowledge. Children mistakenly believe that an object will continue to have the same color even when the lighting source is changed, that objects emit their own shadows, and that a green object will have a green shadow (37-39; see also 40). Children's explanations about such phenomena omit crucial components, such as the source and nature of light illuminating an object (37). Similar inconsistencies between scientific and intuitive theories have been observed in numerous other knowledge domains (e.g., physics: 41, 1983; biology: 42; psychology: 43). Even when educated adults and experts report strong confidence in their own understanding, their explanations for how things work are coarse and incomplete (44). Future work is needed to understand the ways in which intuitive theories of color among sighted and blind people share features with and depart from scientific color theories.

Open questions remain about how blind and sighted people acquire causal intuitions about color. Linguistic communication likely plays a crucial role. From a young age, children use testimony as well as more implicit linguistic cues (e.g., labeling) to inform intuitive theories of physical, biological, and mental phenomena (45, 1, 46). For many previously studied domains of knowledge, language-induced learning could in principle piggyback on pre-existing structured knowledge built through sensory observation. For example, learning that the earth is round might piggyback on learning roundness through vision and touch (47). Even in the case of mental phenomena, simulation of one’s own feelings and thoughts has been offered as a source of “first-person” information about others’ minds (48, 49). Analogously, a sighted person might construct a representation of a novel animal described as blue and large by referencing physical knowledge previously built up through sensory experience of color and size (20). In the case of color knowledge among blind individuals, there is no directly pertinent sensory information. Nevertheless, inferentially rich knowledge is constructed through linguistic communication from the ground up.

Recent text corpus analyses also find that language is a rich source of semantic information, including that of physical appearance. Associative algorithms are able to extract semantic information using word co-occurrence and word neighborhood statistics (e.g., 50-52, 22-23). The available evidence suggests, however, that people's learning of appearance from language differs in important respects from the statistical tracking used by current text analysis algorithms. For example, in the case of animal appearance, people born blind know more about shape, texture, and size, than what is extracted by co-occurrence tracking algorithms (22, 53). People born blind appear to use taxonomy and habitat to infer physical characteristics rather than relying purely on explicit statements about appearance (20). In the case of object color consistency, it is not clear how tracking word co-occurrence alone would provide the correct information. Analyses of object and color label co-occurrence suggest that object-color label co-occurrence consistency does not line up with real-life object-color consistency, nor with sighted people's intuitions about the object's actual color (23). For example, 'crow' co-occurs with 'black' and 'white' with similar frequencies. Nevertheless, we find that blind and sighted people's intuitions about color consistency are similar. Moreover, blind participants generate canonical colors (i.e. black for crow) more often than non-canonical ones (i.e. white for crow), even when the canonical and non-canonical colors are equally likely to co-occur with the object in text (23, 54).

One source of information blind individuals could use to arrive at color consistency is generic language (e.g. tomatoes are red). Generics provide evidence that a property is pervasive to an object, as opposed to specific to a particular instance of that object (e.g., 26, 55-57). For example, hearing "this car is red," as opposed to "tomatoes are red" and "stop signs are red" could provide evidence with respect to which objects tend to have a consistent color. Generic language could also facilitate learning which of two possible color labels is the canonical one (e.g. this crow is white vs. crows are black). Generic language alone cannot, however, explain how blind and sighted people make similar judgments about novel objects, which they have not heard being described in a generic sentence, or how people born blind generate coherent explanations of color. Linguistic cues such as generic language, must make contact with early-emerging intuitive theories and the inferential machinery to transform these theories in light of the linguistic evidence (57-59). We hypothesize that people born blind use linguistic cues, such as generics, to fill in the color-specific elements of intuitive theories about objects and their physical properties. For example, hearing people talk about "favorite colors", together with evidence that a particular personal item (e.g. cars) varies in color, might lead one to conclude that personal items are sometimes colored according to the preferences of the owner. Future work is needed to uncover precisely how blind and sighted people use language as a source of information when constructing intuitions about color. One important direction for future research is testing the acquisition of such knowledge by blind and sighted children. Such studies would reveal exactly when, how, and with what information such intuitions are constructed.

In summary, we find that blind and sighted individuals alike possess theory-like, inferentially rich knowledge about the relationship between objects and their colors. These intuitive theories of color support consistent generalizations in the face of limited information (e.g., for novel objects), invoke deep causes (e.g., object function), support the generation of sophisticated explanations, apply to broad categories (e.g., all plants) as well as to specific instances (e.g., polar bears), and are specific to color. Interestingly, such structured and inferentially rich color knowledge appears to be more resilient to the lack of first-person sensory experience than knowledge of associative color facts. This observation directly contradicts the common intuition that blind people's knowledge of color consists of meaningless arbitrary facts. Language appears to support the updating of causal models much more robustly than it does the acquisition of arbitrary facts.

Methods

Participants

Twenty congenitally blind (14F/6M, age: M=30.85, SD=10.59, years of education: M=15.4, SD=2.23) and nineteen sighted (14F/5M, age: M=31.21, SD=11.21, years of education M=15.79, SD=1.82) participants took part in the study (participant table can be found in Supplemental Materials, Table S1). All blind participants reported no experience with color, shape, or motion, and had at most minimal light perception. All blind participants were tested at the 2018 National Federation of the Blind Convention in Orlando, Florida. Subtests of the Woodcock Johnson III Tests of Achievements (Word ID, Word Attack, Synonyms, Antonyms, and Analogies) were administered to sighted and blind participants, and anyone scoring below two SDs from their own group's mean was excluded from further analyses. This resulted in one sighted participant (participant 20) being excluded. The study consisted of three experiments administered to all participants within the same session. Experimental procedures were approved by the Johns Hopkins Homewood Institutional Review Board, and all participants provided informed consent.

Experimental Procedures Overview

Experiment 1 and 3 queried knowledge of and inferences about the colors of real objects (30 objects in Experiment 1, 24 in Experiment 3). In Experiment 2, participants made color inferences about 15 novel objects. Experiment 2 was always administered first to prevent the real object judgments from influencing inferences made about novel objects. Within each experiment, two different trial orders were used, one for half of the participants within each group. Experimenters read aloud instructions and trials, and participant answers were audio-recorded and later transcribed for scoring. The full list of stimuli and instructions can be found in the Appendix (Supplemental Materials).

Experiment 1: Knowledge of Real Object Colors

In each trial of Experiment 1, participants were asked two questions about an everyday object (Fig. 1). Three types of questions were asked: color consistency (30 objects), usage consistency (20 objects), and fillers (20 objects). Objects used for color trials were either natural (10 objects) or manmade (20 objects), and manmade artifacts could have function-relevant color (FC, 10 objects) or non-function-relevant color (NFC, 10 objects). Usage trials consisted of 10 natural kinds and 10 artifacts. On filler trials, participants were asked questions about non-color features (size, shape, and texture). Filler trials consisted of 5 natural kinds and 15 artifacts in order to balance the overall number of natural kind and manmade trials. The full list of items used in color and usage trials can be found in main Figure 1.

On color trials, participants were first asked, "What is one common color of (a) [object name]?", followed by, "If you picked two [object name]s at random, how likely are they to be the same color? Rate on a scale of 1 to 7 (1: 'unlikely', 3: 'somewhat likely', 5: 'likely', 7: 'very likely')."

For usage trials, the questions were, "What is one common thing you can do with (a/some) [object name]?" and, "If you picked two people at random and asked them each to do something with (a/some) [object name], how likely are they to do the same thing, on a scale of 1 to 7?" Usage trials served as a control condition to ensure blind and sighted participants showed equivalent performance and were willing to rate artifacts as having some consistent properties.

Experiment 2: Color inferences about novel objects

In Experiment 2, in order to elicit inferences about novel objects parallel to in Experiment 1, participants were first presented an "Explorer on an Island" scenario:

"Imagine that you're an explorer, and on your travels, you've discovered an island in a remote corner of the world… You learn that the people on this island call themselves Zorkas… The Zorka people appear to have a highly advanced culture. They have their own language, tools, machines, buildings, vehicles, foods, customs, and so on. The ecology on this island is also

different from what we're used to: it has its own plant and animal life, unusual rocks, minerals, and so on. You're trying to learn about how things work on this island...."

Participants then heard 35 short vignettes, each of which described an encounter with a novel object (natural kind, artifacts with functional color, and artifacts with non-functional color; Fig. 1). In each trial, several appearance features were noted (e.g., "green gem that is spiky and the size of a hand"). The object was then named (e.g., "The miners tell you that this gem is called an Enly.").

As in Experiment 1, participants were next asked to rate the likelihood that another instance of the same object would have the same color (e.g., "How likely is it that the next time you come across another Enly, it is also green?). In usage trials, the question asked the likelihood that the novel object would be used in the same way if encountered at another time (e.g., "How likely is it that the next time you come across another Irve, it is also being ripped out of the ground?"). In addition, there were 10 filler trials (7 natural kind, 3 manmade objects), in which participants were asked about the likely repeat occurrence of a non-color feature (e.g., shape, texture, size).

Color trials consisted of 5 natural kinds (plant, algae, gem, liquid from a plant, fruit), 5 artifacts with function-relevant color (coin, road symbol, toxic waste container, ceremonial clothing, clear substance being used to build a wall), and 5 artifacts with function-irrelevant color (an odor-emitting gadget, roof cleaning machine, two devices with ambiguous functions).

Usage trials consisted of 5 natural kind (creature, boulder, stone, flower, plant) and 5 artifacts (machine that makes square holes, storage device, toy, machine that turns stones into goo, and one contraption with ambiguous function).

Filler trials contained 7 natural kind (fruit, two creatures, rock, two plants, gem) and 3 artifacts (game device, type of pool, one contraption with ambiguous function).

Experiment 3: Explanations about the cause of object color

For an additional list of 24 real objects (8 natural kind, 9 manmade with functional color, 7 function-irrelevant color), we asked participants to report their common colors (as in Experiment 1). Common color reports for these 24 objects are collapsed with those from Experiment 1 in main Figure 2. For these objects, we additionally asked why objects had the particular color (or colors) that the participant provided: "*Why* are [object name]s that/those color[s]?" Participants were instructed to provide whatever explanation felt right to them. Participants were also asked whether the object has different colored parts, and if an object's color varies across instances, to report the other colors. The answer to these questions were not analyzed for the present study.

Quantifying color naming agreement for real objects

Across Experiments 1 and 3, participants named the color of 54 objects (Fig. 1) (Experiment 1: 30 objects, "What is one common color of…?" and Experiment 3: 24 objects, "What is the most common color of…?"). For each object, we quantified naming agreement by using the Simpson's Diversity Index (SDI) (32, 20). For unique color words (1 to R) provided for each object across all participants within a group (blind or sighted), a naming agreement score was calculated according to the equation below. N is the total number of words used across participants for each object, and n is the number of times each unique word (1 to R) was provided. The index ranges for 0 to 1, where 0 indicates that the same color word was never used by two participants (i.e., low color naming agreement), and 1 suggests all participants provided the same color (i.e., high naming agreement).

$$SDI = \frac{\sum_{i=1}^{R} n_i(n_i - 1)}{N(N-1)}$$

Although participants were instructed to provide one color, a few participants provided multiple colors (at most three, e.g., "red, white, and blue"). All of these colors were included in the analysis. Further, a small proportion of participants said "I don't know" or provided words that were not typical color terms (dark, light, beige, neon). These responses were treated the same as color terms ("I don't know" was counted as one word, coded "IDK"). Since SDIs were not normally distributed, they were log-transformed. To examine differences in color naming agreement across groups, we then performed linear mixed effects regression on log-transformed SDIs, using lmer in R (60), with objects as random effects.

Color consistency inference analysis

Consistency likelihood judgments were analyzed using ordinal logistic regression using the ordinal (61) package in R. Participants and objects were always included as random effects, and separate models were used in each analysis described (e.g., for real vs. novel objects).

We first compared group differences for natural kinds and artifacts with non-functional color only, since artifacts with functional color are a special category. This also allowed us to look at a group (blind vs. sighted) x object kind (natural vs. artifact) x trial type (color vs. function) three-way interaction. Baselines were coded as sighted group, usage trial, and artifact. We then compared across groups for color trials only, this time including all three kinds of objects (natural, artifact with functional color, artifact with non-functional color), with sighted group and artifact with non-functional color as the baseline.

Correlation with functional relevance of color for artifacts

We obtained ratings from Amazon Mechanical Turk (n=20) for the functional relevance of color to artifacts. Participants were asked "How important is the color of a [object] to its function?" and had to rate on a scale of 1 to 7 (not at all to very relevant). Artifacts designated as 'artifacts with functional colors' were those that received an average rating of 4 or above, and artifacts 'non-functional colors' all had ratings below 3 (Table S2). We correlated the average functional relevance ratings for each object with the average color consistency judgments, for blind and sighted groups separately (Spearman correlation).

Analysis of explanations

Explanation types were decided by the experimenters based on examining all the explanations (while blind to group and object). We decided on 9 types of explanations: 'process', ‘depends on’, ‘just is’, ‘material’, ‘social/aesthetic’, ‘maker’, ‘visibility’, ‘convention’, and ‘I don’t know’. A key of explanations can be found in Supplemental Materials (Table S3).

Explanations were coded by four coders who did not know which object or group each explanation came from. Note, however, that in a small number of instances participants said the object’s name in their explanations, and at other times, it was fairly easy to discern the object from the explanation.

There was large variability in how many words participants used in their explanations (range=1 to 165 words, M=13 words). This meant that each explanation (i.e., what one participant said for one object) could contain multiple explanation types. For example, a participants’ answer that the color of a wedding dress is due to “symbolism, or personal style”, was coded as containing ‘convention’ (for symbolism) and ‘social/aesthetic’ (for personal style) explanations. However, the same word or phrase (e.g., “personal style”) was never coded for more than one explanation type.

Some participants gave lengthier explanations than others, without necessarily providing additional information (e.g., often telling anecdotal stories to make a point). For wedding dress, for instance, another participant explained: “well, there’s something about tradition, and white being associated with purity and virginity and all that, but beyond that it’s just a matter of demand, if you want a baby barf green wedding dress that’s your problem”. This explanation was also coded with ‘convention’ and ‘social/aesthetic’.

Coding was then filtered according to the criteria that at least three out of four coders have to agree. The first author (5th coder) made some additional changes, again keeping group and objects blind, and overruled tagging for <5% explanations. After this process, the number of explanation types per explanation (again, a single explanation from one participant for one object) only ranged from 1-3 (mean=1.26).

We compared explanations across groups within each object kind. Within a group and kind (e.g., sighted group, natural kinds), we calculated how frequently participants (across all participants within group) used each of the 9 explanation types. The counts were then calculated as a percentage of all explanations (within group and object kind). We then computed Spearman's correlations over the percentages (for 9 types) across groups, as well as across object kinds within groups.

References

1. Gelman, S. A. (2009). Learning from others: Children's construction of concepts. Annual review of psychology, 60, 115-140.

2. Lombrozo, T. (2019). “Learning by Thinking” in Science and in Everyday Life. In The scientific imagination. Oxford University Press.

3. Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and brain sciences, 22(4), 577-660.

4. Prinz, J. J. (2005). The return of concept empiricism. In Handbook of categorization in cognitive science (pp. 679-695). Elsevier Science Ltd.

5. Machery, E. (2006). Two dogmas of neo-empiricism. Philosophy Compass, 1(4), 398-412.

6. Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. Infant Behavior and Development, 21(2), 181-200.

7. Carey, S. (2009). The origin of concepts. Oxford university press.

8. Meltzoff, A. N., & Gopnik, A. (2013). Learning about the mind from evidence: Children's development of intuitive theories of perception and personality. Understanding other minds, 3, 19-34.

9. Locke, J. (1924). 1690. An essay concerning human understanding, 1.

10. Hume, D. (1938). An Abstract of a Treatise of Human Nature, 1740. CUP Archive.

11. Jackson, F. (1982). Epiphenomenal qualia. The Philosophical Quarterly (1950-), 32(127), 127-136.

12. Jackson, F. (1986). What Mary didn't know. The Journal of Philosophy, 83(5), 291-295.

13. Dennett, D. C. (1993). Consciousness explained. Penguin uk.

14. Tye, M. (2000). Knowing what it is like: The ability hypothesis and the knowledge argument. Protosociology, Collection of Essays for David Lewis.

15. Chalmers, D. J. (2004). 13 Phenomenal Concepts and the Knowledge Argument. There's Something About Mary: Essays on phenomenal consciousness and Frank Jackson's knowledge argument, 269.

16. Gleitman, L., & Landau, B. (1985). Language and experience: Evidence from the blind child.

17. Shepard, R. N., & Cooper, L. A. (1992). Representation of colors in the blind, color-blind, and normally sighted. Psychological Science, 3(2), 97-104.

18. Marmor, G. S. (1978). Age at onset of blindness and the development of the semantics of color names. Journal of Experimental Child Psychology, 25(2), 267–278. https://doi.org/10/dpdnkw

19. Saysani, A., Corballis, M. C., & Corballis, P. M. (2018). Colour envisioned: Concepts of colour in the blind and sighted. Visual Cognition, 26(5), 382–392. https://doi.org/10/ggmt29

20. Kim, J. S., Elli, G. V., & Bedny, M. (2019a). Knowledge of animal appearance among sighted and blind adults. Proceedings of the National Academy of Sciences, 116(23), 11213-11222.

21. Connolly, A. C., Gleitman, L. R., & Thompson-Schill, S. L. (2007). Effect of congenital blindness on the semantic representation of some everyday concepts. Proceedings of the National Academy of Sciences, 104(20), 8241-8246.

22. Lewis, M., Zettersten, M., & Lupyan, G. (2019). Distributional semantics as a source of visual knowledge. Proceedings of the National Academy of Sciences, 116(39), 19237-19238.

23. Ostarek, M., Van Paridon, J., & Montero-Melis, G. (2019). Sighted people's language is not helpful for blind individuals' acquisition of typical animal colors. Proceedings of the National Academy of Sciences, 116(44), 21972-21973.

24. Keil, F. C., Smith, W. C., Simons, D. J., & Levin, D. T. (1998). Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge. Cognition, 65(2-3), 103-135.

25. Keil, F. C. (1992). Concepts, kinds, and cognitive development. mit Press.

26. Gelman, S. A. (2003). The essential child: Origins of essentialism in everyday thought. Oxford University Press, USA.

27. Greif, M. L., Nelson, D. G., & Kemler, K. FC, & Gutierrez, F.(2006). What do children want to know about animals and artifacts, 455-459.

28. Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. Cognitive psychology, 20(1), 65-95.

29. Bloom, P. (1996). Intention, history, and artifact concepts. Cognition, 60(1), 1-29.

30. Springer, K., & Keil, F. C. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. Child development, 62(4), 767-781.

31. Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. Oxford handbook of causal reasoning, 515-548.

32. Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O'grady, L., ... & Shayan, S. (2018). Differential coding of perception in the world's languages. Proceedings of the National Academy of Sciences, 115(45), 11369-11376.

33. Grojean, R. E., Sousa, J. A., & Henry, M. C. (1980). Utilization of solar radiation by polar animals: an optical model for pelts. Applied Optics, 19(3), 339-346.

34. Keil, F. C., Greif, M. L., & Kerner, R. S. (2007). A world apart: How concepts of the constructed world are different in representation and in development. Creations of the mind: Theories of artifacts and their representation, 231-245.

35. Conceptual differences between children and adults. Mind & Language, 3(3), 167-181.

36. Keil, F. C. (2010). The feasibility of folk science. Cognitive science, 34(5), 826-862.

37. Feher, E., & Meyer, K. R. (1992). Children's conceptions of color. Journal of research in Science Teaching, 29(5), 505-520.

38. Naranjo Correa, F. L., Martinez Borreguero, G., Perez Rodriguez, A. L., Suero Lopez, M. I., & Pardo Fernandez, P. J. (2016). A new online tool to detect color misconceptions. Color Research & Application, 41(3), 325-329.

39. Anderson, C. W., & Smith, E. L. (1986). Children's Conceptions of Light and Color: Understanding the Role of Unseen Rays. Research Series No. 166.

40. Cohen, J., & Nichols, S. (2010). Colours, colour relationalism and the deliverances of introspection. Analysis, 70(2), 218-228.

41. McCloskey, M. (1983). Intuitive physics. Scientific american, 248(4), 122-131.

42. IInagaki, K., & Hatano, G. (2006). Young children's conception of the biological world. Current Directions in Psychological Science, 15(4), 177-181.

43. Malle, B. F., Knobe, J. M., & Nelson, S. E. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. Journal of Personality and Social Psychology, 93(4), 491.

44. Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. Cognitive science, 26(5), 521-562.

45. Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. Annual Review of Psychology, 69, 251-273.

46. Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. Proceedings of the National Academy of Sciences, 114(30), 7900-7907.

47. Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. Cognitive psychology, 24(4), 535-585.

48. Gordon, R. M. (1986). Folk psychology as simulation. Mind & language, 1(2), 158-171.

49. Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. Trends in cognitive sciences, 2(12), 493-501.

50. Bruni, E., Boleda, G., Baroni, M., & Tran, N. K. (2012, July). Distributional semantics in technicolor. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1 (pp. 136-145). Association for Computational Linguistics.

51. Grand, G., Blank, I. A., Pereira, F., & Fedorenko, E. (2018). Semantic projection: recovering human knowledge of multiple, distinct object features from word embeddings. arXiv preprint arXiv:1802.01241.

52. Rubinstein, D., Levi, E., Schwartz, R., & Rappoport, A. (2015, July). How well do distributional models capture different types of semantic knowledge?. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers) (pp. 726-730).

53. Kim, J. S., Elli, G. V., & Bedny, M. (2019b). Reply to Lewis et al.: Inference is key to learning appearance from language, for humans and distributional semantic models alike. Proceedings of the National Academy of Sciences, 116(39), 19239-19240.

54. Kim, J. S., Elli, G. V., & Bedny, M. (2019c). Reply to Ostarek et al.: Language, but not co-occurrence statistics, is useful for learning animal appearance. Proceedings of the National Academy of Sciences, 116(44), 21974-21975.

55. Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. Cognitive science, 34(8), 1452-1482.

56. Tessler, M. H., & Goodman, N. D. (2019). The language of generalization. Psychological review, 126(3), 395.

57. Brandone, A. C., & Gelman, S. A. (2009). Differences in preschoolers' and adults' use of generics about novel animals and artifacts: A window onto a conceptual divide. Cognition, 110(1), 1-22.

58. Csibra, G., & Gergely, G. (2009). Natural pedagogy. Trends in cognitive sciences, 13(4), 148-153.

59. Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. Perspectives on Psychological Science, 7(4), 341-351.

60. Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1).

61. Christensen RHB (2019). "ordinal—Regression Models for Ordinal Data ." R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

Acknowledgments: We thank the National Federation of the Blind convention, the blind community, and all of our participants for making this research possible; **Funding:** This work was supported by the National Institutes of Health (R01 EY027352 to M.B.) and the Johns Hopkins University Catalyst Grant (to M.B.); J.S.K was funded by the William Orr Dingwall Dissertation Fellowship; **Author contributions:** J.S.K, B.A, V.M, and M.B. designed research; J.S.K, B.A, and V.M, performed research, J.S.K analyzed data, and J.S.K and M.B. wrote the paper; **Competing interests:** Authors declare no competing interests; and **Data and materials availability:** All data, code, and materials used in the analysis are available in the repository: https://github.com/judyseinkim/Intuitive-Theories-of-Color and a detailed description of analyses can be found in the following document: https://rpubs.com/judyseinkim/color_theory .