Manchester Metropolitan University

e-space

Manchester Metropolitan University's Research Repository

Please cite the Published Version

Roberts, Pendaran and Schmidtke, Kelly (2019) Folk core beliefs about color. Review of Philosophy and Psychology, 10 (4). pp. 849-869. ISSN 1878-5158

DOI: https://doi.org/10.1007/s13164-019-00437-w

Publisher: Springer Verlag

Version: Published Version

Downloaded from: https://e-space.mmu.ac.uk/622635/

Usage rights: Creative Commons: Attribution 4.0

Additional Information: This is an Open Access article in Review of Philosophy and Psychology published by Springer.

Enquiries:
If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)

Review of Philosophy and Psychology
https://doi.org/10.1007/s13164-019-00437-w

Folk Core Beliefs about Color

Pendaran Roberts[1] · Kelly Ann Schmidtke[2]

Published online: 22 April 2019
© The Author(s) 2019

Abstract
Johnston famously argued that the colors are, more or less inclusively speaking, dispositions to cause color experiences by arguing that this view best accommodates his five proposed core beliefs about color. Since then, Campbell, Kalderon, Gert, Benbaji, and others, have all engaged with at least some of Johnston's proposed core beliefs in one way or another. Which propositions are core beliefs is ultimately an empirical matter. We investigate whether Johnston's proposed core beliefs are, in fact, believed by assessing the agreement/disagreement of non-philosophers with them. Two experiments are run each with large sample sizes, the second designed to address criticisms of the first. We find that non-philosophers mostly agree with the proposed core beliefs, but that they agree with some more than others.

1 Introduction

What is the nature of color? This question has been given numerous answers: colors are micro-structural physical properties of objects (Kripke 1972; Jackson and Pargetter 1987); colors are physical dispositions to reflect light, a.k.a. reflectance dispositions (Tye 2000; Byrne and Hilbert 2003, 2004); colors are appearance dispositions to cause color experiences (McGinn 1983; Johnston 1992); colors are primitive or sui generis properties (Yablo 1995; Chalmers 2006). How should we decide between these views? Philosophers have appealed to intuitive truths, phenomenal truths (e.g. how colors look), to common sense, and to scientific findings in physics and psychophysics. Johnston's well-known meta-methodology is to adjudicate between competing views by examining which view best accommodates our core beliefs about color, where core beliefs set conceptual limits on what counts as *a color* (1992). Core beliefs about color

✉ Pendaran Roberts
Pendaran@mac.com

[1] University of Manchester, Manchester, England

[2] Manchester Metropolitan University, Manchester, England

are beliefs that are widely held by the general population (the folk) such that if a property is *ever so inclusively* a color, as opposed to just *more or less inclusively* a color, then all the core beliefs about color are true of it.[1] An *ever so inclusive* view is just one that captures or is compatible with all our core beliefs, whereas a *more or less inclusive* view fails to capture or be compatible with one or more of our core beliefs. Johnston leaves it open exactly how inclusive we have to be in order to avoid changing the subject matter.[2]

Johnston proposed five core beliefs and argued that ultimately an appearance dispositional view to the effect that colors are dispositions to cause color experiences best accommodates them. Since then, Yablo (1995), Campbell (2005), Kalderon (2007), Gert (2008), Allen (2011), Benbaji (2016), and Roberts (2018) have all engaged with at least some of Johnston's proposed core beliefs. Although Johnston's list may not be exhaustive, it includes those core beliefs most salient to present debates in the ontology of color. We present the list below using the color 'canary yellow,' as Johnston does, but the core beliefs apply to all colors.

> *Paradigms*. Some of what we take to be paradigms of canary yellow things (e.g. some canaries) are canary yellow.
>
> *Explanation*. The fact of a surface or volume or radiant source being canary yellow sometimes causally explains our visual experience as of canary yellow things.
>
> *Unity*. Thanks to its nature and the nature of the other determinate shades, canary yellow, like the other shades, has its own unique place in the network of similarity, difference and exclusion relations exhibited by the whole family of shades.
>
> *Availability*. Justified belief about the canary yellowness of external things is available simply on the basis of visual perception. That is, if external things are canary yellow we are justified in believing this just on the basis of visual perception and the beliefs which typically inform it.
>
> *Revelation*. The intrinsic nature of canary yellow is fully revealed by a standard visual experience as of a canary yellow thing.

Empirical methods would be useful for deciding whether any of these five propositions are, in fact, core beliefs. Without an empirical investigation into people's core beliefs, we have only Johnston's intuitions and those of his proponents to rely on. The intuitions of philosophers alone provide insufficient reason. First, we should doubt

[1] The way we understand the issue these beliefs need not be explicit but may be implicit. We leave it open what exactly it means to have an implicit belief. However, we assume it means this at minimum: implicit beliefs are those beliefs that cannot be openly recalled or stated but influence a person's thoughts and actions.

[2] As Johnston (1992) rejects a sharp boundary between the analytic and the synthetic, he holds that there is a decent amount of vagueness regarding which beliefs are core and so constitutive of the subject matter. He thinks that pragmatic concerns can be used to select certain beliefs as core vs. others. Thus, he thinks we can have a view on color without having to accommodate every core belief or, more precisely, candidate core belief. Nevertheless, ceteris paribus, the more core beliefs that a view can accommodate the stronger a claim it has to be correct.

whether highly-educated philosophers retain folk beliefs. Perhaps philosophers' beliefs have been tainted by theoretic considerations (Goldman 2007). Johnston (1992) when discussing his core beliefs talks of the theoretic work of other philosophers (e.g. Russell and Strawson). So, it is not difficult to imagine that Johnston's beliefs may not be the beliefs of ordinary people. Second, without independent reason to endorse the five beliefs proposed by Johnston (or any others), opponents can just reject their content as question begging. This is exactly how Benbaji (2016) responds to Johnston's Revelation. An empirical investigation can provide some independent justification.

According to Johnston, his proposed core beliefs play a significant role in debates about the nature of color. Paradigms is relevant to the realist/irrealist debate.[3] Irrealists argue that external objects are not actually colored (Boghossian and Velleman 1989; Hardin 1993, 2004; Maund 1995; Chalmers 2006). It has been debated whether Irrealism can be a view on color ever so inclusively speaking (Stroud 2000). If Paradigms is a core belief, the colors ever so inclusively speaking must exist. Paradigms is relevant to the irrealist debate in another way if we assume that anything which is a core belief is intuitive and/or common sense: if so, and Paradigms is a core belief, there is an argument via intuition and/or common sense for realism. It is probably in part because of an argument from intuition and/or common sense that realism is the dominant position in philosophy. Of course, Paradigms, as well as Johnston's other core beliefs, do not have to be core beliefs per se to be common sense or intuitive. If one balks at the idea that certain beliefs are such that they are constitutive of a subject matter, one may prefer to think in terms of common sense. If one balks at the idea that people have such complex beliefs about color, even implicitly, one may prefer to think in terms of intuition.

Explanation, Unity, and Availability are relevant to debates between positions on color which hold that they are lower-level physical properties (micro-structuralism and reflectance physicalism) and positions which hold that they are higher-level properties (appearance dispositionalism and primitivism). If Explanation is a core belief, whatever the colors are, ever so inclusively (for simplicity, we will sometimes drop this qualification entirely), they must causally explain our experiences of color. However, the overdetermination argument suggests that only lower-level physical properties can causally explain our color experiences: any explanation provided by higher-level properties is redundant (Kim 1993a, 1993b; Hardin 1993, p. 61). There is a related reason why Explanation is relevant: if only science tells us what causally explains our color experiences and Explanation is true, we must look to the properties discussed in science if we are to find the colors. It may be this kind of reasoning in part that motivates Johnston's narrowing his search to exclude primitivism (1992, p. 224).

If Unity is a core belief, the various colors must, because of their natures, stand in the appropriate relations of similarity, difference, and exclusion. Unity poses a special problem for lower-level physical views. There are at least 14 features at the micro-structural level that are sufficient to cause normal observers in normal conditions to

[3] We note that Explanation, as phrased by Johnston, also implies that objects are colored. It could be phrased as a conditional so as to avoid this implication and avoid overlap with Paradigms. The other core beliefs do not clearly make any claims about objects being colored. Unity only speaks of the higher order properties of the colors, which they could have regardless of being instantiated. Availability speaks of 'justified belief' not 'justified true belief,' and Revelation speaks of 'visual experience,' not 'veridical visual experience.'

have an experience as of an object being blue, for example (Nassau 1983, 1997).[4] Assuming Explanation, the micro-structuralist should look to these properties to determine what colors are, but it is unknown whether any of these properties can accommodate Unity. There is a well-known argument that reflectance physicalism cannot accommodate Unity (Hardin 1993, p. 66; Maund 1995, p. 126–133; Thompson 1995, p. 124; Pautz 2006): reflectances do not necessarily stand in the required similarity relations, for upon examining the relevant reflectance curves (Pautz 2006), no justified means to hold that Unity is true of them can be found (likewise for types of reflectances).

If Availability is a core belief, this poses a problem for views on color that reduce colors to properties for which perception does not provide justification. Johnston (1992) points out that Unity and Availability taken together pose a problem for lower-level physical views: if a lower-level physical view is adjusted to accommodate Unity, it violates Availability (p. 238). Johnston's argument runs thus: if the colors were lower-level physical properties which accommodate Unity, then we could only know that there were colors if we knew that there were lower-level physical properties which accommodate Unity, but knowing this would require a sophisticated scientific investigation (pp. 238–238). So, if Unity is true of lower-level physical colors, then these colors cannot accommodate Availability. Arguably, Availability alone poses a problem for lower-level physical views: via Leibniz's law (if $x = y$, then x and y share all the same properties) one can deduce that Availability is false for the colors if they are lower-level physical properties, because we are not justified in believing objects have lower-level physical properties based simply on visual perception (Jackson 2012).

Revelation is the purported core belief that especially favors a primitivist view of color. Revelation states that the intrinsic nature of canary yellow is fully revealed by a standard visual experience as of a canary yellow thing, but the colors being appearance dispositions or physical properties is not revealed by a standard visual experience(s) (Johnston 1992, p. 225). So, if Revelation is, in fact, a core belief, the colors are not micro-structural, reflectance dispositions, or appearance dispositions (at least not ever so inclusively speaking). Johnston, of course, believes that Revelation is a core belief. He thinks that no view which satisfies Explanation can satisfy Revelation (p. 224). If this is right, which purported core belief is empirically a better candidate is highly relevant. Johnston favors Explanation over Revelation, but it is unclear whether this decision is warranted. If Revelation is the better candidate, perhaps we should throw out Explanation and adopt primitivism. If Explanation is a better candidate, perhaps we should throw out Revelation and adopt a non-primitivist view on color, for example, micro-structuralism. If they are both equally good candidates, this is important too, for it shows that neither can be preferred. Given Johnston's core belief meta-methodology for adjudicating between competing views on color, empirical research could help to decide what to do.

In this article, we empirically test Johnston's proposed core beliefs by asking the folk how much they agree with them. If a proposition is a core belief, then agreement with it should be high. Core beliefs need to be widely believed by non-philosophers, and it is

[4] These features comprise the following list: Incandescence, gas excitations, vibrations and rotations, transition metal compounds, charge transfer, metals, pure semiconductors, doped semiconductors, color centers, dispersive refraction, scattering, interference, and diffraction.

expected that if participants believe that a proposition is true, then they will agree with it. Just how high this agreement must be is not something we will discuss, as any number seems ad hoc. Rather, we will speak in terms of better or worse candidates for core beliefs. What we find is that agreement is high for all five of Johnston's proposed core beliefs, but, importantly, that agreement is higher for some than others. Ceteris paribus, the higher the agreement, the better a candidate for a core belief. We ran two experiments with large participant numbers. The second was designed to improve on arguable weaknesses of the first. We will present Experiment 1, then Experiment 2, then discuss the results and their impact on the ontology of color. Our goal is not to make a priori research in this area redundant but to add to it with our empirical findings.

For ease of exposition, this paper assumes that Johnston's meta-methodology is worth our time. This paper will make no attempts to argue for this meta-methodology. This is an empirical paper testing whether Johnston's core beliefs are widely believed. It is not an a priori paper on philosophical methodology, and there is not the space for this debate either. This said, one need not ascribe to Johnston's method to recognize his five 'core beliefs' as relevant to the ontology of color more generally. As we have mentioned, Paradigms, as well as Johnston's other core beliefs, do not have to be core beliefs per se to be common sense or intuitive. In other words, it is possible to separate the propositions themselves from the meta-methodology of core beliefs. So, there is no reason this paper cannot be viewed as providing an empirical basis for an argument from common sense or intuition for views which better accommodate the content of Johnston's core beliefs. To save space, we will mostly leave the construction of these arguments to the reader.

2 Experiment 1

2.1 Method

Survey Development In preliminary surveys, we sought to design test items that could be understood by non-philosophers while retaining content accuracy. The survey development was ethically approved by the Humanities and Social Sciences Ethics Committee at the University of Warwick (REF: 14/15–16:DR@W). Initially, survey items were chosen via cooperative reflection between the co-authors. These items were refined iteratively through conversations with available academics (Berit Brogaard, Keith Allen, and Derek Brown) and lay people. Then, 30 participants were recruited using the University's sona-system and entered into a lottery draw for an Amazon gift voucher upon completion. Students at the University of Warwick are actively recruited to join the sona-system. Faculty who wish to join are able to do so, but the primary participant pool is university students.

For all items (test and control), participants were asked how well they understood each on a scale from 1 = I don't understand this statement at all to 10 = I completely understand this statement. Participants were first presented with three control items in a random order: 'A circle is a type of shape,' 'A circle is a type of animal,' and 'A spherical pain dissertates ubiquitously.' Then, they were presented with the test items also in a random order. For each test item, we also asked, 'How could we make this statement easier to understand?'

Based on responses (quantitative and qualitative), the authors then adjusted the test items. After which, another 30 participants were recruited using the sona-system to

complete the adjusted survey and entered into a lottery draw for an Amazon voucher upon completion. As expected, these participants indicated higher understanding for the first two control items (*Modes* = 10; *Medians* = 10) than they did for the third (*Mo* = 1, *Mdn* = 2). Further, and meeting our standards for understanding, these participants indicated high understanding of all the test items (*Mo* = 10; *Mdn* range = 8–10). We thus considered the items ready for the main survey.

As should be clear from Table 1, the items we decided to use, partly as a result of the aforementioned process, differed in various respects from those as stated by Johnston. We rely mostly on the face validity of the items as evidence that they retain a high degree of content accuracy. However, we will explicitly touch on two notable changes between our items and Johnston's. First, the notion of 'paradigms' is not explicitly mentioned in our Paradigm items. Johnston defined Paradigms thus: 'Some of what we take to be paradigms of canary yellow things (e.g. some canaries) are canary yellow.' This seems to be true if and only if some canaries are canary yellow, some oranges are orange, some pumpkins are auburn orange, some grass is green, and so on. Thus, we tested Paradigms by testing these individual statements. We did not explicitly bring up the notion of 'paradigms,' because we think that what is philosophically most interesting is whether people believe that some of the objects in question have the colors that we often associate with them. Whether people think they are *paradigms* of the relevant colors is another question entirely.

Second, we neither used the term 'nature' in the Unity items nor in any of the Revelation items. In the survey development, it became apparent that this term was hard for non-philosophers to understand. For the Unity items, we instead relied on the notion of necessity, conveyed with the term 'must' (e.g. red must be more similar to orange than it is to green). For the Revelation items, we instead relied on a few different terms for which participants reported higher understanding. Revelation is the most obscure of Johnston's core beliefs, with disagreement regarding its correct expression (e.g. see Roberts 2018). As such, there is no definitive statement of this core belief, and so we only intended to capture Revelation as a somewhat amorphous and imprecise notion. Talk of 'complete understanding,' used in the first Revelation item, comes from Gert's version of Revelation (2008). The idea of 'red's intrinsic features,' used in the second Revelation item, is prima facie very close to the idea of 'red's intrinsic nature,' used by Johnston. The last item, 'When we look at a red object, we perceive the characteristic of the object that is red,' was used to capture a central aspect of Revelation simply: that the metaphysical identity of red is available to perception. This formulation is clearly perceptual and similar to that in Roberts (2018): 'The colors as we see them in perception are the intrinsic natures of the colors' (i.e. what they are metaphysically). The results for each item are presented in Appendix A and so one can exclude items at one's discretion if desired.

Participants In total, 333 participants completed the main survey via the University of Warwick's sona-system. This study was ethically approved by the Humanities and Social Sciences Ethics Committee at the University of Warwick (REF: 20/15–16:DR@W). Participants were entered into a new lottery draw for a chance to win an Amazon gift voucher. The participants did not include any that were used in the survey's development. We continued recruitment until 300 participants met our inclusion criteria. The 33 additional participants were removed for the following reasons: 16 had philosophical training, and 17 did not respond appropriately to the control items

Table 1 Items (test and control) for Experiment 1

Paradigms

The program randomly selects one item from each of the three columns below.

P1	P2	P3
- Some tomatoes are red.	- Some pumpkins are auburn orange.	- Some grass is green
- Some lipstick is scarlet red.	- Some bananas are yellow.	- Some trees' leaves are forest green.
- Some oranges are orange.	- Some canaries are canary yellow.	- Sometimes the sky is blue.
		- Some sapphires are sapphire blue.

Explanation

The program randomly selects one of ten colors to complete the sentence.

E1. A liquid being ___ at least sometimes causally explains why it looks ___.

E2. A surface being ___ at least sometimes causally explains why it looks ___.

E3. A light bulb being ___ at least sometimes causally explains why it looks ___.

Unity

The program randomly selects one of the ten colors along with its color complements to complete the sentence.

U1. ___ must be more similar to ___ than it is to ___.

U2. ___ must be more different from ___ than it is from ___.

U3. If something is entirely ___, then it cannot at the same time be entirely ___

Availability

The program randomly selects one of the ten colors to complete the sentence.

A1. When one looks at a(n) ___ car, one is normally justified in believing that the car is ___.

A2. One's belief that something is ___ can normally be justified without scientific investigation.

A3. Normally, we have good reason to believe that a coffee mug is ___ based just on seeing the coffee mug.

Revelation

The program randomly selects one of the ten colors to complete the sentence.

R1. A complete understanding of what ___ is can be had by seeing a(n) ___ thing, like a(n) ___ dress.

R2. When one sees a(n) ___ object, one sees ___'s intrinsic features.

R3. When we look at a(n) ___ object, we perceive the characteristic of the object that is ___

Control items

C1. A circle is a type of shape.

C2. A circle is a type of animal.

(see the section, 'Procedure') (5 for control item 1, and 12 for control item 2). Of the remaining 300 participants, 200 were Female and the median age was 20 years (IQR = 19–22). We also collected details about the native language of participants. 164 identified as native English speakers. Of the remaining 136, 32 had a native language from a country in the European Union, 101 from a non-European Union country, and 3 did not say.

That these participants were all recruited through Warwick's sona-system is a limitation. The sample mainly included young adults who were students of one British University. Experiment 2 uses a different demographic pool by recruiting participants via United States based Amazon Mechanical Turk and includes a wider range of ages (see Sect. 3). Future research could include a more diverse range of peoples and cultures than either of our two studies (see Sects. 3 and 4.). Although Experiment 1 did include participants with a range of native languages, they were not selected in their home countries, and those who chose to attend a British University may not be representative of their home populations.

Materials The survey was created and administered online using Qualtrics software, version 2015. The survey typically took between 5 and 6 min to complete (*Mdn* = 328 s, *IQR* = 249–445). All analyses were performed using IBM SPSS version 22.

Procedure Participants were asked how much they agreed with 3 test items for each core belief on a 10-point Likert scale with only end points indicated (1 = strongly disagree and 10 = strongly agree). To ensure that specific colors would not confound the experiment, 10 versions of each test item were created, using these fine and coarse-grain colors: red, orange, yellow, green, blue, scarlet red, auburn orange, canary yellow, forest green, and sapphire blue. (We did not choose colors that were similar in hue, because it was important that participants differentiate between the colors chosen: if they did not, this might affect their responses. So, as a first test of Unity, we chose to look at colors that were not close in color space; if we did not find Unity for these colors, it is unlikely we would find it for more similar colors.) For Paradigms, 10 separate items were created and divided into 3 groups, and the survey software randomly selected one item from each group to show each participant (see Table 1). For the other core beliefs, 10 versions of each item were created, and the survey software randomly selected one version of each item to show each participant (Table 1).

Every participant responded to two control items using the same Likert scale as before (see Table 1). The purpose of these control items was to remove those participants who did not pay enough attention, were unable to do/understand the task, or demonstrated a high acquiescence bias (agreed no matter the content). The first control item read, 'A circle is a type of shape.' Participants who were paying attention and were able to do/understand the task should agree with this item. Hence, those who responded 5 or below were removed from the analyses. The second control item read, 'A circle is a type of animal.' Unbiased participants who were paying attention and were able to do/understand the task should disagree with this item. Thus, those participants who responded 6 or above were removed from the analyses. We chose these cut-points for this reason: as we used a 10-point Likert scale, 5 and below can be taken to indicate some degree of disagreement and 6 and above can be taken to indicate some degree of agreement. These cut-points are the same cut-points we use in our results section to designate whether participants indicate some degree of agreement with each core-belief. After indicating how much they agreed with all the items in Table 1, demographic questions were asked.[5]

[5] We collected the following demographics: Gender, Age, Experience in Academic Philosophy (None, Current undergraduate, Graduated undergraduate, Current postgraduate, Graduated postgraduate) and Native Language (Free-text).

2.2 Results

Comparing Core Beliefs Composite scores for each core belief (an average of the three items) were compared using descriptive statistics and tests. As there are no a priori hypotheses to guide the analyses in this article, the significance level was set to a conservative 0.005 (rather than the conventional 0.05).[6] To allow readers who wish to more liberally interpret differences to do so, exact *p*-values are given where results are not deemed significant.

Descriptively, participants expressed the strongest agreement with Paradigms (*Mdn* = 9.0, with 94% expressing an agreement of 6 or above). This was followed by Unity (*Mdn* = 8.0; 87%), Availability (*Mdn* = 7.3; 82%), Revelation (*Mdn* = 7.0; 78%), and finally Explanation (*Mdn* = 6.7; 66%) (see Fig. 1). The medians for all items and the percent who answered 6 or higher for the 300 participants are given in Appendix A.

Friedman's test, a non-parametric alternative to the repeated measures ANOVA, was used to compare the composite scores for each core belief. This test was significant, $\chi^2(4) = 262.70$, $p < 0.001$. Post hoc analyses were conducted to see where the differences were. The difference between most core beliefs was significant (*p*'s < 0.001) but not the difference between Availability and Revelation ($p = 0.006$), and Explanation and Revelation ($p = 0.03$). Recall that Revelation especially favors a primitivist view and Explanation is thought to be an issue for primitivism. Lower-level physical views struggle with Availability and Unity.

Fine Versus Coarse-Grain Colors To compare participants' responses to fine and coarse-grain colors, Mann-whitney U tests were performed for each of the 15 items. For completeness, the *p* value of each test is given here in the order of the items in Table 1 (e.g. P1, P2, then P3).

Paradigm: *p*'s = 0.54, <0.001, and 0.01.
Explaination: *p*'s = 0.63, 0.47, and 0.37.
Unity: *p*'s = 0.11, 0.50, and 0.57.
Avalibility: *p*'s = 0.01, 0.87, and 0.14.
Revelation: *p*'s = 0.11, 0.37, and 0.72.

The only item's *p* value that fell below 0.005 was Paradigms 2 ($Z = 6.24$, $p < 0.001$). For this item, more people agreed with the coarse-grain version (involving yellow) than the with the fine-grain versions (involving auburn orange and canary yellow) (*Mdn's* = 10 v. 8). Thus, while agreement is *very* high for all the Paradigm items, it might be that folk Paradigm beliefs regarding some coarse-grain colors are higher than for some fine-grain colors. However, the reader should not make much of this finding, as it was not stable for Paradigms generally.

[6] The *p* value needed for significance is often adjusted when multiple tests are performed with no a priori hypothesis(es). This is done to reduce the chance of a type I error. However, by decreasing the chance of a type 1 error, one also increases the chance of a type 2 error. Commonly, levels of significance are adjusted according to the Bonferroni correction (i.e., 0.05/[the number of tests]), but when the number of tests is high this correction can be overly conservative. So, as we ran many tests, the Bonferroni correction was avoided; rather, the *p* value needed to obtain significance was set universally to a conservative 0.005 (Armstrong 2014).

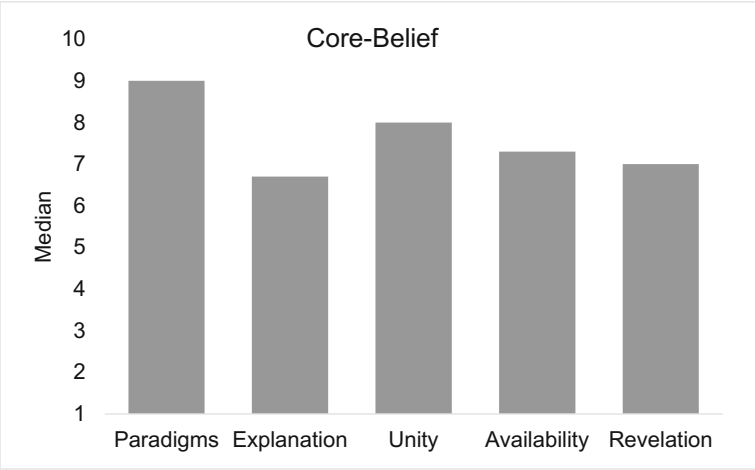

Fig. 1 Experiment 1 medians of the composite scores for each core belief

3 Experiment 2

Experiment 2 was conducted to address four arguable weaknesses of Experiment 1. These weaknesses involve the participant population and the survey design. We now discuss these weaknesses and how they were addressed. We then will look at the methods and results.

Participant Population The participants in Experiment 1 were recruited over the University of Warwick's sona-system. Although our sample included a range of native languages, it was also comprised of mostly young adults attending the University of Warwick in the United Kingdom. Whether the results generalize to different populations is an open question. Experiment 2 uses a different demographic pool by recruiting participants via Amazon Mechanical Turk (a US based service with residents from all over the country) and includes a much wider range of ages. Future research could look at a more diverse range of peoples and cultures (see 'three limitations' in Sect. 4).

Survey Design: Unity Items An arguable weakness of Experiment 1's survey design is that the simple Unity 1 and 2 items used in it were, because of their simplicity, ambiguous as to whether they were about the colors' hue, saturation, or brightness. For example, a Unity 1 item read merely, 'red must be more similar to orange than it is to green.' Participants might have disagreed with this to express that red need not be more similar in every respect. To address this weakness, the original Unity 1 and 2 items were altered for Experiment 2 by placing the following clause in front of them: 'With respect to its hue (as opposed to how saturated or bright it is).' The remaining items from Experiment 1 were unaltered for Experiment 2.

Survey Design: Likert Scale Another arguable weakness of Experiment 1's survey design is that the Likert scale response options in it were only labeled at the end points: 1 = strongly disagree and 10 = strongly agree. As a result, it may not have been obvious to participants that a response of 5 indicates disagreement and a response of 6 indicates

agreement. To address this arguable weakness, in Experiment 2 every Likert point was labeled: 1 = completely disagree, 2 = strongly disagree, 3 = disagree, 4 = moderately disagree, 5 = slightly disagree, 6 = slightly agree, 7 = moderately agree, 8 = agree, 9 = strongly agree, and 10 = completely agree. Thus, in Experiment 2, unlike in Experiment 1, it should be obvious to participants that a response of 5 indicates some level of disagreement, and that a response of 6 indicates some level of agreement.

Survey Design: Acquiescence Bias A final, and perhaps the most significant, potential weakness of Experiment 1's survey design is that participants' agreement may have been tainted by an acquiescence bias: an effect wherein participants tend to agree with any item regardless of its content (Podsakoff et al. 2003). While the attention check item, 'a circle is a type of animal,' was designed to remove those who displayed a strong acquiescence bias, this item was so obvious it might not have tapped into the acquiescence bias effectively. To address this, in Experiment 2, 15 new test items were created that were logical negations of the original 15 (Swain et al. 2008). For example, in addition to the original item, 'Some tomatoes are red,' Experiment 2 included its logical negation: 'No tomatoes are red.' Participants who agree with an original item and who disagree to the same degree with its negation have not displayed an acquiescence bias. Participants in Experiment 2, responded to both the 15 original and the new 15 negation items.

Negation items were not used in Experiment 1, because past research suggests that negatively worded items may confuse participants and so make surveys less reliable (Sonderen et al. 2013). We thought this a good reason to avoid the complication. However, other researchers believe that the careful inclusion of negation items enhances participants' attention to and comprehension of all the items (Weijters and Baumgartner 2012). As Experiment 1 already provided data without negation items, we carefully included them in Experiment 2. Thus, the reader can see the results of two similar surveys, one with and one without negation items.

3.1 Method

Participants In total, 327 participants were recruited and completed the survey via Amazon M(echanical) Turk. This study was ethically approved by the Humanities and Social Sciences Ethics Committee at the University of Warwick (REF: 46/16–17:DR@W). As is common practice when using M Turk, participants were at least required to have a US high school diploma and an assignment approval rate of 95%. Participants who passed the control items and entered the correct code (given to them at the end of the survey) into M Turk received a 0.50 USD payment. Twenty-seven participants' responses were removed for the following two reasons: 20 had philosophical training, and 7 did not respond appropriately to the control items (2 for control item 1, and 5 for control item 2). Of the remaining 300 participants, 151 were female and the median age was 35 years (IQR = 29 to 54). As Amazon M Turk requires a US bank account, almost all participants were native English speakers (295 of the 300). So, Experiment 1 had a demographic advantage of including a more diverse range of native languages, but Experiment 2 had an advantage in having a more diverse range of ages. The participants gathered from M Turk were obviously also not all from one British University.

Materials The survey was created and administered online using Qualtrics software, version 2015. It typically took participants between 5 and 10 min to complete (*Mdn* = 439 s, *IQR* = 332.5–602.5). All analyses were performed using IBM SPSS version 22.

Procedure Participants were asked how much they agree with the 15 original and 15 negation items, using a 10-point Likert scale where each point was designated with a label from completely disagree to completely agree. As in Experiment 1, the order of the items was randomized. All the negation items used in Experiment 2 are in Table 2.

In addition to the original and negation items, every participant responded to the same two control items as in Experiment 1, using the new response options. If a participant answered one of these control items incorrectly, they were immediately sent to the end of the survey and did not receive payment. Participants were told that payment depended on answering the control items correctly on the consent page. After indicating how much they agreed with all the items, participants were asked the same demographic questions as in Experiment 1.

3.2 Results

Comparing Core Beliefs Composite scores for each core belief (the average of the three original items and the three negation items reverse-scored, so, for example, a 1 reversed scored would become a 10) were compared using descriptive statistics and tests.

Descriptively, participants expressed the strongest agreement with Paradigms (*Mdn* = 9.7, with 100% expressing an agreement of 6 or above). This was followed by Availability (*Mdn* = 8.2; 91%), Unity (*Mdn* = 7.3; 90%), Revelation (*Mdn* = 7.2; 79%), and finally Explanation (*Mdn* = 6.3; 59%) (see Fig. 2). Notably, compared to Experiment 1, the order of Unity and Availability switched. This could be due to the population, adjustments to the unity items, adjustments to the Likert scale, and/or due to the inclusion of negation items. We do not have the means to parse this out. The reader should bear in mind that agreement for both Unity and Availability was high for both Experiment 1 and Experiment 2. The medians for all items and the percent who answered 6 or higher for the 300 participants is given in Appendix A (without reverse scoring).

Friedman's test was used to compare the composite scores for each core belief. This test was significant, $\chi^2(4) = 573.93$, $p < 0.001$. Post hoc analyses were conducted to see where the differences lay. The differences between all core beliefs were significant (p's < 0.003). This contrasts with Experiment 1, where Revelation and Availability were not significantly different nor Revelation and Explanation. We prefer not to speculate as to why. The differences could be due to some of or all of the differences between Experiment 1 and Experiment 2.

It is good to remind oneself when considering these results of the general relationships that the core beliefs have with the different views on color. Paradigms is problematic for irrealist views according to which external objects are not colored. Explanation is considered problematic for higher-level views (appearance dispositionalism and primitivism). Availability and Unity are considered problematic for lower-level physical views (micro-structuralism and reflectance physicalism). Revelation is the core belief particularly favorable for primitivism. We will look at these relationships in more depth in the discussion and just mention them here for ease of reference.

Table 2 Negation items for experiment 2

Negation Paradigms		
The program randomly selects one item from each of the three columns below.		
P1.neg	P2.neg	P3.neg
- No tomatoes are red.	- No pumpkins are auburn orange.	- No grass is green
- No lipstick is scarlet red.	- No bananas are yellow.	- No trees' leaves are forest green.
- No oranges are orange.	- No canaries are canary yellow.	- The sky is never blue.
		- No sapphires are sapphire blue.

Negation Explanation

The program randomly selects one of ten colors to complete the sentence.

E1.neg A liquid being ___ never causally explains why it looks ___.

E2.neg A surface being ___ never causally explains why it looks ___.

E3.neg A light bulb being ___ never causally explains why it looks ___.

Negation Unity

The program randomly selects one of the ten colors along with its color complements to complete the sentence.

U1.neg With respect to its hue (as opposed to how saturated or bright it is), ___ does not have to be more similar to ___ than it is to ___.

U2.neg With respect to its hue (as opposed to how saturated or bright it is), ___ does not have to be more different from ___ than it is from ___.

U3.neg If something is entirely ___, then it can at the same time be entirely ___.

Negation Availability

The program randomly selects one of the ten colors to complete the sentence.

A1.neg When one looks at a ___ car, one is not normally justified in believing that the car is ___.

A2.neg One's belief that something is ___ cannot normally be justified without scientific investigation.

A3.neg We do not normally have good reason to believe that a coffee mug is ___ based just on seeing the coffee mug.

Negation Revelation

The program randomly selects one of the ten colors to complete the sentence.

R1.neg A complete understanding of what ___ is cannot be had by seeing a ___ thing, like a ___ dress.

R2.neg When one sees a(n) ___ object, one does not see ___'s intrinsic features.

R3.neg When we look at a(n) ___ object, we do not perceive the characteristic of the object that is ___.

Fine Versus Coarse-Grain Colors To compare participants' responses to fine and coarse-grain colors, Mann-whitney U tests were performed for each of the 30 items. For completeness, the p value of each test is given here for the original and then the negation items in order (e.g. for the Paradigm items, the p-values are given for P1, P2, P3, then P1.neg, P2.neg, and P3.neg).

Paradigm: p's = 0.07, 0.001, <0.001, 0.07, <0.001, and 0.001.
Explanation: p's = 0.43, 0.47, and 0.69, 0.93, 0.55, and 0.31.

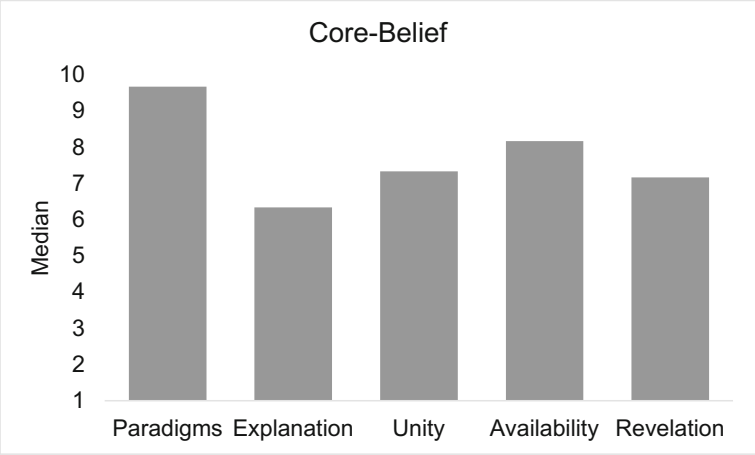

Fig. 2 Experiment 2 medians of the composite scores for each core belief

Unity: p's = 0.07, 0.28, 0.22, 0.23, 0.24, and 0.25.
Availability: p's = 0.001, 0.02, 0.88, 0.30, 0.87, and 0.68.
Revelation: p's = 0.26, 0.07, 0.57, 0.18, 0.37, and 0.98.

As with Experiment 1, differences between fine and coarse-grain colors emerged for the second Paradigm item: yellow ($Mdn = 10$) was rated higher than auburn orange and canary yellow ($Mdn = 8$; P2: $Z > 3.47$, $p = 0.001$). A similar difference emerged for this item's negation: negation yellow ($Mdn = 1$) was rated lower than negation auburn orange and canary yellow ($Mdn = 1$; P2.neg: $Z > 6.45$, $p < 0.001$). An additional difference was found for the third Paradigm item: green and blue ($Mdn = 10$) were rated higher than forest green and saphire blue ($Mdn = 10$; P3: $Z > 3.72$, $p < 0.001$). A similar difference emerged for this item's negation: negation green and blue ($Mdn = 1$) were rated lower than negation forest green and sapphire blue ($Mdn = 1$; P3.neg: $Z > 3.34$, $p = 0.001$). Ceiling and floor effects are likely driving these differences.[7]

The only other item for which a difference between fine and coarse-grain colors was found was the first, original Availability item (A1: $Z = 3.43$, $p = 0.001$) with the median score for fine-grain being 9, and the median score for course-grain being 10. Why participants would treat fine and coarse-grain colors differnetly for just one Availability item is curious; however, as this result does not hold true for both Experiments 1 and 2 nor for the relevant negation item, we do not want to overinterpret the result.

4 Discussion

Summary of General Results Our results for both Experiments 1 and 2 suggest that people generally agree with all five of Johnston's core beliefs but, importantly, with some

[7] A ceiling (or floor effect) occurs in surveys when most participants respond near the top (or bottom) of the scale for a particular item. In so doing, the variability around that item is very low and atypical outliers have an opportunity to cause differences to appear that would have otherwise not appeared.

more than others. Looking at Experiment 2, which is the more rigorous, we see that participants expressed the strongest agreement with Paradigms, followed by Availability, Unity, Revelation, and finally Explanation. A notable difference between Experiment 1 and 2 is that for the former Unity was agreed with more than Availability. We cannot speculate much as to why. Future research could investigate if the difference was due to the different participant pools or something else by keeping everything the same between two numerically different experiments run with the different pools. Possibly, the new Unity items were agreed with less than Availability, because they were less well understood. However, Availability agreement was notably higher in Experiment 2 (Exp. 1: *Mdn* = 7.3; 82% vs. Exp. 2: *Mdn* = 8.2; 91%), whereas Unity agreement was, in fact, rather similar (*Mdn* = 8.0; 87% vs. *Mdn* = 7.3; 90%).

More relevantly, Experiments 1 and 2 agree in more ways than they disagree, and so can be seen largely to be complimentary. Both Experiments 1 and 2 agree that Explanation is descriptively the least agreed with and Paradigms the most. Both Experiments 1 and 2 agree that Revelation is descriptively the second least agreed with but still generally agreed with. Both Experiments 1 and 2 agree that Unity and Availability are in the middle of the other core beliefs in terms of participant agreement. Statistically, Experiments 1 and 2 agree on a lot as well, for example, that Explanation is agreed with less than Unity and Availability, that Paradigms is agreed with more than Explanation, and that Unity is agreed with more than Revelation.

Three Limitations 1. Our experiments are based on the presupposition that if a proposition is a candidate for being a core belief, agreement should be high, and that the higher the agreement, the better a candidate, ceteris paribus.[8] There is a limitation to testing core beliefs in this way: high agreement alone does not entail that a proposition is a core belief. To avoid this limitation, an alternative method might use counterfactual scenarios. While such a method could yield interesting results, we chose not to use it, for it would have required more conceptually complicated items and so would decrease the number of items we could realistically expect a sizable number of non-philosophers to answer. We wanted to develop a survey that could comprehensively assess a sizable number of non-philosophers' responses with respect to all five of Johnston's core beliefs, so that we could compare participant agreement across them. We encourage researchers to employ many methods, including psycholinguistic methods (Fischer and Engelhardt 2016, 2017, 2019), because *the triangulation of data from different methods will certainly yield a richer understanding of folk core beliefs about color.*

2. Our experiments focused on participants in the United States and the United Kingdom. One area that experimental philosophy has investigated is the cross-cultural stability of philosophically relevant concepts (Weinberg et al. 2001; Nagel et al. 2013; Machery et al. 2004). If people's concepts diverge between cultures that might call into question the methods of philosophy. Specifically, it might call into question whether we can say what the necessary and sufficient conditions are for a philosophically relevant concept. If they are stable, that might reinforce the value of traditional philosophical methods. No cross-cultural

[8] We are hesitant to think that everything is equal between Paradigms and the four other proposed core beliefs. There is a special concern with Paradigms, for with this proposed core belief in particular we seem to be able to conceive of it being false, but perhaps this is too fast. There are worries about whether irrealism is really coherent. Stroud (2000) says, 'it [is] a necessary condition of our acknowledging the presence in the world of perceptions of and beliefs about the colors of things that we believe that some objects are colored' (p. 157).

research into philosophical color concepts is known to us. Do people the world over agree on color core beliefs? Future research could translate our items into different languages and test them in non-English speaking countries. We encourage that this work be done but cannot do it here.

3. Our method does not allow us to determine whether participants understand the items in the way philosophers would. This said, it should be recognized that philosophers do not agree on precise ways to understand the core beliefs. For example, regarding Explanation, what is the nature of causal explanation? Does Hume's Dictum that the relata must be distinct apply (Allen 2016, p. 100–101)? Regarding Revelation, what exactly is it to be an intrinsic nature? Is it just the intrinsic and necessary properties of something, or is it something stronger, the *essential* properties of something, where this requires a more fine-grained notion than necessity (Fine 1994)? Roberts (2018) uses 'intrinsic nature' to mean what something is most fundamentally. We could go on, but the point should be clear: there is no one precise understanding available. This is why we chose to approach participant understanding in an inclusive fashion: if they said they understood the items, we accepted that as a minimally sufficient condition. The present experiment provides a starting place for future experiments that may take a more stringent approach. For example, in-depth interviews could be conducted with fewer participants and then thematically analyzed to determine precisely how participants understand each item.[9]

Relevance of Results to the Ontology of Color Our goal in this section is not to do much heavy philosophical work but to more survey plausible consequences our results could have for the ontology of color given Johnston's meta-methodology of how to adjudicate between competing ontological views on color. We focus on Experiment 2 for simplicity and because it is the more rigorous of the two (remember that the studies largely agreed with each other).

As Paradigms is ranked the highest, it is (ceteris paribus) the best candidate for being a core belief.[10] If Paradigms is a core belief, the colors can only not be instantiated by objects more or less inclusively speaking. Our finding that people are generally realist about color is consistent with Roberts and Schmidtke's (2016) study. They asked participants to read a vignette involving a disagreement about the color of an object and then indicate their agreement with (among other things) the following: 'the object really has a color (or colors).' They found very high agreement with this item with a mean of 8.6 on a 10-point likert scale. Importantly, notice their use of the word 'really,' which presumably tracks realism even tighter. An empirically backed argument from intuition/common sense is available for color realism given the two studies mentioned. Intuition/common sense should not be overturned without significant reason. We have empirical evidence that realism is the intuitive/common sense view. So, realism should be assumed the default position only to be overturned in light of significant arguments to the contrary (for discussion of realism/irrealism see Hardin 1993,

[9] We avoid engaging with meta-philosophical concerns about the value of intuitions in philosophy and/or metaphysics (for this debate see Ladyman and Ross 2007, Dorr 2010, and Maclaurin and Dyke 2012).

[10] The vast majority by far agreed with the Paradigm items but not everyone. Potentially some disagreed because they took 'some' in, for example, 'Some tomatoes are red' to mean 'only some.' Logically 'some' does not mean 'only some,' but of course this does not mean that some participants did not read the item this way. As some may have disagreed for this reason, we cannot assume that those who disagreed endorse irrealism.

2004; Mclaughlin 2003; Chalmers 2006; and Byrne and Hilbert 2007a). More generally, and as previously mentioned in the introduction, there is no reason this paper cannot be viewed as providing an empirical basis for an argument from common sense or intuition for views which better accommodate the content of Johnston's core beliefs.

Explanation is the weakest core belief, according to our results, with only 59% agreeing, compared to Revelation with 79% agreeing.[11] Thus, there is reason to think Johnston was wrong to prefer Explanation over Revelation (notice that this would also hold if Revelation and Explanation were statistically the same, as was the case with Experiment 1). So, if core beliefs are to adjudicate between views on color, perhaps philosophers should consider giving up the idea that colors cause our color experiences, in favor of a primitivist view whereby the physical realizers of primitive color properties cause our experiences.[12] More generally, our results suggest that overdetermination arguments against higher-level views, such as primitivist and appearance dispositionalist views, may not be as strong as thought. Our data also supports the claim that Unity is a significantly stronger candidate (90% agreeing) for a core belief than Explanation (ceteris paribus). Lower-level physical views, such as micro-structuralism and reflectance physicalism, are known to struggle with Unity more than higher-level views (Johnston 1992; Hardin 1993, p. 66; Maund 1995, p. 126–133; Thompson 1995, p. 124; Pautz 2006). On the other hand, higher-level views are known to struggle more with Explanation than lower-level views (Johnston 1992; Hardin 1993, p. 61; Kim 1993a, 1993b; Tye 2000, p. 148). Which type of view is better? Our results suggest that we should prefer Unity over Explanation, and hence that higher-level views, such as appearance dispositionalism and primitivism, are preferable (ceteris paribus) to lower-level physical views, such as micro-structuralism and reflectance physicalism.

Availability is the second strongest core belief, according to our results, with 91% agreeing. This is reason to think that Johnston (1992) was right to worry about lower-level views, such as micro-structuralism and reflectance physicalism, accommodating Availability. As we said, Johnston's argument runs thus: if the colors were physical properties which accommodate Unity, we could only know that there were colors if we knew that there were Unity accommodating physical properties, but we cannot know this without a scientific investigation (pp. 238–238). Neither rejecting Unity nor Availability is a good strategy to avoid this argument given our results for both are highly agreed with. As we said, Availability alone poses a problem for lower-level physical views on color: such properties do not have the property of being such that justified belief about them is available simply based on visual perception, but then by Leibniz's law neither do the colors, if the colors are physical properties (Jackson 2012). Of course, this argument is going to be controversial. We are just surveying the types of issues Availability raises. Further discussion is better left for a separate article.

Unity is the third strongest core belief after Availability, according to our results, with 90% agreeing. So, a plausible case can be made that any view which can be said to

[11] This might be seen as consistent with Roberts et al.'s (2016) finding to the effect that the causal condition on perception is not a conceptual truth. They found that for certain cases (non-blocker cases) many lay people did not seem to think that it was required to perceive something that the thing be a cause of the perception.

[12] Johnston (1992, note 12) worries that Explanation is a prerequisite for the colors being visible properties. In reply, empty space neither reflects nor emits light, but it seems we do visually represent it. So, more needs to be said for Johnston's worry to be convincing. This is not something we can discuss further in this paper.

be *a view on color* must hold onto Unity. As we said, the lower-level views micro-structuralism and reflectance physicalism are known to struggle with Unity. Thus, there is good reason in particular to be concerned about these views. It seems plausible that these views cannot just bite the bullet and reject Unity if they are to remain *views on color*. Many proponents of micro-structuralism and reflectance physicalism, of course, understand the importance of Unity and try to accommodate it in various ways (Pautz 2006). However, some views purposefully reject Unity. For example, Cohen's relationalism (Cohen 2004, 2009) is purposefully designed so that everyone is right in variation cases (cases where there is perceptual disagreement about the color of an object).[13] Unity includes in it that certain colors exclude each other: if something is entirely blue, then it cannot at the same time be entirely yellow. In fact, it is this element of Unity about which participants indicated the strongest agreement (see Appendix A). Thus, our results pose a problem for Cohen's relationalism: as it rejects Unity, in fact, the strongest element of Unity, can it be said to even be *a view on color*? Perhaps not.[14]

Revelation is the most contentious of the core beliefs proposed by Johnston with many different interpretations of it discussed (Byrne and Hilbert 2007b; Kalderon 2007; Gert 2008; Campbell 2005; Allen 2011; Roberts 2018). Our items were just meant to capture Revelation as a somewhat amorphous and imprecise notion. Future work could attempt to empirically test the separate versions of revelation proposed in the literature (Roberts 2018). Revelation, amorphously understood, is not, given our results, a particularly strong core belief comparatively speaking: taking the composite of the 6 Experiment 2 items, it is the weakest candidate besides Explanation. This said, our results are that a large majority of people (79%) agree with Revelation. Looking at the individual items, we find that for every item the majority of people agree with it (see Appendix A) and so that people agree is a robust finding that is not easily dismissed. Benbaji (2016) rightly discounted Revelation as being question begging. At that time, there was no independent motivation for Revelation but only the opinions of Benbaji's opposition. Our results provide some independent motivation. Hence, future work should engage with Revelation more seriously and cannot simply discount it as question begging.

Acknowledgements We would like to thank Nat Hansen for raising the acquiescence bias concern with the first experiment. Keith Allen also provided useful feedback in particular regarding the discussion. An earlier version of this article was presented at the Seventh Conference of the Experimental Philosophy Group at the University of Reading in the UK. We would like to thank the organizers for a chance to present this work and the conference participants for useful feedback. Finally, we are grateful to an anonymous referee for helpful comments.

[13] We acknowledge that there is an added complication here that arises due to Cohen's (2009) fine-grain/coarse-grain distinction. Coarse-grain colors are supposed to accommodate Unity to some degree. Other papers have already engaged with this distinction (e.g. see Roberts 2014), and so we will not do so here.

[14] Our results for Unity are inconsistent with Cohen and Nichols' (2010) results. They find that about half of their participants support a relationalist view of color according to which everyone is right in variation cases. However, Cohen and Nichols' results have been seriously called into question by subsequent studies by Roberts, Andow and Schmidtke (2014), and Roberts and Schmidtke (2016). Our results are consistent with these newer studies. (For some interesting meta-discussion of all these studies, see Adams 2016, p. 13–18; Adams and Hansen Forthcoming.)

Appendix A

Table 3 The median agreement indicated for each item and the percent who answered 6 or higher for each item

Experiment 1			Experiment 2					
Item	Mdn	% 6 or above	Item	Mdn	% 6 or above	Item	Mdn	% 6 or above
P1	10	91.7%	P1	10	99%	P1.neg	1	1%
P2	9	90.0%	P2	10	98%	P2.neg	1	1%
P3	10	91.7%	P3	10	97%	P3.neg	1	1%
E1	7	68.3%	E1	7	68%	E1.neg	5	47%
E2	7	79.7%	E2	8	74%	E2.neg	5	43%
E3	6.5	63.0%	E3	7	69%	E3.neg	5	42%
U1	8	82.3%	U1	8	84%	U1.neg	4	34%
U2	8	72.7%	U2	8	76%	U2.neg	4	33%
U3	10	88.7%	U3	10	98%	U3.neg	1	13%
A1	9	92.7%	A1	9	98%	A1.neg	2	14%
A2	8	73.7%	A2	8	87%	A2.neg	3	13%
A3	8	69.7%	A3	8	86%	A3.neg	3	20%
R1	7.5	72.0%	R1	7	72%	R1.neg	5	40%
R2	7	67.0%	R2	7	78%	R2.neg	4	26%
R3	8	84.7%	R3	8	94%	R3.neg	3	14%

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Adams, Z. 2016. *On the genealogy of color: A case study in historicized conceptual analysis*. Routledge.

Adams, Z. & Hansen, N. (Forthcoming). The myth of the common sense conception of color. In Å. Wikforss and T. Marques, eds., Shifting Concepts: The Philosophy and Psychology of Conceptual Variability. *Oxford University Press.*

Allen, K. 2011. Revelation and the nature of color. *Dialectica 65*: 153–176.

Allen, K. 2016. *The naive realist theory of color*. Oxford University Press.

Armstrong, R.A. 2014. When to use the Bonferroni correction. *Ophthalmic & Physiological Optics 34*: 502–508.

Benbaji, H. 2016. Why color primitivism? *Australasian Journal of Philosophy 94*: 243–265.

Boghossian, P.A., and J.D. Velleman. 1989. Color as a secondary quality. *Mind 98*: 81–103.

Byrne, A., and D.R. Hilbert. 2003. Color realism and color science. *Behavioral and Brain Sciences 26*: 3–21.

Byrne, A., and D.R. Hilbert. 2004. Hardin, Tye and color physicalism. *Journal of Philosophy 101*: 37–43.

Byrne, A., and D.R. Hilbert. 2007a. Truest blue. *Analysis 67*: 87–92.

Byrne, A., and D.R. Hilbert. 2007b. Color primitivism. *Erkenntnis 66*: 73–105.

Campbell, J. 2005. Transparency versus revelation in color perception. *Philosophical Topics 33*: 105–115.

Chalmers, D.J. 2006. Perception and the fall of Eden. In *Perceptual Experience*, ed. T. Gendler and J. Hawthorne. Oxford University Press.
Cohen, J. 2004. Color properties and color ascriptions: A relationalism manifesto. *Philosophical Review 113*: 451–506.
Cohen, J. 2009. *The red and the real*. Oxford University Press.
Cohen, J., and S. Nichols. 2010. Colors, color relationalism, and the deliverances of introspection. *Analysis 70*: 218–228.
Dorr, C. 2010. Review of James Ladyman and Don Ross, every thing must go: Metaphysics naturalized. *Notre Dame Philosophical Reviews 6*: 16.
Fine, K. 1994. Essence and modality. *Philosophical Perspectives 8*: 1–16.
Fischer, E., and P. Engelhardt. 2016. Intuitions' linguistic sources: Stereotypes, intuitions and illusions. *Mind & Language 31*: 67–103.
Fischer, E., and P. Engelhardt. 2017. Stereotypical inferences: Philosophical relevance and psycholinguistic toolkit. *Ratio 30*: 411–442.
Fischer, E. & Engelhardt, P. (2019). Eyes as Windows to Minds: Psycholinguistics for Experimental Philosophy. In E. Fischer and M. Curtis (eds.): *Methodological Advances in Experimental Philosophy* (pp. 43–100). London: Bloomsbury.
Gert, J. 2008. What colors could not be: An argument for color primitivism. *The Journal of Philosophy 105*: 128–155.
Goldman, A. (2007). Philosophical intuitions: Their target, their source, and their epistemic status *Grazer Philosophische Studien*, *74*, 1–26.
Hardin, C. L. (1993). *Color for Philosophers*: *Unweaving the rainbow* (expanded edition). Hackett.
Hardin, C.L. 2004. A green thought in a green shade. *Harvard Review of Philosophy 12*: 29–38.
Jackson, F. (2012). Leibniz's Law and The Philosophy of Mind. *Proceedings of the Aristotelian Society, 112,* 269–283.
Jackson, F., and R. Pargetter. 1987. An objectivist guide to objectivism about color. *Revue Internationale do Philosophie 41*: 127–147.
Johnston, M. 1992. How to speak of the colors. *Philosophical Studies 68*: 221–263.
Kalderon, M.E. 2007. Color pluralism. *The Philosophical Review 116*: 563–601.
Kim, J. 1993a. Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research 52*: 1–26.
Kim, J. 1993b. The non-reductivist's troubles with mental causation. In *Mental Causation*, ed. John Heil and Alfred R. Mele. Oxford University Press.
Kripke, S. 1972. *Naming and necessity.* Harvard University Press.
Ladyman, J., and D. Ross. 2007. *Every thing must go: Metaphysics naturalized*. Oxford University Press.
Machery, E., Mallon, R., Nichols, S. & Stich, S. P. (2004). Semantics, cross-cultural style *Cognition, 92*, 1–12.
Maclaurin, J., and H. Dyke. 2012. What is analytic metaphysics for? *Australasian Journal of Philosophy 90*: 291–306.
Maund, B. 1995. *Colors: Their nature and representation*. Cambridge University Press.
McGinn, C. 1983. *The Subjective View: Secondary Qualities and Indexical Thoughts*. Oxford University Press.
Mclaughlin, B. P. (2003). The place of color in nature. In R Mausfeld & D. Heyer, eds., *Color Perception: Mind and nature.* Oxford University Press.
Nagel, J., S.J. Valerie, and R.A. Mar. 2013. Lay Denial of Knowledge for Justified True Beliefs. *Cognition 129*: 652–661.
Nassau, K. 1983. *The physics and chemistry of color. The fifteen causes of color*: Wiley.
Nassau, K. 1997. The causes of color. In *The Philosophy of Color, Volume 1: Readings on Color*, ed. Alex Byrne and David Hilbert. MIT Press.
Pautz, A. 2006. Can the physicalist explain color structure in terms of color experience? *Australasian Journal of Philosophy 84*: 535–564.
Podsakoff, P.M., S.B. MacKenzie, J.Y. Lee, and N.P. Podsakoff. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology 88*: 879–903.
Roberts, P. 2014. Color relationalism, ordinary illusion, and color incompatibility. *Philosophia 9*: 1085–1097.
Roberts, P. (2018). Another look at color primitivism. *Synthese*, https://doi.org/10.1007/s11229-018-1821-9.
Roberts, P., J. Andow, and K. Schmidtke. 2014. Color relationalism and the real deliverances of introspection. *Erkenntnis 79:* 1173–1189.
Roberts, P., and K. Schmidtke. 2016. Relationalism about perceptible properties and the principle of charity. *Synthese 193*: 2779–2803.

Roberts, P., Allen, K. & Schmidtke, K. (2016). Folk intuitions about the causal theory of perception. *Ergo*: *An Open Access Journal of Philosophy, 3*.

Stroud, B. (2000). *The Quest for Reality: Subjectivism and the Metaphysics of Color*. Oxford University Press.

Sonderen, E., R. Sanderman, and J.C. Coyne. 2013. Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS One 8*. https://doi.org/10.1371/journal.pone.0068967.

Swain, S.D., D. Weathers, and R.W. Niedrich. 2008. Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research 45*: 116–131.

Thompson, E. 1995. *Color Vision: A study in cognitive science and philosophy of perception*. Routledge.

Tye, M. 2000. *Consciousness, color, and content*. MIT Press.

Weijters, B., and H. Baumgartner. 2012. Misresponse to reversed and negated items in surveys: A review. *Journal of Marketing Research 49*: 737–747.

Weinberg, J.M., S. Nichols, and S. Stich. 2001. *Normativity and Epistemic Intuitions Philosophical Topics 29*: 429–460.

Yablo, S. 1995. Singling out properties. *Philosophical Perspectives 9*: 477–450.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.