The Baltic International Yearbook of
Cognition, Logic and Communication

December 2014
pages 1-30

Volume 9: *Perception and Concepts*
DOI: 10.4148/biyclc.v9i0.1085

STEVEN GROSS
THITAPORN CHAISILPRUNGRAUNG
ELIZABETH KAPLAN
JORGE AURELIO MENENDEZ
JONATHAN FLOMBAUM
Johns Hopkins University

PROBLEMS FOR THE PURPORTED COGNITIVE PENETRATION OF PERCEPTUAL COLOR EXPERIENCE AND MACPHERSON'S PROPOSED MECHANISM

ABSTRACT: Fiona Macpherson (2012) argues that various experimental results provide strong evidence in favor of the cognitive penetration of perceptual color experience. Moreover, she proposes a mechanism for how such cognitive penetration occurs. We argue, first, that the results on which Macpherson relies do not provide strong grounds for her claim of cognitive penetrability; and, second, that, if the results do reflect cognitive penetrability, then time-course considerations raise worries for her proposed mechanism. We base our arguments in part on several of our own experiments, reported herein.

1. INTRODUCTION

Experimental results suggest that stored information relating shapes and colors can affect one's perceptual experience of color. For instance, Delk & Fillenbaum (1965) found that participants match heart-shaped reddish-orange objects to a redder background than they do square-shaped reddish-orange objects. Fiona Macpherson (2012) argues that this result and others provide strong evidence in favor of the cognitive penetration of perceptual color experience. Moreover, she proposes a mechanism for how such cognitive penetration occurs.

Cognitive penetrability has been of great interest to philosophers for a variety of reasons. Philosophers share psychologists' interest in the relation of thought and perception and in cognitive architecture generally. But they concern themselves as well with the possible upshot of cognitive penetrability for questions regarding, for example, perceptual justification and the objectivity of scientific theorizing (reviewed in Stokes 2013). In addition, one of the studies Macpherson cites — concerning perception and race (Levin & Banaji 2006, discussed below) — is relevant to issues in social cognition and normative social theory of more general significance.

In what follows, however, we argue, first, that the results on which Macpherson relies do not provide strong grounds for her claim of cognitive penetrability; and, second, that, if the results *do* reflect cognitive penetrability, then time-course considerations raise worries for her proposed mechanism.

2. BACKGROUND

The term 'cognitive penetrability' was coined by Zenon Pylyshyn, who characterizes the phenomenon as one where "the function [a system] computes is sensitive, in a semantically coherent way, to the organism's goals and beliefs." (Pylyshyn 1999, p. 343) There has been some debate concerning the most fruitful way to further gloss, or reframe, what Pylyshyn puts his finger on. (Stokes 2013; Machery forthcoming) We can work with Macpherson's own characterization, applied by her in particular to perceptual experience:

> . . . perceptual experience is cognitively impenetrable if it is not possible for two subjects (or one subject at different times) to have two different experiences on account of a difference in their cognitive systems which makes this difference intelligible when certain facts about the case are held fixed, namely, the nature of the [effect of the] proxi-

> mal stimulus on the sensory organ, the state of the sensory organ, and the location of attentional focus of the subject. (Macpherson 2012, p. 29)

Cognitive systems — which would be the *source* of the penetration — are those responsible for such states as beliefs, desires, preferences, and intentions. Note that cognition as *opposed* to perception is meant. Of course, 'cognitive' is also often used in a way that *includes* perception, as in the phrase 'cognitive science.' But the question is not whether some system influences itself. (One might have spoken instead of higher cognition or of conception.)[1] The *target* of penetration would be perceptual experience, where 'experience' here indicates that phenomenally conscious perceptual states are at issue. Note that this differs from Pylyshyn's own target: his primary concern is to argue that a significant stage of visual processing — so-called early vision — is cognitively impenetrable.

In developing the case for cognitive penetrability of perceptual experience, Macpherson appeals to several experiments that appear to find color experience effects. Delk & Fillenbaum (1965) had participants match the color of two-dimensional orange-red cardboard cutouts to a color-adjustable background. Some of the cut-outs were shaped like objects associated with red: hearts, lips, and apples. Others were shaped like objects not associated with red: squares, bells, horseheads, etc. Their main result was that cut-outs with shapes associated with red were color-matched to redder backgrounds than were cut-outs with shapes not associated with red. Levin & Banaji (2006) produced gray-scale images of faces with features stereotypical of black people and with features stereotypical of white people. Images constructed to be in fact identical in average luminance were deemed lighter if they had features stereotypical of white people. Moreover, racially ambiguous faces were deemed lighter if labeled 'white' than if labeled 'black.' Macpherson's discussion focuses on these two papers. But she also mentions in a footnote work by Gegenfurtner and colleagues. (Macpherson cites Hansen et al. 2006. See also Olkkonen et al. 2008, 2012; and Witzel et al. 2011.) They had participants adjust a computer screen to reduce to neutral gray both realistic images of color-associated objects (e.g., bananas) and images of non-color-associated objects. Compared to the grays to which non-color-associated objects were reduced, color-associated objects were reduced to a gray with greater color in a direction opposite that of their associated color. For instance, bananas were reduced to a bluer gray. Gegenfurtner and colleagues interpret this result as indicating that participants' memories of color affected their experience of the stimulus such that, at the gray point for non-color associated objects, color-associated objects still appeared characteristically colored to some extent (for example, the bananas still retained some yellow to be grayed-out).

Macpherson argues that these results support the cognitive penetrability of perceptual color experience, and she proposes a mechanism by which such penetration might occur. The mechanism is multi-staged. First, the perceived shape of the stimulus activates a belief that associates the shape with a color — e.g., the belief that hearts are red.[2] Second, the activation of this belief generates non-perceptual phenomenal visual imagery of the heart-shaped stimulus as red — non-perceptual because generated in imagination.[3] Third, this visual imagery interacts with the phenomenal character of the participant's visual experience so as to yield a redder color experience than the participant would have had of a differently shaped stimulus.[4] Macpherson supports her proposed mechanism by arguing that independent evidence exists for the second and third steps. In support of the second step, she cites commonplace examples of imagination, dreams, and hallucinations in which beliefs cause visual imagery. In support of the third step, she cites the Perky effect, the incorporation of external stimuli into dreams (e.g., one's ringing alarm clock), and mixed perceptual and hallucinatory states (when one hallucinates something as being in the scene one is otherwise veridically perceiving). The Perky effect (Perky 1910; Segal 1972) involves an unnoticed stimulus, just above perceptual threshold, affecting subjects' visual imagining. For example, some subjects, asked to imagine a New York skyline while looking at a screen that was blank except for an unnoticed just-above-threshold projection of a tomato, described imagining New York at sunset.

In the next section, we argue that the experiments on which Macpherson relies do not provide strong grounds for the cognitive penetrability of perceptual color experience. In the section after that, we raise a problem, in any event, for Macpherson's proposed mechanism.

3. PROBLEMS WITH THE CASE FOR COLOR COGNITIVE PENETRABILITY

The case for cognitive penetrability, based on the cited experiments, can and has been challenged on various grounds. After reviewing some, we will present further grounds based on experiments we carried out. (See Machery, forthcoming, for a review of challenges to cognitive penetrability results more generally. We mention here only those challenges we have reason to refer back to later.)

For one, it is unclear whether the results reflect a difference in perceptual *experience*, as opposed to perceptual (or, perception-based) *judgment*. To be sure, when participants color-match, the case can seem strong that results are based on their color experiences. But some of these tasks are perceptually difficult — for example, Delk and Fillenbaum had subjects peering through wax paper (intended both to blur the edges between targets and background and to reduce visual acuity) in a dimly lit room. The resulting perceptual ambiguity–concerning, what's more, colors that are borderline relative to participants' color categories — could render the colors difficult to discriminate and leave room for perceptual judgment to affect performance. Moreover, Delk and Fillenbaum did not have participants adjust background color themselves; instead, participants told experimenters what adjustments to make. These verbal commands could activate semantic representations of color. Finally, Zeimbekis (2013) suggests that Gegenfurtner and colleagues' task is subject to an anchoring effect whereby the starting (conventional) color modulates what shall count as gray on that trial. (Note that the task does not involve matching two perceived colors, as in Delk and Fillenbaum, but adjusting a stimulus' color to what one considers gray.) That perceptual *judgment* is subject to cognitive penetration is a rather different claim from Macpherson's (and uncontroversial).

There is another way, worth highlighting separately, that participants' performance might reflect perceptual judgment. Claims of alleged cognitive penetrability have proven quite fragile when confronted with the possibility of task compliance. Experimental tasks place participants in a situation that encourages them to comply with the perceived demands of the task. It is well-known that even an implicit conception of what the experiment is about can affect performance. (Orne 1962) The worry, then, is that just this may have occurred in the studies Macpherson cites. Indeed, in a series of papers, Frank Durgin and his students (e.g., Russell & Durgin 2008; Durgin et al. 2009 — in response to Bhalla & Proffitt 1999) have demonstrated the effects of experimental demand specifically in tasks thought to demonstrate cognitive penetrability of, or more generally top-down effects on, perceptual experience. These demonstrations do not concern specifically the experiments concerning color experience to which Macpherson appeals. But they raise a worry more generally for cognitive penetrability claims based on experiments that don't carefully guard against task compliance.[5]

Further, even if stored information associating shapes and colors does affect perceptual *experience* of color, it doesn't yet follow that there is cognitive penetration. For not all top-down effects are cognitive penetration: it matters *where* the information is stored, thus from where the penetration comes. In particular, cognitive penetration requires that the effect originate in beliefs, goals, and the like, not in associations stored elsewhere — for example, in vision itself or in some other relatively modularized system. Deroy (2012) and Bitter (2014) develop this sort of objection in reply to Macpherson. That an effect is top-down but not an instance of cognitive penetration, of course, does not drain it of interest; but it is important to interpret correctly the nature of such effects — indeed, all the more so, given their interest.[6]

Finally, we turn to the challenge that motivated our own experiments. Firestone & Scholl (2014) suggest a strategy for testing top-down claims generally, not just claims of cognitive penetrability. They take their cue from a discredited hypothesis concerning El Greco's elongated figures (Firestone 2013) — viz., that they reflect visual distortions owing to astigmatism. This hypothesis can't be right, since, if astigmatism were the cause, it would likewise have distorted the visual appearance of the canvas on which El Greco was painting, cancelling out the supposed effect. Similarly, Firestone and Scholl predict that, where there's a top-down effect, the effect should disappear if the means for measuring the alleged effect are likewise subject to the effect. If instead one continues to find an apparent perceptual distortion, then the participant's performance does not reflect a top-down effect after all, but rather something else—perhaps task compliance.

For example, Stefanucci & Geuss (2009) found that holding a long pole horizontally makes apertures look narrower — or so they concluded from participants' performance. Firestone and Scholl first replicated their result by having participants turn away from the aperture and tell an experimenter how far to draw out a measuring tape so as to equal the width of the aperture. They then repeated the experiment but with the experimenter adjusting, not a measuring tape, but an aperture just like the target participants had seen. Their prediction was that, had there been a real top-down effect in the replication, the effect should disappear in the second experiment (the aperture used for measurement should likewise look narrower than it is). But the effect did not disappear: participants holding the pole still matched the target aperture to a narrower object of comparison than did participants without the pole — in this case to an object of comparison that was an aperture just like the target. This suggests that something else explains participants' performance—likely an effect on *judgment*, perhaps driven by task compliance.

With this in mind, we attempted to apply Firestone and Scholl's strategy to results like those on which Macpherson bases her case. In fact, we made two kinds of attempt. (For details, see below: Appendix, Experiments 1A, 1B, and 2.) In the first, participants were presented two shapes and instructed to indicate which was redder. The colors ranged, in varying degrees across trials, from reddish-orange to red. As for the shapes, there were three conditions: either both were not characteristically red (e.g., bunny, oval), both were characteristically red (e.g., heart, lips), or one was and the other wasn't. Accuracy and response times were recorded. If participants' accuracy and response times in the 'one is/one isn't' condition and the 'both are' condition agreed with each other but differed (in the right way) from those of the 'neither are' condition, this would amount to a Firestone-Scholl-type result. In the other attempt, we had participants color-match a reddish-orange heart to one of two backgrounds. In the first condition, the color-match was to the entire background screen; in the second, it was to a larger heart. A Firestone-Scholl-type result would have both exhibiting apparent top-down effects like those of Delk and Fillenbaum.

It turned out that our results do not raise a Firestone-Scholl-type challenge. But nor do they support the conclusion that perceptual color experience was cognitively penetrated or even just subject to top-down effects. In fact they raise a more basic challenge. For unlike the case of Firestone and Scholl's examination of Stefanucci and Geuss's results, we failed to replicate the original results. Both experiments had attempts at conceptual replication built-in; but, in each, participants' performance did not suggest that the characteristically red stimuli looked redder than the stimuli that were not characteristically red. In the first experiment, participants' accuracy and response times differed only with the degree of difference in the presented colors. They did not differ across the three shape conditions. Moreover, participants' performance when colors were identical did not display any top-down effect. In 'one is/one isn't' trials with identical colors, participants were at chance in answering the forced-choice question: which is redder? In the second experiment, participants in *both* conditions adjusted the background to a color very close to the true one, with no bias towards red. It is particularly striking that we failed to replicate in Experiment 1. With Experiment 2, it was at least *possible* for participants to correctly match the target's true color. But in Experiment 1, it was *im*possible for participants to give a correct answer to the forced-choice question when colors were in fact identical. Arguably, this provided added opportunity for a top-down effect — even if only on perceptual judgment. Yet no significant effect was found.

There are of course many reasons why an attempted replication might fail. But it is worth noting that among the features that made these experiments conceptual — not direct — replication attempts was the use of computer screens rather than Delk and Fillenbaum's cardboard cut-outs viewed in dim lighting through wax paper. We noted earlier that this feature of Delk and Fillenbaum's set-up has been questioned, at least insofar as one wishes to draw conclusions concerning perceptual experience. Indeed, Macpherson (2012, p. 38, fn. 10) herself notes this feature and suggests that "…this detail is unimportant for my discussion of the case. One can imagine the experiment being repeated in modern conditions on a computer screen where such paper would not be required." But we carried out that experiment, and our results suggest that perhaps it is not unimportant after all and that the challenge has some support. Parallel remarks hold for our having had participants adjust the background themselves instead of via instructions to an experimenter.

Our results are consonant as well with a failure of conceptual replication found by Gegenfurtner's group (Olkkonen et al. 2008). Their gray-out effect mostly vanished when realistic images (photographs) were replaced by uniformly, conventionally colored shapes that matched the shape of the realistic image. (Cf. Machery forthcoming, who discusses more generally the problem of replication failure in studies of cognitive penetrability.) Regarding their failure, however, it might be replied that the replacements for at least some of the more realistic stimuli no longer carried sufficient information to evoke the conventional color. The replacement of their realistic image of an orange, for example, would basically look like a uniformly orange disk. Interestingly, Delk and Fillenbaum's heart-shape has a direct association with its conventional color driven entirely by border curvature.

That we did not demonstrate task compliance of course does not mean there was none in the original Delk and Fillenbaum experiment. Indeed, differences between our experiments and theirs could perhaps explain our failure to produce a Firestone-Scholl-type result. For example, participants in Experiment 2 were randomly assigned to just one condition and performed their task in a significantly shorter time than did Delk and Fillenbaum's. The opportunity to discern the point of the task was thus greatly reduced. In Experiment 1, the lack of a Firestone-Scholl-type result (despite many participants guessing the point of the experiment — see Appendix) could be owing to our not sufficiently indicating which shape was the "target" being "measured" by the other shape. But, because of the replication failure, we did not rerun the experiment with this feature adjusted.

4. PROBLEMS FOR MACPHERSON'S PROPOSED MECHANISM

We turn now to Macpherson's proposed mechanism. Suppose there is a top-down effect. Does Macpherson's proposed mechanism offer a plausible account of how it might occur? We raise a worry concerning its required time-course: work on the generation of mental imagery suggests it simply takes too long for Macpherson's mechanism to be plausible. Other work suggests shorter latencies for such top-down effects, but at the cost of possibly undermining the claim of *cognitive* penetrability.

Macpherson's proposed mechanism, recall, is that the shape of the stimulus activates a belief that relates that shape to a color; this causes the generation of phenomenal visual imagery, which in turn interacts with color experience generated directly from the stimulus to yield an experience of color other than what one would have had without the top-down effect. Recall also that Macpherson offers independent evidence for the existence of the various stages taken individually: beliefs causing visual imagery in imagination, dreaming, and hallucination; and, for the interaction of non-perceptual visual imagery and perceptual experience, the Perky effect, the incorporation of external stimuli in dreaming, and the placing of a hallucinated object in perceived space. There is room to question how much support some of this evidence provides. For example, the causal direction in the Perky effect and in the incorporation of external stimuli in dreaming is opposite to that needed in her account of the cognitive penetrability of color experience. But this is not our focus.

Macpherson's mechanism has it that the distorted color experience is produced by the interaction of two experiences: the perceptual color experience generated directly by the stimulus and the non-perceptual visual imagining generated indirectly via the representation of the stimulus' shape that activates the relevant shape-color belief. This means that there are two causal pathways leading to the combined color experience. And these causal pathways may take different times to yield their contributing experiential states. In particular, we should expect the indirect path to take longer. The worry, then, is that we might expect participants to experience a *change* in color. For example, they may first experience the heart as reddish-orange, given the faster direct component, and then experience the heart as redder once the indirect component arrives. But, as far as we are aware, participants do not report any such change in color experience.

For this worry to be significant, two points about perceptual thresholds must obtain. First, the difference between the reddish-orange and the final redder color must be above the threshold for noticeable difference (similarly for other cases). Second, the temporal lag must be sufficiently long to allow the subject to first see the one color and then the other. That is, the lag must surpass a temporal threshold necessary for perceiving the color change — i.e., necessary for the directly generated color to be experienced prior to the phenomenal effect of imagination. There is in fact some reason to think both points obtain.

Regarding the first: Gegenfurtner and colleagues, on whom Macpherson relies, themselves report that the distortion in color experience is significantly above perceptual threshold. More specifically, they calculated a distortion effect approximately three to five times greater than the minimum difference in color required for discrimination. (Hansen et al. 2006)

Regarding the second: Various results suggest (1) that the time required to generate the visual imagining[7] would on average exceed 1000ms, and (2) that this would allow participants to experience the directly generated color first, before visual imagining has an effect. We briefly review these results before turning to studies that might suggest otherwise, but at the price of possibly removing the source of penetration from cognition.

Kosslyn et al. (1983, Experiment 3) had participants read descriptions of and then study various 2-d geometrical line drawings. After covering the drawing, participants were given a cue to construct a mental image of it, pushing a key when the image was complete. Times ranged from 1100-2200ms, depending on the complexity of the shape. Arguably, a 2-d heart-shape is not relevantly simpler than Kosslyn et al.'s simplest shapes — a thick cross, a triangle embedded in a triangle, and two over-lapping squares. Weber and Harnish (1974, Experiment II) had similar results in a task that required generating an orthographic image of a word from a spoken stimulus. Prior to the stimulus, participants were given a number corresponding to letter position. They then had to report whether the letter imagined in that position was vertically large or not (e.g., 'b' vs. 'a'). Average response times ranged from 1040-1350ms, depending on the cued position. Finally, Brockmole et al. (2002) presented a partially filled 4×4 grid to participants who then, at variable times after offset, had to integrate a mental image formed on the basis of that grid with another perceptually presented to them, so as to determine which grid position was left empty after integration. They calculated that the amount of information available in the mental image plateaus at 1300ms, suggesting that 1300ms is required to completely form the image. These three experiments differ in various ways, not least in the content of their imagery. Still, they suggest that 1000ms is a plausible lower bound for the average time of imagery formation.

Would this duration enable participants to experience the directly generated color first? Despite theorists' differing views of perceptual consciousness, we can arrive at a reasonable estimate of the time-course of color experience. Prinz (2012, p. 88), who considers intermediate-level representations modulated by attention sufficient for consciousness, cites Plendl et al. (1993) for the claim that intermediate-level visual areas respond to color by 100ms and work by Connor et al. (2004) that records (physiologically) the activation of attention to color by 125ms. Lamme (2003), who identifies the core neural correlates of consciousness with the presence of recurrent processing in areas supporting early perception, and Dehaene et al. (2006), who argue that perceptual consciousness further requires activation of, e.g., prefrontal areas, agree that visual consciousness occurs within 200-300ms of stimulus onset. Thus, if one uses the latencies given above, 700ms (=1000ms-300ms) constitutes a conservative lower estimate of the time-lag between perceptual color consciousness and the formation of mental imagery that would occur in Delk and Fillenbaum's task. This lag would be plausibly more than sufficient for participants to notice a change in color, assuming the difference in color was itself sufficiently large.[8]

Note that this estimate is conservative, not only because of the bounds on which it's based, but also because it ignores other possible sources of time-lag. Participants need *first* to generate the (not necessarily yet conscious) shape representation that, *second*, activates the concept HEART and *then* the belief that hearts are red. Only *then* does the clock start ticking on the generation, from that, of the visual imagining. *Further*, the interaction of the visual imagining and the directly generated perceptual color experience might itself require some time. But while these may indeed increase a lag already plausibly above threshold, we are conservative here in part because the mental imagery time-course results cited above arguably also include other elements.

The time-lag might also be lengthened if color is processed in perception more quickly than shape. Indeed, our Experiment 3 suggests just this (though cf. Viviani & Aymoz 2001). Participants were trained to press one key for certain shapes and for certain colors and to press another key for other shapes and colors. They were then presented colors without shapes (swaths without blackened borders) and shapes without colors. They were also presented stimuli that combined shape and color. Some of these combined stimuli were congruent (the shape and color were associated with the same key); others were incongruent. Regarding incongruent stimuli, some participants were instructed to press the key associated with the shape, others the key associated with the color. Average response times for colors were significantly faster than those for shapes. This suggests that Macpherson's indirect pathway — from shape representation to an effect via belief and imagination on color experience — may already start with a temporal disadvantage compared to the direct pathway. Interestingly, average response times with incongruent stimuli trended faster for participants instructed to use color association than for participants instructed to use shape association—albeit just beyond statistical significance (p = .06), perhaps owing to our smallish sample size. (Also, all but one of the former had faster reaction times than any of the latter.) If such a result could be established, this would suggest that color also *dominates* shape in perception. This would be of interest because Delk and Fillenbaum's reddish-orange hearts could be construed as (slightly) incongruent relative to the conventional association of heart-shapes with the color red. Thus, a top-down effect driven by this association would need to be sufficiently powerful to overcome the dominance of shape by color.

There is a potentially relevant difference, however, between the visual imagery considered in the cited literature and that invoked by Macpherson's mechanism. The results above concern visual imagery that was generated voluntarily, intentionally, and consciously. (Perhaps not in the Weber and Harnish experiment, but let's suppose so to be concessive.) The imagery invoked by Macpherson's model would be generated involuntarily and unconsciously — cf. Macpherson (2012, p. 55). Its *product* might be conscious, but there would be no conscious intention to produce it, nor would one be conscious of the product as an imagining or conscious of the *producing* of it as imagining. Conscious, deliberate mental activity tends generally to be slower and more effortful than involuntary, unconscious activity — as emphasized, for example, in the dual processing literature (Evans & Frankish 2008; Kahneman 2011). Macpherson could thus reply that the involuntary mental imagery her cases involve might possess a significantly shorter time-course than voluntary mental imagery.

Involuntary mental imagery has been much less explored than voluntary mental imagery. But there do exist studies that might bear on the question of time-course. Schlack & Albright (2007), for example, trained rhesus monkeys to associate motion direction and static shapes. In this case, motion-sensitive neurons in cortical area MT became shape-selective as well, with similar latencies for each feature. In this case (given the neurons they targeted and the analysis-window they accordingly selected), the latencies were no shorter than 580ms, significantly shorter than our 1000ms lower bound. Following Sakai & Miyashita (1991), Schlack and Albright suggest that the associative learning induces connectivity between motion- and shape-sensitive neurons, so that the motion-sensitive neurons' shape-selectivity is a top-down effect whereby the perception of the relevant shape generates motion imagery. The time-course from shape perception to the activation of motion-sensitive neurons would be no doubt significantly shorter than 580ms.

Do such results support Macpherson's mechanism? There are several grounds for caution. First, while it is not the interpretation they favor, Schlack and Albright point out that their results could involve the co-opting of MT-neurons for shape processing rather than a top-down influence of existing shape-sensitivity. Second, it is unclear that the shape-sensitivity in MT amounts to conscious imagery. Third, it would be a major step (one not considered or taken by Schlack and Albright) to conclude, even if there is a top-down generation of conscious imagery, that the learned association is *cognitive* — for example, a *belief* that that motion and that shape often co-occur (even assuming rhesus monkeys are capable of belief). It's notable that Schlack and Albright, in supporting their favored interpretation, advert to Freyd (1987) on representational momentum (our tendency to misperceive a moving object as slightly further along in its trajectory). Finke & Freyd (1989) reject the suggestion that representational momentum involves cognitive penetrability, in part owing to the rapidity of the effect. They take the phenomenon to occur within perception itself. Such a reply is available regarding other cases of involuntary imagery as well. Printed stimuli, for example, can generate phonological representations in skilled readers within 45ms, suggesting that phonological imagery would not lag much behind visual phenomenology. (Orden & Kloos 2005) But it is reasonable to maintain that this reflects the workings of an acquired modular capacity — or at least not penetration by belief.

Macpherson's mechanism thus arguably faces a dilemma: Insofar as the penetration genuinely comes from belief, there is a worry that it requires an implausible time-course; whereas insofar as the time-course does not raise problems, it may be implausible that the alleged top-down effect counts as cognitive. This certainly does not constitute a knock-down argument. We have not argued that there could not be involuntary imagery originating in belief; indeed, some intrusive imagery suggests otherwise. It remains to be shown, however, that involuntary imagery originating in belief would have an appropriate time-course. Our considerations present a *prima facie* challenge.

We note, finally, that even if the effect in Delk and Fillenbaum's particular experiment was owing to involuntary non-perceptual imagery, the distinct pathways of Macpherson's proposed mechanism would *prima facie* seem to *allow* the non-perceptual imagery to be initiated voluntarily and deliberately, in ways that would not involve the targeted stimulus itself. For example, the non-perceptual imagery could arise in response to an experimenter's request to imagine that the stimulus is red.[9] Because the initiation would be independent of the stimulus, the shape wouldn't matter: reddish-orange *squares* should thus look redder in such cases as well. This prediction could presumably be tested. Though we did not attempt such a test ourselves, we doubt it would yield an effect (even if the original effect were replicated). After all, one needn't wait upon an experimenter's request: look at something and then see if you can make your experienced color shift by imagining the thing to be a different color. If our hunch is correct, Macpherson would need to explain why the effect is *only* found with involuntary imagining. (Cf. Macpherson 2012, p. 58.) It is unclear how such an explanation would run if the involuntary imagery stemmed from color *beliefs*.

5. CONCLUSION

Perceptual experience may indeed be cognitively penetrable — perhaps in ways and to extents that matter for questions of justification, objectivity, and normative social theory. But we have argued that arguments from "one particularly difficult case to account for by those who don't believe that cognitive penetration of perceptual experience can take place" (Macpherson 2012, p. 59) should be treated with caution. The effect is not readily replicated — at least not without changes that would strengthen other grounds for skepticism.

We have also raised worries for Macpherson's intriguing proposal concerning the mechanism by which such cognitive penetration might occur. Our time-course considerations do not *rule out* an indirect path through mental imagery, but they do indicate what sort of further support the proposal would need. They thus point in a constructive direction if the cognitive penetration of color experience can indeed be established.[10]

APPENDIX

Experiment 1A

Method

Task

Each participant performed a color comparison task whereby they indicated which of the two stimuli presented simultaneously on each trial was redder than the other. Participants had unlimited time to respond by either pressing 'z' on standard keyboard if they thought the left figure was redder or 'm' if the right. A 2s grating mask screen with high spatial frequency was inserted in between trials to remove any lingering visual after-effects. After the experiment, a survey questionnaire was administered to the participants probing their awareness of the experiment's hypothesis.

Stimuli

Similar but slightly different from those used in the original Delk & Fillenbaum (1965)'s experiment, stimuli were ten pictures with colors ranging from red to reddish-orange. Five had characteristically red (henceforth CR) shapes (heart, apple, lips, fire extinguisher, stop sign), while the other five had shapes that are not characteristically red (henceforth NCR) (bunny, mushroom, bell, pentagon, oval). This led to three task conditions: those with two CR shapes, two NCR shapes, or one CR and one NCR shape. All stimuli had approximately the same surface areas.

Each of the ten figures could appear with any of the 15 color shades, which were created by adjusting the hue value (keeping the brightness and saturation constant) on the color wheel feature of Adobe Photoshop CS2. These 15 hue values were 0° (RGB 255-0-0), 6°, 10°, 11°, 12°, 13°, 14°, 15°, 16°, 17°, 18°, 19°, 20°, 21°, and 22°. (We skipped 1°-5° and 7°-9° to avoid making the task too difficult. Our intuition and pilot data revealed that it was much more difficult to distinguish between two extremely red pictures. In fact, preliminary data revealed that performance on such trials was at chance level.) Relabeling these hue values as 1°-15° according to their respective order, we associated these colors with the shapes such that there would be 1°, 2°, 3°, or 4° degree of hue difference between any two figures in a stimulus pair. A complete set of stimuli conditions (2,750) was then created by combining total number of possible shape pairs (55) with total number of possible hue pairs (50). Two complete sets consisting of 5,500 were divided among 25 subjects such that each participant completed 220 trials in an experiment.

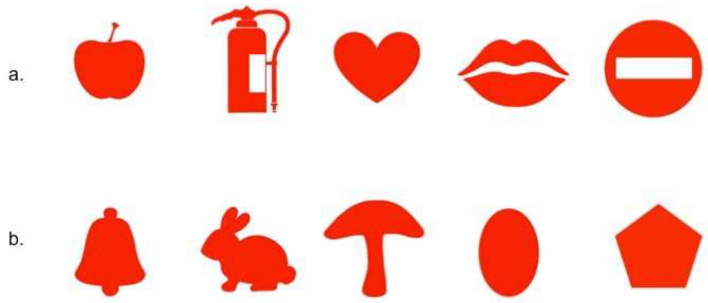


Figure 1: 1 CR (a) and NCR (b) cutouts used in the experiment.

Participants

25 JHU undergraduate students with age ranging from 17-24 participated in the experiment for a course credit. All had normal or corrected-to-normal vision. None were colorblind. The procedure for this and all experiments reported was approved by the JHU IRB.

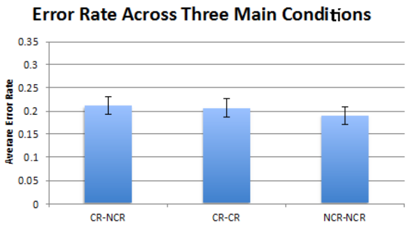
Experimental Setting

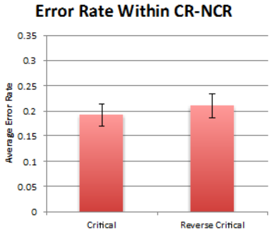
The experiment was conducted using Psychopy Builder Software v.1.80.01 (Peirce 2009) on a 2010 13" MacBook Pro LCD Display with a spatial resolution of 1280 × 800 pixels.

Results

Performance Accuracy

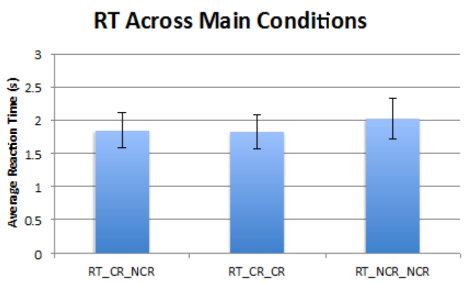
ANOVA analysis reveals no significant difference in error rate across the three main conditions: CR-NCR, CR-CR, and NCR-NCR ($p>0.1$). Within the CR-NCR condition, t-test analysis reveals no significant difference between 'critical trials' (trials where participants indicated that CR cutouts were redder when in fact the opposite was true) and 'reverse critical trials' (participants indicated NCR as being redder when the opposite was true) ($p>0.5$; mean error rate $= 0.31$ and 0.33 respectively). Using ANOVA and pairwise comparison test with Bonferroni adjustment, we however found significant difference across all degrees of hue difference (1°, 2°, 3°, 4°) that could exist between any two stimuli ($p<0.01$).





Reaction Time

Reaction time data shows similar patterns. No significant effect of characteristic redness was found across the three main conditions ($p>0.5$) or across the critical vs. reverse critical trials ($p>0.1$). In contrast to error rate data, no significant RT difference was found among different levels of trial difficulty ($p>0.5$).



Survey Questionnaire

10 out of 25 participants correctly guessed the hypothesis of the experiment. A guess was counted as correct if it made the slightest mention of association with redness of CR stimuli. To our knowledge, previous experiments like this one, including that of Delk and Fillenbaum, did not probe participants to determine if they had guessed the hypothesis. (Levin and Banaji informed participants in advance of the point of their experiment.)

Experiment 1B

Experiment 1A did not contain any trials in which true colors were identical. The smallest difference in hue was just 1° (in Adobe Photoshop). It's not unlikely that this difference was smaller than the effect calculated by Gegenfurtner and colleagues (three to five times greater than the minimum difference required for discrimination — see Hansen et al. 2006). If so, then Experiment 1A contained a conceptual replication failure. To be sure, however, we ran Experiment 1B which included trials with identical true colors.

Method

Similar to experiment 1, except the hue differences between any two figures in a stimuli pair were now 0°, 1°, 2°, and 3°. In addition, only one complete set of 2,750 trials (plus an extra 110 trials) was divided among 13 volunteer participants (aged 21-50), such that each completed 220 trials.

Results

Performance Accuracy

ANOVA analysis reveals a significant difference among the three main conditions ($p<0.01$). However, this significant difference was attributable to participants' superior performance in the NCR-NCR condition over the other two conditions (rather than inferior performance in the CR-NCR). Pairwise comparison tests after Bonferroni adjustment show significant difference between NCR-NCR and CR-NCR ($p=0.005$), and between NCR-NCR and CR-CR ($p<0.05$), but not between CR-NCR and CR-CR ($p>0.1$). More importantly, in the CR-NCR trials where there is no hue difference between the two cutouts (0° trials), participants did not significantly choose one cutout over the other ($p=0.1$). As with Experiment 1A, no significant difference was found across the critical vs. reverse critical condition for the 1°, 2° and 3° trials ($p>0.1$). Significant drop in performance, however, was found as the task became increasingly difficult ($p<0.01$).

Reaction Time

RTs of Experiment 1B replicated those of 1A. No significant RT difference was found across main conditions ($p>0.5$), task difficulty levels ($p>0.5$), or critical vs. reverse critical trials ($p>0.1$). Participants did not take significantly longer time to choose one shape over the other in the 0° trials ($p>0.1$).

Survey Questionnaire

12/13 participants correctly guessed the hypothesis of the experiment.

Experiment 2

Method

Task

Each participant was randomly assigned one of two different tasks. In the 'Replication' task, participants performed a task analogous to that of the original Delk & Fillenbaum (1965) study but on a computer monitor. They were presented with a small reddish orange heart on top of a colored display (Fig. 2A) and instructed to change the color of the background using the right and left arrow keys until its color was identical to that of the center foreground heart (see color details below).

In the 'El Greco' task, participants performed the exact same task but, instead of changing the color of the background of the whole screen, they changed the color of a bigger background heart displayed behind the smaller foreground heart (Fig. 2B). The bigger background heart was presented on top of a grey background that filled the remaining empty space in the display.

Figure 2: Sample displays from 'Replication' task (A) and 'El Greco' task (B). See text for details.

Colors

The colors used for the background and foreground hearts were taken from a continuum of HSV colors converted to RGB values using the hsv2rgb function from MATLAB Psychtoolbox v3.0.10. The continuum was constructed by setting the V and S values fixed (S=0.7990, V=204.0) and then varying the H value from 0.9 to 1.0 and then from 0.0 to 0.1 in intervals of .001. (These values were taken from color wheels used elsewhere in the color working memory literature — e.g., Emrich and Ferber 2012. S and V values were kept constant across all colors insofar as the computer color rendering is faithful to these values — see Bae et al. 2014.) This yielded a continuum of colors that spanned from pink to orange (Fig. 3). These colors were the ones subjects cycled through when pressing the left and right arrow keys. What is typically considered standard 'red' — defined here as the H value returned by the rgb2hsv function for the RGB triplet of pure red ([1 0 0]) — is at the middle at the H value of 0, depicted in Fig. 3 by the left triangle.

The color used for the foreground heart, following the method of Delk and Fillenbaum, was a reddish-orange hue, arbitrarily determined to be significantly more orange than true red but not so much that it was saliently distinguishable from red. Because of the limitations of the colors that can be produced on a computer monitor, we could not replicate the original study's colors exactly. The color of the foreground square can be seen in the right triangle in Fig. 3.

As is visible from Fig. 2, we included a dark grey outline around the heart to simulate the fact that in the original study there was an actual physical distance between the foreground and the background and hence a possible shadow to define the edges of the foreground. If this outline was not introduced, pilot data showed that the task was in fact trivial and participants easily achieved perfect performance on the task.

The color of the background heart in the 'El Greco' experiment and of the entire background in the 'Replication' experiment was set to start at a random color above or below the true color of the foreground heart. Whether it was above or below was counterbalanced across subjects, but was fixed to be at least 20 color values away from the true foreground color in the randomly assigned direction. The grey background in the 'El Greco' experiment was RGB = [125 125 125].

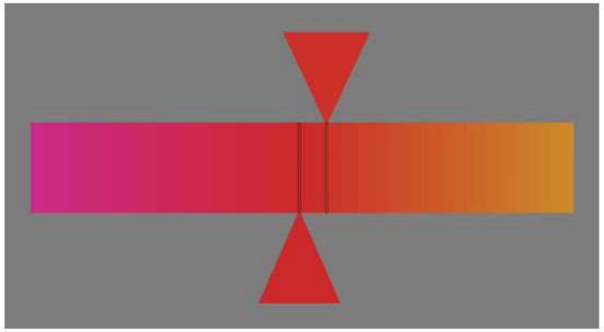


Figure 3: Continuum of colors used in the experiment. Right triangle is the true color of the foreground heart. Left triangle is 'true' red.

Participants

10 participants were tested on the 'El Greco' experiment and 9 participants on the 'Replication' experiment. Ages ranged from 16-40 years of age and all had normal or corrected-to-normal vision with no color blindness. Each participant performed either the 'El Greco' or the 'Repli-

cation' experiment once, performing only a single trial.

Computer Monitor

Both experiments were administered on a 2009 15" MacBook Pro LCD Display with a spatial resolution of 1440 × 900 pixels.

Results

The main finding was that participants were in general very good at the task, always selecting a color close to the true one, with no systematic bias. A two-tailed t-test showed that the mean color response did not significantly differ from the foreground heart's true color in either experiment ($p > .05$ for both). Because the 'Replication' experiment did not replicate the pattern of results from the original study, it is not surprising that we did not observe the expected pattern of results in the 'El Greco' version.

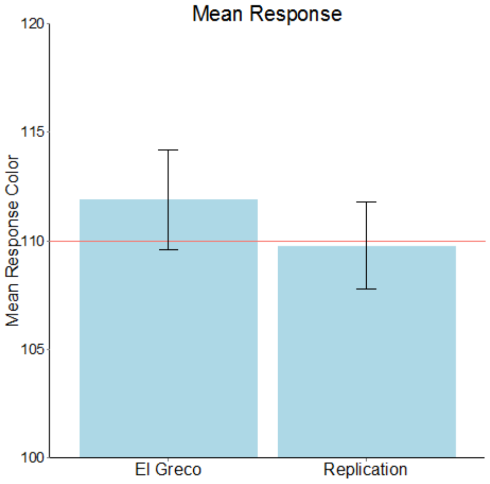


Figure 4: Mean response color for each experiment. Error bars reflect 95% confidence interval. Red line is the true color of the foreground heart (i.e. the "correct answer").

Experiment 3

Method

Participants

10 participants, aged 22 to 31 years old, participated in the Color vs. Shape task. Five participants were randomly assigned to the "Shape Trump" condition and five participants were randomly assigned to the "Color Trump" condition. All subjects had normal or corrected-to-normal vision with no color blindness.

Stimuli

All images were presented on a 13" MacBook Pro using Psychopy software. (Peirce 2008)

The "shape only" stimuli consisted of black outlines of circles, diamonds, triangles, and squares. There was no fill for the "shape only" stimuli, which resulted in the appearance of a white shape on a white background. All "shape only" stimuli were adjusted to be approximately the same size on the laptop screen (approximately 4" by 4").

The "color only" stimuli consisted of rectangular color swaths of green, red, blue, and yellow that did not have any outline. The swath measured 5" by 2" on the laptop screen. The green swath had an RGB value of 0-204-0; the red swath had a RGB value of 255-0-0; the blue swath had an RGB value of 0-0-255; the yellow swath had an RGB value of 255-255-0.

The colored shapes, or combined stimuli, were made from combining the colors from the swaths with the outlines from the shape stimuli. These combined stimuli were colored and had distinct black outlines to designate them as colored shapes as opposed to the color swaths.

Participants were explicitly told before starting the experiment that they would see shapes (black and white, outlined shapes), colors (rectangular swaths with no outline) and colored shapes (colored shapes outlined in black) in the experiment. They were instructed that the rectangle of color was not one of the target shapes, and that only the four outlined shapes that were practiced (circles, diamonds, triangles, and squares) denoted shape stimuli.

Task

Participants were randomly assigned to one of two conditions: either the “Shape Trump” condition or the “Color Trump” condition. All subjects sat in a quiet room to participate in and were given a 13” MacBook Pro on which to complete the experiment. The experimenter remained in the room the entire time to explain the experiment and answer any questions that the participant may have had during the experiment.

Participants were given written and spoken instructions that they were going to play a game with two rules — a shape rule and a color rule — and that their responses would be based on these two rules. Participants were instructed to make their responses as quickly as possible, while still being accurate.

First, all participants learned the “shape” rule, where they were instructed to press “z” when they saw either a circle or diamond and press “m” when they saw a triangle or square. Participants then practiced the shape rule in isolation for 20 trials.

Next, all participants learned the “color” rule, where they were instructed to press “z” when they saw either a green or red color swath, and press “m” when they saw a blue or yellow color swath. Participants then practiced the color rule in insolation for 20 trials.

Then, participants were instructed that they were going to need to combine the two rules to finish playing the game. They were told that they would see plain shapes (and use the shape rule to respond), plain color swaths (and use the color rule to respond), and colored shapes. For the consistent colored shapes (e.g. a green circle), participants were told to respond according to the previously learned rules. For the inconsistent colored shapes (e.g. blue circle), participants in the Shape Trump condition were instructed to respond using the learned shape rule. In other words, the shape rule would always trump the color rule for inconsistent stimuli. Participants in the Color Trump condition were instructed to respond using the learned color rule for inconsistent stimuli. In other words, the color rule would always trump the shape rule for inconsistent stimuli.

The combined stimuli (shape only, color only, and colored shapes) were presented randomly to the participants, who were reminded to respond as quickly as possible while still being accurate. Each individual image was shown a total of three times, for a total of 72 trials: 12 shape only stimuli, 12 color only stimuli, 24 congruent colored shapes, 24 incongruent colored shapes.

Results

Shape and Color Isolation Trials

Because all participants learned and completed trials for both the shape only and color only rules in the exact same way, the data from both Shape Trump condition participants and Color Trump condition participants were combined for the analysis of the shape vs. color isolation data. For all participants, the average number of correct responses for the shape-only stimuli was 19 out of 20 trials ($SD = 0.82$) and the average number of correct responses for the color-only stimuli was 19.2 out of 20 trials ($SD = 1.32$). These averages did not differ significantly from chance. The average reaction time for the shape-only stimuli was 0.98 seconds ($SD = 0.35$), while the average reaction time for the color-only stimuli was 0.68 seconds ($SD = 0.24$). These reaction times are statistically different from each other ($p = 0.03$) suggesting that participants’ reaction times were significantly faster for the color-only stimuli than the shape-only stimuli.

Shape Trump vs. Color Trump Conditions

Baseline responses for shape-only and color-only stimuli were excluded from analyses since they did not reflect any sort of combination of rules in the game, and these trials were duplicates of the shape-only and color-only data reported above. Within the Color Trump and Shape Trump conditions, congruent and incongruent trials were collapsed together into combined stimuli. This was done to get a more general picture of participants’ performance within each condition.

On average, participants in the Color Trump condition got 46.8 out of 48 trials correct ($SD = 0.84$), and participants in the Shape Trump condition answered correctly for an average of 43 out of 48 trials ($SD = 5.83$). These averages did not differ significantly from each other ($p = 0.19$); however, this may have been due to an outlier in the Shape Trump condition that strongly influenced the standard deviation. The trend suggests that the number of correct responses in the Color Trump condition is likely to be greater than the number of correct responses in the Shape Trump condition.

The average reaction time for participants in the Color Trump condition was 1.11 seconds ($SD = 0.36$), and the average reaction time for participants in the Shape Trump condition was 1.74 seconds (SD =

0.52). Although these reaction times did not quite differ significantly from each other (p = 0.056), the trend again suggests that, the reaction times of the participants in the Color Trump condition are likely faster than the reaction times of participants in the Shape Trump condition.

Restricted to incongruent trials, the average reaction time for participants in the Color Trump condition was 1.11 seconds (SD = 0.84), and the average reaction time for participants in the Shape Trump condition was 1.76 seconds (SD = 1.52). Again, the reaction times did not quite differ significantly from each other (p = 0.06). But, in addition to the trend in average reaction times, it was also the case that all but one of the participants in the Color Trump condition had faster reaction times than participants in the Shape Trump condition. Perhaps this could suggest that color dominates shape in perception.

Notes

[1] We won't take up how best to characterize the perception-conception distinction. Cf. Burge (2010). Some such distinction is required for the question of cognitive penetrability even to arise, at least as standardly formulated.

[2] In laying out her model, Macpherson concentrates on steps two and three, not mentioning step one — but presumably only because it's so obvious. There's perhaps a question whether the perceived shape of the stimulus *alone* activates the belief. Perhaps the perceived color of the stimulus plays a role as well. (This needn't be a conscious color perception. But if it were, this would add grist to the mill for our time-course worry.) Delk and Fillenbaum's stimuli were not red, but reddish-orange. But perhaps a color sufficiently similar to the associated color is required to generate the effect. This could be tested by seeing whether white heart-shapes, for example, have a pinker appearance than white squares. Note that Gegenfurtner and colleagues' interpretation of their result suggests that the effect should obtain for gray. Even if a sufficiently similar color is not necessary, however, it could still play a role.

[3] When we speak of a mental image, we shall always mean something with phenomenal character, i.e. conscious (experiential). We follow Macpherson (2012, p. 50-1) in counting the states thus generated in imagination as non-perceptual.

[4] Macpherson (2012, pp. 50, fn. 16, and 51-2) suggests that, rather than posit two interacting phenomenal states (one perceptual, one not), we might posit one phenomenal perceptual state that results from the interaction of vision and imagination. As we'll see, this does not affect our time-course argument. Macpherson's reason for suggesting this way of characterizing things is that "it doesn't seem plausible to suggest subjects are aware of two states or two phenomenal characters." But her own examples — mentioned presently in the main text —might indicate that one can be in two such states without being aware of it. Note, by the way, that it's crucial that the resulting state be a phenomenal *perceptual* state. It does not establish the cognitive penetration of perceptual experience to show that cognition can generate a non-perceptual phenomenal state that co-occurs with a perceptual phenomenal state—even if the subject of the non-perceptual state confuses it for a perceptual state. The cognitive penetration of visual imagination is not a controversial claim (cf. Macpherson 2012, p. 51).

[5] It's important that the perceived task demand affect perceptual judgment—or at least not perceptual experience. Otherwise, one might conclude that task compliance is itself a form of cognitive penetrability. Studies that seemed to unmask top-down claims would then rather have simply eliminated the top-down effects, not shown them to have been absent in the original experiments.

[6] Given the particular reasons Levin and Banaji's claims are of interest, it's worth mentioning a further worry about their results — viz., that it seems their stimuli were not in fact uniform in their actual color features. To yield identical *average* luminance, Levin and Banaji of necessity created *local* differential shading across their face-stimuli, which could affect perceptions and judgments of lightness. Firestone and Scholl (forthcoming) have now verified that the resulting low-level perceptual effects account for Levin and Banaji's results.

[7] Or to generate the total phenomenal state that incorporates the phenomenal contribution of imagination — see n. 4 above.

[8] Indeed, we considered testing whether participants could notice color differences three to five times above threshold introduced into stimuli 700-1500ms after onset. But the psychologists considered a positive result too obvious for it to be worth thus utilizing our limited resources.

[9] The voluntary imagery's effect in such an experiment would have to be confined to the experience of the stimulus, so that the experienced background color is not also affected — as perhaps it could be were the experimenter to request simply that one imagine red. Similarly, Macpherson's account of Delk and Fillenbaum's results must involve, not just the generation of red imagery, but imagery that affects *only* the experienced color of the, e.g., heart-shaped stimuli, *not* also that of the background. Cf. Macpherson (2012, p. 51).

[10] This paper grew out of projects in Gross and Flombaum's course on "Thought and Perception." Our thanks to the other course participants. Thanks also to Gi-Yeul Bae, Joel Pearson, Brenda Rapp, and Robert Wiley. Special thanks to Fiona Macpherson for the stimulation of her wonderfully thorough and pellucid paper. Finally, thanks to Jurgis Skilters and the other organizers of the excellent Riga conference.

References

Bae, G.-Y., Olkkonen, M., Allred, S. R., Wilson, C. & Flombaum, J. I. 2014. 'Stimulus-specific variability in color working memory with delayed estimation'. *Journal of Vision* 14: 1–23.

Bhalla, M. & Proffitt, D. R. 1999. 'Visual–motor recalibration in geographical slant perception'. *Journal of Experimental Psychology: Human Perception and Performance* 25, no. 4: 1076–96.

Bitter, D. 2014. 'Is low-level visual experience cognitively impenetrable: A critical analysis of some of the purported best evidence'. This volume.

Brockmole, J. R., Wang, R. F. & Irwin, D. E. 2002. 'Temporal integration between visual images and visual percepts'. *Journal of Experimental Psychology: Human Perception and Performance* 28, no. 2: 315–34.

Burge, T. 2010. *Origins of Objectivity*. Oxford University Press.

Connor, C. E., Egeth, H. E. & Yantis, S. 2004. 'Visual attention: Bottom-up versus top-down'. *Current Biology* 14, no. 19: R850–2.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J. & Sergent, C. 2006. 'Conscious, preconscious, and subliminal processing: A testable taxonomy'. *Trends in cognitive sciences* 10: 204–211.

Delk, J. L. & Fillenbaum, S. 1965. 'Differences in perceived color as a function of characteristic color'. *The American Journal of Psychology* 78: 290–3.

Deroy, O. 2012. 'Object-sensitivity versus cognitive penetrability of perception'. *Philosophical Studies* 162: 1–21.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K. & Waymouth, S. 2009. 'Who is being deceived? The experimental demands of wearing a backpack'. *Psychonomic Bulletin & Review* 16, no. 5: 964–9.

Emrich, S. M. & Ferber, S. 2012. 'Competition increases binding errors in visual working memory'. *Journal of Vision* 12: 12.

Evans, J. & Frankish, K. 2008. *In Two Minds: Dual Processes and Beyond*. Oxford University Press.

Finke, R. & Freyd, J. 1989. 'Mental extraction and cognitive penetrability: Reply to Ranney and proposals for evaluative criteria'. *Journal of Experimental Psychology: General* 118: 403–8.

Firestone, C. 2013. 'On the origin and status of the" El Greco fallacy"'. *Perception* 42, no. 6: 672.

Firestone, C. & Scholl, B. J. 2014. '"Top-down" effects where none should be found: The el Greco fallacy in perception research'. *Psychological Science* 25: 38–46.

——. forthcoming. 'Can you experience 'top-down'effects on perception?: The case of race categories and perceived lightness'.

Freyd, J. J. 1987. 'Dynamic mental representations'. *Psychological Review* 94, no. 4: 427–38.

Hansen, T., Olkkonen, M., Walter, S. & Gegenfurtner, K. R. 2006. 'Memory modulates color appearance'. *Nature Neuroscience* 9: 1367–8.

Kahneman, D. 2011. *Thinking, fast and slow*. Macmillan.

Kosslyn, S. M., Reiser, B. J., Farah, M. J. & Fliegel, S. L. 1983. 'Generating visual images: Units and relations'. *Journal of Experimental Psychology: General* 112: 278–303.

Lamme, V. 2003. 'Why visual attention and awareness are different'. *Trends in Cognitive Sciences* 7: 12–18.

Levin, D. T. & Banaji, M. R. 2006. 'Distortions in the perceived lightness of faces: The role of race categories'. *Journal of Experimental Psychology: General* 135: 501–12.

Machery, E. forthcoming. 'Cognitive penetrability: A no progress report'. In J. Zeimbekis & A. Raftopoulos (eds.) 'Cognitive Effects on Perception: New Philosophical Perspectives', Oxford University Press.

Macpherson, F. 2012. 'Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism'. *Philosophy and Phenomenological Research* 84: 24–62.

Olkkonen, M., Hansen, T. & Gegenfurtner, K. R. 2008. 'Color appearance of familiar objects: Effects of object shape, texture, and illumination changes'. *Journal of Vision* 8: 1–16.

Olkkonen, M., Gegenfurtner, K. R. & Hansen, T. 2012. 'High-Level Perceptual Influences on Color Appearance'. In G. Hatfield & S. Allred (eds.) 'Visual Experience: Sensation, Cognition, and Constancy', 179–98. Oxford University Press.

Orden, G. C. Van & Kloos, H. 2005. 'The question of phonology and reading'. In M. Snowling & C. Hulme (eds.) 'The science of reading: A handbook ', 61–78. Blackwell.

Orne, M. T. 1962. 'On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications'. *American psychologist* 17: 776–83.

Peirce, J. W. 2008. 'Generating stimuli for neuroscience using PsychoPy'. *Frontiers in neuroinformatics* 2, no. 10.

Perky, C. W. 1910. 'An experimental study of imagination'. *The American Journal of Psychology* 21: 422–52.

Plendl, H., Paulus, W., Roberts, I. G., Bötzel, K., Towell, A., Jr, Pitman, Scherg, M. & Halliday, A. M. 1993. 'The time course and location of cerebral evoked activity associated with the processing of colour stimuli in man'. *Neuroscience Letters* 150: 9–12.

Prinz, J. 2012. *The Conscious Brain*. Oxford University Press.

Pylyshyn, Z. 1999. 'Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception'. *Behavioral and brain sciences* 22: 341–365.

Russell, R. & Durgin, F. 2008. 'Demand characteristics, not effort: The role of backpacks in judging distance'. *Journal of Vision* 8: 755a.

Sakai, K. & Miyashita, Y. 1991. 'Neural organization for the long-term memory of paired associates'. *Nature* 354, no. 6349: 152–5.

Schlack, A. & Albright, T. 2007. 'Remembering visual motion: Neural correlates of associative plasticity and motion recall in cortical area MT'. *Neuron* 53, no. 6: 881–890.

Segal, S. J. 1972. 'Assimilation of a stimulus in the construction of an image: The Perky effect revisited'. In P. Sheehan (ed.) 'The Function and Nature of Imagery', 203–230. Academic Press.

Stefanucci, J. K. & Geuss, M. N. 2009. 'Big people, little world: The body influences size perception'. *Perception* 38: 1782–95.

Stokes, D. 2013. 'Cognitive penetrability of perception'. *Philosophy Compass* 7: 646–663.

Viviani, P. & Aymoz, C. 2001. 'Colour, form, and movement are not perceived simultaneously'. *Vision Research* 41: 2909–18.

Weber, R. & Harnish, R. 1974. 'Visual imagery for words: The Hebb test'. *Journal of Experimental Psychology* 3: 409–14.

Witzel, C., Valkova, H., Hansen, T. & Gegenfurtner, K. R. 2011. 'Object knowledge modulates colour appearance'. *i-Perception* 2: 13–49.

Zeimbekis, J. 2013. 'Color and cognitive penetrability'. *Philosophical Studies* 165: 167–175.