

12-15-2021

Filled/Non-filled Pairs: An Empirical Challenge to the Integrated Information Theory of Consciousness

Amber R. Hopkins
Chapman University

Kelvin J. McQueen
Chapman University, mcqueen@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/philosophy_articles



Part of the [Cognition and Perception Commons](#), [Neurology Commons](#), [Other Psychiatry and Psychology Commons](#), and the [Psychological Phenomena and Processes Commons](#)

Recommended Citation

Hopkins, A. R., & McQueen, K. J. (2022). Filled/non-filled pairs: An empirical challenge to the integrated information theory of consciousness. *Consciousness and Cognition*, 97, 103245. <https://doi.org/10.1016/j.concog.2021.103245>

This Article is brought to you for free and open access by the Science and Technology Faculty Articles and Research at Chapman University Digital Commons. It has been accepted for inclusion in Philosophy Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Filled/Non-filled Pairs: An Empirical Challenge to the Integrated Information Theory of Consciousness

Comments

NOTICE: this is the author's version of a work that was accepted for publication in *Consciousness and Cognition*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Consciousness and Cognition*, volume 97, in 2022. <https://doi.org/10.1016/j.concog.2021.103245>

The Creative Commons license below applies only to this version of the article.

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Copyright

Elsevier

Filled/non-filled pairs: an empirical challenge to the integrated information theory of consciousness*

Amber R. Hopkins^a and Kelvin J. McQueen^b

^a*Institute for Interdisciplinary Brain and Behavioral Sciences, Chapman University.*

^b*Philosophy Department, Chapman University.*

November 19, 2021

Abstract

Perceptual filling-in for vision is the insertion of visual properties (e.g., color, contour, luminance, or motion) into one's visual field, when those properties have no corresponding retinal input. This paper introduces and provides preliminary empirical support for *filled/non-filled pairs*, pairs of images that appear identical, yet differ by amount of filling-in. It is argued that such image pairs are important to the experimental testing of theories of consciousness. We review recent experimental research and conclude that filling-in involves brain activity with relatively high integrated information (Φ) compared to veridical visual perceptions. We then present filled/non-filled pairs as an empirical challenge to the integrated information theory of consciousness, which predicts that phenomenologically identical experiences depend on brain processes with identical Φ .

Keywords: visual perception, consciousness, filling-in, integrated information theory, feedback connectivity, filled/non-filled pairs

*Forthcoming in *Consciousness and Cognition*.

Contents

1	Introduction	3
2	Filled/non-filled pairs	5
2.1	Petals	5
2.2	Watercolor	6
3	Neural basis of filling-in	7
3.1	Blind spot filling-in	8
3.2	Contour and figure filling-in	9
3.3	Color filling-in	10
4	The empirical challenge to IIT	11
4.1	The challenge	12
4.2	Possible responses to the challenge	14
5	Conclusion	16
6	Acknowledgements	16
	Bibliography	18
A	Appendix: Experimental support for filled/non-filled pairs	24
A.1	Summary	24
A.2	Methods	25
A.2.1	Subjects	25
A.2.2	Apparatus	26
A.2.3	Stimuli	26
A.2.4	Procedure	27
A.2.5	Attribute matching task	27
A.2.6	Method of constant stimulation (MCS)	28
A.2.7	Validation	30
A.3	Results	31
A.3.1	Attribute matching task	31
A.3.2	Method of constant stimulation (MCS)	32
A.3.3	Final reconstructions	33
A.3.4	Validation	34
A.3.5	Conclusion	36
B	Appendix: additional filled/non-filled pairs	38

1 Introduction

Perceptual filling-in in the case of vision is the insertion of visual properties (e.g., color, contour, luminance, or motion) into the visual field, when those properties have no corresponding retinal input. The character of the filled-in properties is influenced by the character of the surrounding visual field (Komatsu 2006 & Devinck and Kenneth Knoblauch 2019).

Filling-in happens often and can be immediately appreciated by finding your natural blind spot (Ramachandran 1992), see Figure 1. An imperceptible portion of the visual field results from a void of photoreceptors on the retina where the optic nerve connects. Despite lacking retinal input from this region of visual space, there is no sense-less void in our conscious visual experience. Rather, the blind spot is filled-in and a seamless or complete visual field is perceived.

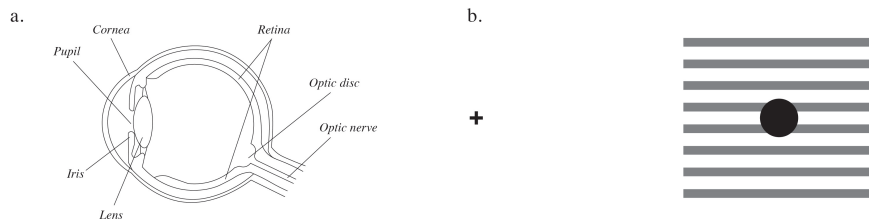


Figure 1: The natural blind spot. (a): The retina is lined with light sensitive cells that relay visual information from the eye, through the optic nerve, to visual areas in the brain. The optic disc, the part of the eye at which the optic nerve meets the retina, contains no light sensitive cells and thus corresponds to a blind spot in each visual hemifield. (b): You can find your blind spot by covering your left eye and focusing on the small black cross with your right eye. Move forward until the black dot at the center of the gray stripes on the right vanishes. The void is filled-in with the surround: continuous gray lines are perceived in place of the black dot.

Filling-in does not require a deficit in visual input, like at the blind spot. Careful arrangements of shapes and colors can induce different kinds perceptual filling-in, such as that of contour and figure, luminance, and color. In Figure 2, we offer three prevalent examples of filling-in-inducing stimuli. These examples demonstrate that it is often far from obvious when filling-in takes place in one's experience. On an initial glance, it is natural to conclude that rows of diamonds towards the bottom of 2b are darker than the rows at the top or that there are colored circles of the center of each black shape in 2c. But with a closer look, one sees that there are no real circles in 2c and every diamond in 2b is exactly the same.

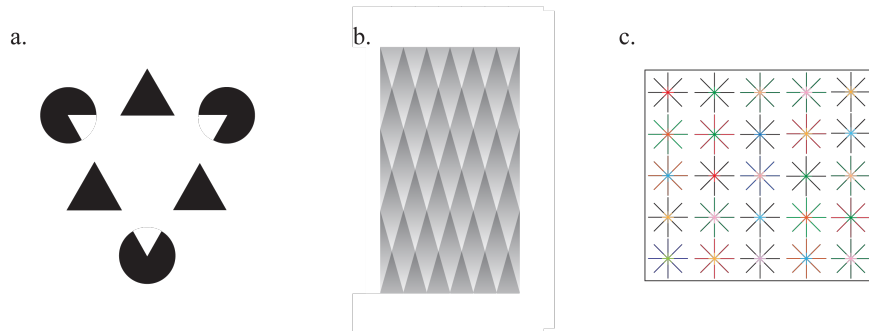


Figure 2: Examples of filling-in stimuli. (a): Kanizsa figure (Kanizsa 1979). The arrangement of black shapes, also called inducers or elements, prompts the perception of a bright-white overlaying triangle. The pronounced contours of the triangle as well as its seemingly high luminance relative to the background are filled-in. Thus, the Kanizsa figure is an example of contour and figure as well as luminance filling-in. (b): Diamond Craik-O'Brien-Cornsweet (COBC) illusion (O'Brien 1959; D. Craik and K. J. Craik 1966; Watanabe, Cavanagh, and Anstis 1995; Cornelissen et al. 2006). Identical rows of grey diamonds with a dark to light gradient from top to bottom, respectively, are stacked on top of each other. Lower rows of diamonds appear darker than rows at the top. COBC is an example of luminance filling-in. (c): Neon color spreading with Ehrenstein figures (Tuijl 1975; Redies and Lothar Spillmann 1981). Each Ehrenstein figure is a set of intersecting black lines with another set of shorter and colored lines overlapping them. At the center of each Ehrenstein figure, color from the inner segments of each line leaks into its surround and a transparent circle or disc taking the color of the lines is perceived. Neon color spreading is thus an example of contour and figure as well as color filling-in.

This paper proposes what we call *filled/non-filled pairs* and argues for their significance to the study of consciousness. Resembling metamers (Freeman and Simoncelli 2011; Mims et al. 2020), these are pairs of images or stimuli that appear identical or are introspectively indistinguishable—hereafter, *phenomenologically identical*—yet differ in feature(s). Distinct from existing metamers, filled/non-filled pairs are designed such that one image in the pair invokes more perceptual filling-in than the other. In the next section, we present two filled/non-filled pairs as examples. In Appendix B, we offer an additional 15 pairs. In Appendix A, we explore the feasibility of filled/non-filled pairs with an experiment.

In section 3 we review literature on the neural basis of filling-in. It is found, generally, that the neural mechanisms responsible for filled-in perceptions not only involve distinct brain areas from, but also recruit (more) long-range feedback connectivity than, sensory-based perceptions. Thus, filled/non-filled pairs comprise stimuli that are phenomenologically identical but differ in neural basis. Such stimuli could not only help to further isolate the neural activity responsi-

ble for filling-in and more generally the neural basis of consciousness, but could also open a novel avenue to empirically test existing theories of consciousness.

Section 4 explores a particular application for filled/non-filled pairs: we argue that they pose an empirical challenge to a leading theory of consciousness, the integrated information theory of consciousness (IIT). According to IIT, if filled/non-filled pairs are indeed phenomenologically identical, then their neural correlates must possess the same amount of integrated information (Φ). However, we will argue that because filling-in significantly increases long-range feedback connectivity, it thereby increases Φ . Accordingly, filled/non-filled pairs challenge IIT's predictions. Section 4.1 presents the objection. Section 4.2 then considers various possible responses on behalf of IIT and emphasizes their testability. Lastly, section 5 concludes that filled/non-filled pairs are important to the experimental testing of theories of consciousness.

2 Filled/non-filled pairs

Filled/non-filled pairs comprise stimuli that appear identical, yet differ in their physical properties and the amount of filling-in they generate. To create these pairs, we take an initial, established perceptual filling-in illusion. Then, we isolate and reconstruct the illusory or filled-in properties perceived in the initial illusion. Lastly, the reconstruction is inserted into the initial image to create the non-filled version. Filled/non-filled pairs face several challenges, which we raise and address as we describe two example pairs below.

2.1 Petals

The petals illusion (Figure 3a; Kennedy 1981) invokes luminance filling-in—the illusory perception of brightness or darkness across a visual region. The center of the figure to which all of the petals point appears brighter than the background despite being identical in luminance to the background. A petals filled/non-filled pair can be made by isolating and reconstructing the illusory brightness perceived in the filled image (Figure 3b) and then inserting it into the initial image to create the non-filled image (Figure 3c).

The petals filled/non-filled pair (Figure 3) may not appear identical to every observer. After all, it is unlikely that everyone fills in an image in exactly the same way. However, it seems plausible that the reconstruction used to create the

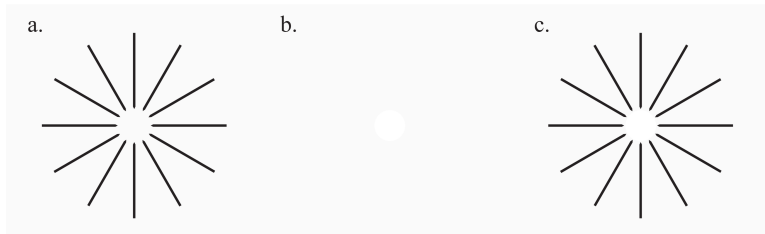


Figure 3: Filled/non-filled petals pair. (a): Petals filled image. Brightness is filled-in at the center of the petals creating an illusory discrepancy in luminance between the center of the figure and the background. (b): Petals reconstruction. A circle with higher luminance than the background. (c): Petals non-filled petals image. The reconstruction (b) is added to the initial illusion (a). Images (a) and (c) make up the petals filled/non-filled pair.

non-filled member of the pair could be *tuned* or *matched* for a given person such that the images in the pair are identical for that person. We have conducted a modest experiment (see Appendix A) to demonstrate how tuning or matching, namely an attribute matching task and the method of constant stimulation, can be used to create filled/non-filled pairs.

2.2 Watercolor

The watercolor effect (Figure 4a; (Pinna 1987)) is an instance of contour-dependent surface color filling-in. Within the area enclosed by wavy two-color contour(s), with the inner contour being a light chromatic and the adjacent or fringe contour being a darker chromatic, there is a shift of surface appearance towards that of the inner contour(s).¹ The colored surface perceived in the initial watercolor effect stimulus can be isolated and reconstructed (see Figure 4b). By then adding this reconstruction back to the initial illusion, we create the non-filled match (see Figure 4c).

One worry may be that the filled-in property is irreducibly distinct from the real property in the reconstruction. In other words, maybe there is some phenomenological difference between a real color (e.g., see Figure 4b) and a filled-in one (e.g., see Figure 4a). This would make the images in our pair seem different, always. However, filled-in colors have been matched with real colors both in our experiment (Appendix A) as well as others.²

¹Pinna, Brelstaff, and Spillmann (2001), Baingio Pinna and Grossberg (2005), Coia and Crognale (2018), and Devinck and Kenneth Knoblauch (2019).

²Filled-in color perceived in watercolor effect illusions have been matched to real colors with, for example, color matching and hue-cancellation tasks, see Devinck, Delahunt, Hardy, Spillmann, et al. 2004; Devinck, Delahunt, Hardy, Spillmann, et al. 2005; Heydt and Pierson

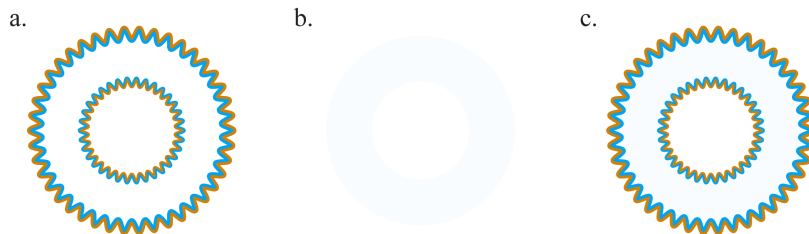


Figure 4: Watercolor effect filled/non-filled pair. (a): Watercolor filled image. The annulus is filled-in with a faint color taking the hue of the inducing contour—the innermost contour of the annulus. (b): Watercolor reconstruction. An annulus is set to the inner contour’s hue but at low-opacity or saturation. (c): Watercolor non-filled image. Adding (b) to (a) completes a non-filled version of the watercolor effect image. (a) and (c) comprise the watercolor filled/non-filled pair.

This paper is rooted in an assumption that the images in filled/non-filled pairs differ by amount of filling-in they generate. We suggest that the presence of the reconstruction in the non-filled image “blocks” filling-in.³ A challenge arises if filling-in is not blocked. If the reconstruction fails to block filling-in, then filling-in would *overlay* the reconstruction added to the non-filled image. This would inevitably lead to a perceivable difference between the filled and non-filled members of the pair. We can rule out this scenario neurophysiologically, by showing that the neural dynamics implicated in filling-in differ when viewing each image in the pair, or behaviorally, by demonstrating that non-filled images made with reconstructions can look identical to filled images. We offer preliminary behavioral support for filled/non-filled pairs in Appendix A.

3 Neural basis of filling-in

Filling-in recruits neural activity distinct from sensory-based perception (Komatsumatsu 2006). Below we review neuroscientific research on blind spot, contour and figure, as well as color, filling-in to highlight that feedback connectivity is implicated in filling-in while sensory-based percepts rely more heavily on feed-forward activations. Crucially, long-range feedback connectivity is known to

2006; Devinck, Hardy, et al. 2006.

³Filling-in has been blocked consistently in watercolor effect experiments by interleaving or braiding the bands of color comprising the contours, eliminating the stimulus configuration that causes filling-in (Devinck and Knoblauch 2012; Devinck, Peggy Gerardin, et al. 2014; Gerardin, Abbatecola, et al. 2018). However, this does not allow for the kind of phenomenological identity we aim for here, since the initial illusion does not have braided contours.

generate large amounts of integrated information, making this distinction important for IIT.

3.1 Blind spot filling-in

Single-unit recordings of neurons in monkey primary visual cortex (V1) led to the discovery of “blind spot neurons”—neurons that represent the blind spot. These neurons exhibit several relevant qualities. Blind spot neurons are found in deep layers of V1 and possess large receptive fields that extend beyond the blind spot (Komatsu, Kinoshita, and Murakami 2000). Furthermore, they are selectively activated based on the character of the filled-in properties at the blind spot.⁴ A later fMRI study with humans confirmed that the blind spot is both represented in deep portions of V1 and that this region is activated in accordance with what is perceived (Tong and Engel 2001).

In addition to feedforward input from the eyes and lateral geniculate nucleus (LGN), V1 receives feedback from many other cortical regions, including higher visual areas (including V2, V3, V4, and V5).⁵ Within each cortical area, feedback signals predominantly target deep and superficial layers, while feedforward signals arrive in middle layers.⁶ Moreover, excitatory pathways between deep V1 layers and superficial and middle layers facilitate horizontal or intracortical communication (Douglas and K. A. Martin 2004). Thus, the location of blind spot neurons in deep layers of V1 suggests that they receive feedback signals from higher visual areas and participate in horizontal or intracortical communication.⁷

Blind spot neurons possess large receptive fields that extend beyond the blind spot, which demonstrates that blind spot neurons respond to visual information not only within the blind spot but also from the surround. These large receptive fields seem necessary if the sensitivity of blind spot neurons to the surround facilitates blind spot filling-in. In addition, blind spot neurons are selectively activated in accordance with what is perceived despite receiving no

⁴Fiorani Júnior et al. (1992), Komatsu, Kinoshita, and Murakami (2002), and Matsumoto and Komatsu (2005).

⁵Muckli and Petro (2013), Markov et al. (2014), and Bastos et al. (2015).

⁶Rockland and Pandya (1979), Felleman and Van Essen (1991), Anderson and K. Martin (2009), and Harris and Mrsic-Flogel (2013).

⁷Interestingly, Matsumoto and Komatsu (2005) found that response latencies in V1 are significantly longer for stimuli presented at the blind spot than outside of the blind spot. This time lag could allow for feedforward and feedback signals between V1 and higher visual areas to fill-in at the blind spot.

retinal input corresponding to the region of visual space they represent. Thus, their activity must not be determined by feedforward input from the eye but by connected neurons. Given their location in deep layers of V1, horizontal projections from neighboring neurons or feedback projections from higher visual areas may modulate the activity of blind spot neurons based on properties in the surround or global image properties, respectively. In sum, blind spot neurons appear to integrate a large amount of information from connected neurons.

3.2 Contour and figure filling-in

Contour and figure filling-in have also been linked to feedback activation of deep layers in early visual areas. Using invasive laminar recordings in macaque V1 during figure filling-in, Self et al. (2013) report that figure filling-in recruits activity in superficial and deep layers of V1, targets of feedback projections, while sensory input activated middle layers of V1, targets of feedforward projections. Similarly, but with fMRI and humans, Kok et al. (2016) report that perceiving the illusory white triangle in the Kanizsa image (see Figure 2a) activates deep layers of V1 while sensory-based stimulation activated all layers of V1, but especially middle and superficial layers. Kok et al. (2016) also report that the activity in deep layers of V1 associated with perceiving the illusory triangle figure was modulated by, and dependent upon, sensory stimulation from the black elements or inducers. In both experiments, the selective activation deep layers of V1 implicates feedback connections in figure filling-in, given established feedback connectivity from higher visual areas to this part of V1. Contour filling-in seems to follow the same patterns as figure filling-in. In an fMRI study with monkeys, Lee and Nguyen (2001) found neurons in V1 and V2 with receptive fields on the illusory contours of the Kanizsa triangle that are active during contour filling-in. The neuronal response to illusory contours was strongest in deep layers of V2, evidencing the recruitment of feedback connectivity for contour filling-in.

Interestingly, Self et al. (2013) also report temporal distinctions between neural activation by veridical visual input and brain activity associated with filling-in. Sensory-based edge detection began in middle and superficial layers around 70ms after stimulus onset, while activity in superficial and deep layers corresponding to figure filling-in 100ms after stimulus presentation. Similarly, Lee and Nguyen (2001) also report that neurons in V1 respond to the illusory

contours later than the neurons in V2 and also that responses in deep layers occurred after responses in superficial layers (in both V1 and V2). The temporal sequence of activity in V1 and V2 could suggest that neurons in V2, which possess large receptive fields and thus integrate edge signals from a larger portion of the visual field, send feedback input to V1 to modulate contour filling-in.

3.3 Color filling-in

Perceiving veridically-colored surfaces also activates networks in the brain distinct from those implicated in surface color filling-in. An fMRI study with humans and watercolor effect color filling-in by Gerardin, Abbatecola, et al. (2018) ties surface color filling-in to activity in dorsal areas of V3, and real surface color perception to ventral portions of V4 and the lateral occipital complex. After all, the vast region of visual space that is filled-in with color requires whole-image or global context appreciation (Hale and Brown 2021) and thus the participation of cortical areas with sufficiently large receptive fields and contour curvature sensitivity. Thus, Devinck and Kenneth Knoblauch (2019) argues that visual cortex beyond V1 and V2 (especially V3) and feedback connectivity (which has been previously tied to contextual processing; Zipser, V. A. Lamme, and Schiller (1996)), are implicated in watercolor illusion surface color filling-in. This is consistent with reports that color filling-in does not activate early visual areas at all (Cornelissen et al. 2006).

However, V1 seems important for color filling-in in detecting edges or contours. In a single-unit recording study with monkeys viewing a color filling-in disc stimulus, Heydt, Friedman, and Zhou (2003) recorded neurons in V1 and V2 with receptive fields either inside the disc or on its border. The activity of neurons with receptive fields inside the disk held steady during filling-in. However, neurons with receptive fields on the disc’s contours increased in activity, which could mean that early visual areas provide contour information used to fill-in color. Gerardin, Abbatecola, et al. (2018) emphasize the role of V1 even further, using connectivity analyses to argue that feedback modulation from V3 to V1 and lateral occipital complex is responsible for watercolor filling-in while feedback connections from lateral occipital complex modulating V1 and V3 implicated in perceiving veridical surface color.

In sum, sensory-based perceptions are given rise to by neural networks distinct from filled-in ones. Surface color filling-in seems to recruit higher visual

areas V3 and V4 but also feedback connections from higher to lower visual areas. Perceiving real contours activates feedforward activity, mostly in middle layers of V1, while feedback activation of deep and superficial layers in V1 or V2 with a lag is implicated in contour and figure filling-in. And blind spot filling-in is attributed to perceptually-correlated activity of blind spot neurons found in deep layers of V1 with large receptive fields. Across aforementioned kinds of filling-in, feedback connectivity from higher visual areas to lower visual ones is implicated in filling-in while sensory-based percepts are given rise to predominately by feedforward activations. Here, a filled stimulus would recruit feedback connectivity while its non-filled match would involve more feedforward activations.

4 The empirical challenge to IIT

Our experiments (Appendix A) suggest that filled/non-filled pairs are possible. Meanwhile, our literature review (section 3) suggests that the neural mechanisms responsible for filling-in are distinct from those responsible for sensory-based perceptions. In particular, perceptual filling-in involves more long-range feedback connectivity. We may therefore draw two conclusions relevant for theories of consciousness:

1. We have provided preliminary support for the general claim that two visual experiences can be phenomenologically identical yet depend on quite distinct neural mechanisms. This raises a problem for theories of consciousness that start from phenomenology and try to derive the neural basis of phenomenology from there. Such theories may not attribute different neural bases to phenomenologically identical experiences.
2. We have provided preliminary support for the more specific claim that two visual experiences can be phenomenologically identical despite significant differences in underlying feedback connectivity. This raises a problem for theories of consciousness that give a central role to feedback connectivity. This is especially the case if they predict that phenomenologically identical experiences should not differ with respect to underlying feedback connections.

In this section, we show that filled/non-filled pairs may play a central role

in testing theories of consciousness by arguing that an influential theory of consciousness, the integrated information theory (IIT)⁸, faces both problems.⁹ IIT starts from phenomenology in its methodology, and predicts that phenomenologically identical experiences share identical neural bases. Because of the central role that feedback connectivity plays in IIT, IIT appears to predict that phenomenologically identical filled/non-filled pairs cannot differ with respect to amount of underlying feedback connectivity. Section 4.1 explains this challenge to IIT in more detail. Section 4.2 considers possible responses available to IIT and emphasizes that they may be testable.

4.1 The challenge

IIT postulates a mathematical measure of *amount of consciousness* in a system: the integrated information (Φ) of the system. A system with zero Φ is unconscious. A system with nonzero Φ is conscious, and its level of consciousness corresponds to its level of Φ . Experimental support for IIT is limited, especially because it is impractical to measure and calculate Φ in biological systems. However, some measurable approximations of Φ have been shown to correlate with level of consciousness.¹⁰

Nonetheless, the Φ measure was not initially derived from experiments, rather, it was derived from phenomenology:

“IIT starts from phenomenological axioms: information says that each experience is specific – it is what it is by how it differs from alternative experiences; integration says that it is unified – irreducible to non-interdependent components; exclusion says that it has unique borders and a particular spatio-temporal grain. These axioms are formalized into postulates that prescribe how physical mechanisms, such as neurons or logic gates, must be configured to generate experience (phenomenology).” (Oizumi, Albantakis, and Tononi 2014)

Starting from phenomenology means using introspection to reflect on the

⁸See Tononi (2004), Tononi (2008), Oizumi, Albantakis, and Tononi (2014), Tononi and Koch (2015), Tononi, Boly, et al. (2016), Tegmark (2016), Tsuchiya (2017), and Mayner et al. (2018).

⁹The Recurrent Processing Theory of consciousness (V. Lamme 2006) may also face the second problem, due to its emphasis on feedback connections.

¹⁰See Massimini et al. (2005), Casarotto et al. (2016), Leung et al. (2020), and Afrasiabi et al. (2021).

essential properties of conscious experience. IIT therefore seems committed to saying that two experiences have identical phenomenology if and only if they are introspectively indistinguishable. Furthermore, IIT predicts that two experiences with identical phenomenology have neural bases with identical Φ . However, this is challenged by our findings, which suggest that phenomenologically identical experiences can have significantly different neural bases in a way that affects Φ .

Consider what IIT predicts about filled/non-filled pairs: If the two visual experiences are phenomenologically identical, IIT predicts that their neural bases have identical Φ . However, feedback connectivity makes a significant difference to Φ : the more feedback connectivity in a system, the more Φ in the system.¹¹ IIT therefore seems to wrongly predict that the neural bases of filled/non-filled pairs will not differ with respect to amount of feedback connectivity. This is the challenge to IIT.

Interestingly, filled/non-filled pairs may be interpreted as concrete examples of a more general phenomenon, identified by Doerig et al. (2019) in their unfolding argument against IIT. Theoretically, any feedback network can be unfolded into a feedforward network that implements the same functions. In particular, any behavioural experiment can be seen as an input-output function, and so theoretically can be implemented by both feedback and feedforward networks. If we assume functionalism (specifically, that functional identity among two networks is sufficient for phenomenal identity), then it follows that the very same percept can be created by very different networks with different Φ . Thus, either IIT is false or functionalism is false (in the latter case, the authors argue this takes IIT outside the realm of experimental science).

IIT might escape the unfolding argument by denying that an unfolded cortex is physically plausible. After all, unfolding complex feedback networks into feedforward networks typically requires many more neurons, which raises questions about how such a mass increase will affect functionality. However, filled/non-filled pairs illustrate the physical possibility of unfolding in the human cortex. In particular, the neural basis of the non-filled experience can be thought of as a (partially) unfolded version of the filled experience's neural basis. Perceiving

¹¹Doerig et al. (2019, p51) interpret IIT such that Φ is always greater than zero in feedback systems (they are always conscious) and always equal to zero in feedforward systems (they are never conscious). Chalmers and McQueen (forthcoming, Sec.7) assert that feedback connectivity is a good heuristic for estimating Φ in complex systems, in part because of certain non-zero Φ feedforward systems, which we discuss in the responses section below.

the non-filled image of course does not involve purely feedforward networks. But the two experiences seem to at least *differ in amount* of feedback activity. Due to difference in feedback activity, the non-filled member has less Φ than the filled member, but they are phenomenologically identical, contradicting a key claim of IIT.

4.2 Possible responses to the challenge

The challenge to IIT can be presented as a dilemma: either explain how it could be that filled/non-filled pairs are impossible or explain how their existence is consistent with IIT, in light of what we know about the brain mechanisms. We emphasize that our analysis has not experimentally refuted IIT. It poses a challenge, but there are several ways IIT may respond. Many of these responses seem experimentally testable. We therefore hope our analysis spurs further experimental research into filled/non-filled pairs and IIT. In what follows, we take each horn of the dilemma in turn.

By taking the first horn, one denies that phenomenologically identical filled/non-filled pairs are possible. One could claim that there is some inevitable phenomenal difference between filled-in experiences and sensory-based experiences. Or one could claim that affixing reconstructed properties to filled images does not ‘block’ filling-in but causes overlay. However, our experiments (Appendix A) suggest otherwise. We therefore welcome further experimental research into this question.

By taking the second horn of the dilemma, one accepts phenomenologically identical filled/non-filled pairs, but denies that they contradict IIT. However, as we shall see, this strategy requires making either substantive, testable postulates about the neural basis of filling-in, or substantive revisions to IIT. Below we consider two revisions, and then two testable postulates.

The first revision allows that at least some phenomenology is introspectively inaccessible. This means denying the strict assumption that two experiences have identical phenomenology if and only if they are introspectively indistinguishable. The upshot is that IIT can allow the Φ -differences we’ve identified to correspond to phenomenal differences, even though they don’t correspond to introspectively accessible differences. Perhaps Φ differences must meet a certain threshold before a subject can identify them. However, it is not immediately clear how to square this with IIT’s phenomenology first approach. Furthermore,

the Φ -relevant differences do not seem negligible, they seem substantive. Still, without exact quantification, this remains an open possibility.

The second revision states that Φ , rather than measuring amount consciousness generally, just measures amount of filling-in. We think this is an interesting testable hypothesis that is worth putting on the table. It would, however, involve significantly rethinking IIT. Let us now consider postulates that would allow IIT to avoid revisions.

The first testable postulate denies that there is any significant difference in amount of feedback connectivity, between the neural bases of the two phenomenologically identical experiences. This means denying that the results of our literature review isolate the neural bases of the experiences. For example, one might claim that there are many more feedback processes underlying non-filled perception (enough so that IIT makes the right predictions about filled/non-filled pairs), we've just failed to identify them. We would invite advocates of this response to identify these processes. Alternatively, one might claim that the large- Φ (feedback) mechanisms necessary for filling in are actually unconscious processes, which are not part of the physical basis of the experiences. However, to render them unconscious in light of their nonzero Φ , IIT must identify nearby neural mechanisms that have even greater Φ . This is because of IIT's exclusion postulate, which states that Φ (no matter how large) is not sufficient for consciousness: to be conscious a system must be a *maximum* of Φ : it must have more Φ than any of its subsystems and must have more Φ than any supersystem it is a part of. Again, we invite advocates of this response to identify these nearby large- Φ mechanisms.

The second testable postulate accepts that there is a significant difference in amount of feedback connectivity between the neural bases of the two phenomenologically identical experiences, but it denies that this falsifies IIT. In particular, it denies that the neural bases have different Φ , just because they have different amounts of feedback connectivity. Here, one may appeal to the special case in which a purely feedforward network has nonzero Φ : when it has overlapping inputs and overlapping outputs.¹² Thus, one might contend that these conditions obtain for the neural basis of the non-filled experience, and that this renders the Φ of the two experiences equivalent. It is an experimental question as to whether these conditions are met by the non-filled experience.

¹²For illustration and calculation see Oizumi, Albantakis, and Tononi (2014, Fig. 7(B)).

However, this final strategy comes into conflict with another IIT prediction. Recall that Φ is just a correlate of the *amount* of consciousness in a given system. IIT also postulates a correlate of a system’s *total experience* (or total state of consciousness). These correlates are called *Q-shapes*, and geometrically model the “shape” or “pattern” of integrated information in the neural basis of the experience. IIT postulates a 1-to-1 correspondence between experiences and Q-shapes: each experience has a unique Q-shape and vice versa.¹³ The Q-shape of a brain state depends on the distribution of integrated information throughout that state. The problem is that integrated information would appear to be distributed quite differently for feedback and feedforward network states. Moreover, the neural bases of filled and non-filled experiences have a distinct spatiotemporal structure: they recruit distinct brain regions and filled percepts take longer. But different Q-shapes entail qualitatively distinct experiences. Thus, there is no obvious way out of the dilemma for IIT.

5 Conclusion

The integrated information theory starts from phenomenology, and deduces the mechanisms that must underly conscious experiences. Filled/non-filled pairs suggest that this strategy may be wrong-headed, since identical phenomenology can arise out of distinct neural processes. Here, we have argued that mechanisms responsible for filling-in have greater integrated information than sensory-based neural responses. We created filled/non-filled pairs to press this challenge. This is just an example of the theoretical significance of filled/non-filled pairs. We believe that they may play a pivotal role in finding the complete neural correlates of consciousness. We therefore hope to have established filled/non-filled pairs as a tool for consciousness research.

6 Acknowledgements

We would like to thank David Chalmers, Lucas Jeay-Bizot, Alie Male, Brennan McDavid, Michael Pace, Marco Panza, and Aaron Schurger for helpful feedback. We are also grateful to the reviewers for thought-provoking discussion and sug-

¹³For detailed calculation of both the Q-shape and the Φ of the simplest feedback network, see Chalmers and McQueen (forthcoming, appendix).

gestions that strengthened our manuscript substantially. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Bibliography

- Afrasiabi, Mohsen et al. (2021). “Consciousness depends on integration between parietal cortex, striatum, and thalamus”. *Cell systems* 12.4, pp. 363–373.
- Anderson, JC and KAC. Martin (2009). “The Synaptic Connections between Cortical Areas V1 and V2 in Macaque Monkey”. *The Journal of Neuroscience* 29.36, pp. 11283–11293.
- Bastos, Andre Moraes et al. (2015). “Visual areas exert feedforward and feedback influences through distinct frequency channels”. *Neuron* 85.2, pp. 390–401.
- Cao, B, A Yazdanbakhsh, and E Mingolla (2011). “The effect of contrast intensity and polarity in the achromatic watercolor effect”. *Journal of vision* 11.3, pp. 18–18.
- Casarotto, S. et al. (2016). “Stratification of unresponsive patients by an independently validated index of brain complexity”. *Annals of Neurology* 80.5, pp. 718–729.
- Chalmers, D.J. and K.J. McQueen (forthcoming). “Consciousness and the collapse of the wave function”. In: *Consciousness and Quantum Mechanics*. Ed. by S. Gao. Oxford University Press. URL: <https://arxiv.org/abs/2105.02314>.
- Coia, Andrew J. and Michael A. Crognale (2018). “Contour adaptation reduces the spreading of edge induced colors”. *Vision Research* 151, pp. 135–140. ISSN: 0042-6989.
- Coia, Andrew J., Christopher Jones, et al. (2014). “Physiological correlates of watercolor effect”. *Journal of the Optical Society of America A* 31.4, A15–A22. DOI: 10.1364/JOSAA.31.000A15. URL: <http://josaa.osa.org/abstract.cfm?URI=josaa-31-4-A15>.
- Cornelissen, Frans W. et al. (2006). “No Functional Magnetic Resonance Imaging Evidence for Brightness and Color Filling-In In Early Human Visual Cortex”. *The Journal of Neuroscience* 26.14, pp. 3634–3641.
- Craik, David and Kenneth JW Craik (1966). *The nature of psychology*. Cambridge University Press. ISBN: 0521047544.
- DeCarlo, Lawrence T. (2013). “Signal detection models for the same–different task”. *Journal of Mathematical Psychology* 57.1, pp. 43–51. ISSN: 0022-2496.

- Devinck, F, P Delahunt, J Hardy, L Spillman, et al. (2004). “Watercolor Spreading Quantified by Matching and Cancellation”. *Journal of Vision* 4.8, pp. 333–333.
- Devinck, F, P Delahunt, J Hardy, L Spillmann, et al. (2005). “The watercolor effect: Quantitative evidence for luminance-dependent mechanisms of long-range color assimilation”. *Vision Research* 45.11, pp. 1413–1424. ISSN: 0042-6989.
- Devinck, F, Peggy Gerardin, et al. (2014). “Spatial selectivity of the watercolor effect”. *JOSA A* 31.4, A1–A6.
- Devinck, F, J Hardy, et al. (2006). “Illusory spreading of watercolor”. *Journal of Vision* 6.5, pp. 7–7. ISSN: 1534-7362. DOI: 10.1167/6.5.7.
- Devinck, F and K Knoblauch (2012). “A common signal detection model accounts for both perception and discrimination of the watercolor effect”. *Journal of Vision* 12.3, pp. 19–19. ISSN: 1534-7362. DOI: 10.1167/12.3.19.
- Devinck, F and Kenneth Knoblauch (2019). “Central mechanisms of perceptual filling-in”. *Current Opinion in Behavioral Sciences* 30, pp. 135–140.
- Devinck, F and L Spillmann (2009). “The watercolor effect: Spacing constraints”. *Vision Research* 49.24, pp. 2911–2917. ISSN: 0042-6989.
- Doerig, Adrien et al. (2019). “The unfolding argument: Why IIT and other causal structure theories cannot explain consciousness”. *Consciousness and cognition* 72, pp. 49–59.
- Douglas, Rodney J and Kevan AC Martin (2004). “Neuronal circuits of the neocortex”. *Annu. Rev. Neurosci.* 27, pp. 419–451.
- Felleman, Daniel J and David C Van Essen (1991). “Distributed hierarchical processing in the primate cerebral cortex”. In: *Cereb cortex*. Citeseer.
- Fiorani Júnior, M. et al. (1992). “Dynamic surrounds of receptive fields in primate striate cortex: a physiological basis for perceptual completion?” *Proceedings of the National Academy of Sciences of the United States of America* 89.18, pp. 8547–8551.
- Freeman, Jeremy and Eero P Simoncelli (2011). “Metamers of the ventral stream”. *Nature neuroscience* 14.9, pp. 1195–1201.
- Gerardin, P, C Abbatecola, et al. (2018). “Neural circuits for long-range color filling-in”. *NeuroImage* 181, pp. 30–43.
- Gerardin, P, F Devinck, et al. (2014). “Contributions of contour frequency, amplitude, and luminance to the watercolor effect estimated by conjoint

- measurement”. *Journal of Vision* 14.4, pp. 9–9. ISSN: 1534-7362. DOI: 10.1167/14.4.9.
- Gerardin, P, M Dojat, et al. (2018). “Effects of background and contour luminance on the hue and brightness of the Watercolor effect”. *Vision Research* 144, pp. 9–19. ISSN: 0042-6989.
- Hale, Ralph G. and James M. Brown (2021). “Influence of context on spatial expanse of color spreading in the watercolor illusion”. *Attention, Perception, Psychophysics*. ISSN: 1943-393X. DOI: 10.3758/s13414-021-02306-0.
- Hamburger, Kai et al. (2006). “Filling-in with colour: Different modes of surface completion”. *Vision Research* 46.6-7, pp. 1129–1138.
- Harris, Kenneth D and Thomas D Mrsic-Flogel (2013). “Cortical connectivity and sensory coding”. *Nature* 503.7474, pp. 51–58.
- Hermann, L. (1870). “Eine Erscheinung simultanen Contrastes”. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 3.1, pp. 13–15.
- Heydt, Rüdiger von der, Howard S Friedman, and Hong Zhou (2003). “Searching for the neural mechanisms of color filling-in”. *Filling-in: From perceptual completion to cortical reorganization*, pp. 106–127.
- Heydt, Rüdiger von der and Rachel Pierson (2006). “Dissociation of color and figure-ground effects in the watercolor illusion”. *Spatial Vision* 19.2-4, pp. 323–340.
- Kanai, Ryota et al. (2006). “Discrete color filling beyond luminance gaps along perceptual surfaces”. *Journal of Vision* 6.12, pp. 4–4.
- Kanizsa, Gaetano (1979). *Organization in vision: Essays on Gestalt perception*. Praeger Publishers. ISBN: 0030490715.
- Kennedy, J. M. (1981). “Illusory brightness and the ends of petals: change in brightness without aid of stratification or assimilation effects”. *Perception* 10.5, pp. 583–585. DOI: 10.1068/p100583.
- Kitaoka, A (2001). “Illusion designology (7): Illusory light perception that cannot be explained by the Fourier analysis”. *Nikkei Sci* 31, pp. 66–68.
- Kok, Peter et al. (2016). “Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback”. *Current Biology* 26.3, pp. 371–376.
- Komatsu, Hidehiko (2006). “The neural mechanisms of perceptual filling-in”. *Nature Reviews Neuroscience* 7.3, pp. 220–231.

- Komatsu, Hidehiko, Masaharu Kinoshita, and Ikuya Murakami (2000). “Neural Responses in the Retinotopic Representation of the Blind Spot in the Macaque V1 to Stimuli for Perceptual Filling-In”. *The Journal of Neuroscience* 20.24, pp. 9310–9319.
- (2002). “Neural responses in the primary visual cortex of the monkey during perceptual filling-in at the blind spot”. *Neuroscience research* 44.3, pp. 231–236.
- Lamme, Victor (2006). “Towards a true neural stance on consciousness”. *Trends in cognitive sciences* 10.11, pp. 494–501.
- Lee, Tai Sing and My Nguyen (2001). “Dynamics of subjective contour formation in the early visual cortex”. *Proceedings of the National Academy of Sciences* 98.4, pp. 1907–1911.
- Leung, Angus et al. (2020). “Integrated information structure collapses with anesthetic loss of conscious arousal in *Drosophila melanogaster*”. *bioRxiv*. DOI: 10.1101/2020.05.17.090001.
- Lu, Zhong-Lin and Barbara Doshier (2013). *Visual psychophysics: From laboratory to theory*. MIT Press.
- Macmillan, N.A. and C.D. Creelman (2004). *Detection Theory: A User’s Guide*. Taylor Francis. ISBN: 9781135634575.
- Markov, Nikola T et al. (2014). “Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex”. *Journal of Comparative Neurology* 522.1, pp. 225–259.
- Massimini, M et al. (2005). “Breakdown of cortical effective connectivity during sleep”. *Science (New York)* 309.5744, pp. 2228–2232.
- Matsumoto, Masayuki and Hidehiko Komatsu (2005). “Neural responses in the macaque V1 to bar stimuli with various lengths presented on the blind spot”. *Journal of neurophysiology* 93.5, pp. 2374–2387.
- Mayner, W. et al. (2018). “PyPhi: A toolbox for integrated information theory”. *PLOS Computational Biology* 14, pp. 1–21.
- Mims, Callie et al. (2020). “Are perceptual metamers causal metamers? A study of multisensory integration and metacognitive access in cue combination”. *Journal of Vision* 20.11, pp. 724–724.
- Muckli, Lars and Lucy S Petro (2013). “Network interactions: Non-geniculate input to V1”. *Current opinion in neurobiology* 23.2, pp. 195–201.

- O'Brien, Vivian (1959). "Contrast by contour-enhancement". *American Journal of Psychology* 72.2, p. 299.
- Oizumi, M., L. Albantakis, and G. Tononi (2014). "From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0". *PLoS Computational Biology* 10, pp. 1–25.
- Otten, Marte et al. (2016). "The Uniformity Illusion: Central Stimuli Can Determine Peripheral Perception". *Psychological Science* 28.1, pp. 56–68.
- Pinna, B (1987). "Un effetto di colorazione". In: *Il laboratorio e la città. XXI Congresso degli Psicologi Italiani*. Edizioni SIPs, Società Italiana di Psicologia, Milano, p. 158.
- Pinna, B, G Brelstaff, and L Spillmann (2001). "Surface color from boundaries: a new 'watercolor' illusion". *Vision Research* 41.20, pp. 2669–2676.
- Pinna, Baingio and Stephen Grossberg (2005). "The watercolor illusion and neon color spreading: a unified analysis of new cases and neural mechanisms". *JOSA A* 22.10, pp. 2207–2221. ISSN: 1520-8532.
- Ramachandran, Vilayanur S (1992). "Blind spots". *Scientific American* 266.5, pp. 86–91.
- Redies, Christoph and Lothar Spillmann (1981). "The Neon Color Effect in the Ehrenstein Illusion". *Perception* 10.6, pp. 667–681.
- Rockland, Kathleen S and Deepak N Pandya (1979). "Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey". *Brain research* 179.1, pp. 3–20.
- Schachar, RA (1976). "The 'pincushion grid' illusion". *Science* 192.4237, pp. 389–390. DOI: 10.1126/science.1257775.
- Self, Matthew W et al. (2013). "Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation". *Current Biology* 23.21, pp. 2121–2129.
- Tegmark, M. (2016). "Improved measures of integrated information". *PLoS Computational Biology* 12.11, pp. 195–199.
- Tong, Frank and Stephen A Engel (2001). "Interocular rivalry revealed in the human cortical blind-spot representation". *Nature* 411.6834, pp. 195–199.
- Tononi, G. (2004). "An information integration theory of consciousness". *BMC Neurosci* 5.42.
- (2008). "Consciousness as integrated information: a provisional manifesto." *Biol Bull* 215.3, pp. 216–242.

- Tononi, G., M. Boly, et al. (2016). “Integrated information theory: from consciousness to its physical substrate.” *Nat Rev Neurosci* 17, pp. 450–461.
- Tononi, G. and C. Koch (2015). “Consciousness: here, there and everywhere?” *Phil. Trans. R. Soc.* B37020140167.
- Tsuchiya, N. (2017). ““What is it like to be a bat?”—a pathway to the answer from the integrated information theory”. *Philosophy Compass* 12.3, e12407.
- Tuijl, HFJM van (1975). “A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours”. *Acta psychologica* 39.6, 441–IN1.
- Watanabe, I, P Cavanagh, and S Anstis (1995). “Shaded diamonds give an illusion of brightness”. In: *Investigative Ophthalmology & Visual Science*. Vol. 36. 4, S468–S468.
- Weerd, Peter De et al. (1995). “Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma”. *Nature* 377.6551, pp. 731–734.
- Wichmann, Felix A and N Jeremy Hill (2001). “The psychometric function: I. Fitting, sampling, and goodness of fit”. *Perception & psychophysics* 63.8, pp. 1293–1313.
- Wollschläger, Daniel, Antonio M Rodriguez, and Donald D Hoffman (2002). “Flank transparency: The effects of gaps, line spacing, and apparent motion”. *Perception* 31.9, pp. 1073–1092.
- Zipser, Karl, Victor AF Lamme, and Peter H Schiller (1996). “Contextual modulation in primary visual cortex”. *Journal of Neuroscience* 16.22, pp. 7376–7389.

A Appendix: Experimental support for filled/non-filled pairs

A.1 Summary

Filled/non-filled pairs comprise images that appear identical but differ by amount of filling-in they generate. They are created by taking an established filling-in illusion, which we call the filled image, and reproducing the illusory properties perceived in it with another stimulus, called the reconstruction. Then, the reconstruction is added to the initial illusion to create a non-filled image. In the non-filled image, filling-in is not necessary because the reconstruction provides the properties anticipated by the brain that would otherwise be filled in. Importantly, the reconstruction is *tuned* such that the resulting non-filled image appears identical to the filled one.

There are two scenarios that would render filled/non-filled pairs impossible. In the first scenario, filled experiences are just phenomenologically different from non-filled or physically-based experiences. So, reconstructions could never make non-filled images look like filled images. In the second scenario, the reconstruction fails to block filling-in. In other words, filling-in overlays the added reconstruction in the non-filled image. This would inevitably lead to a perceivable difference between the filled and non-filled members of the pair. We can rule out the first scenario by demonstrating that reconstructions can be tuned to match filled-in properties. We can rule out the second by demonstrating that non-filled images made with reconstructions can appear identical to filled images.

To demonstrate that filled/non-filled pairs are possible, we conducted an experiment with two stimuli that induce filling-in: a petals illusion (Figure 5a.i) and a watercolor illusion (Figure 5a.iv).¹⁴ We used two strategies to tune reconstructions for each subject: one in which our participants adjusted the reconstructions to reproduce the filled-in properties they perceived in a filled image (an attribute matching task, see section A.2.5 and Figure 5b) and another in which we showed participants many filled/non-filled pairs with reconstructions at various intensities and asked subjects whether the images were the same or different (the method of constant stimulation or MCS, see section A.2.6 and

¹⁴The petals stimulus was from Kennedy 1981 and the watercolor stimulus was based off of Cao, Yazdanbakhsh, and Mingolla 2011.

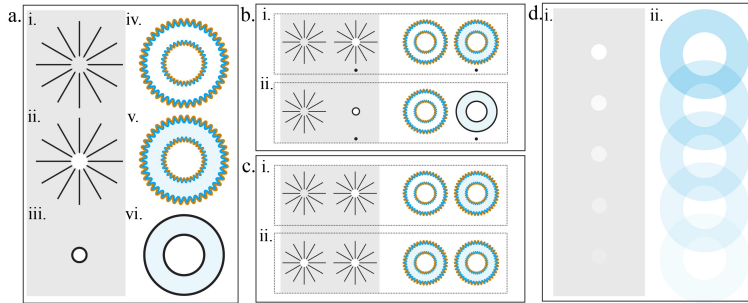


Figure 5: Stimuli and task diagrams. (a): Stimulus types. (i): Petals filled stimulus. (ii): Petals non-filled stimulus. The white circle is the reconstruction, which is tuned to match the illusory properties perceived in (i). (iii): Petals reconstruction with neutral border. (iv): Watercolor filled stimulus. (v): Watercolor non-filled stimulus. The blue annulus is the tunable reconstruction. (vi): Watercolor reconstruction with neutral border. (b): Attribute matching task. Subjects tune a reconstruction (indicated by a black dot) to match illusory properties perceived in a filled stimulus. (i): Inducer-bordered matching trial. Subjects tune a reconstruction within the initial illusion’s inducers (the black petals for petals and the bi-chromatic wavy contour for watercolor). (ii): Neutrally-bordered matching trial. Subjects tune a reconstruction within neutral borders. (c): MCS and validation same-different task pairs. Subjects judge whether the images in a pair are the same or different. (i): Signal trial. Signal here refers to a difference. A filled stimulus is paired with a non-filled stimulus. The latter contains a reconstruction that varies in intensity, making the stimuli different. (ii): Catch trial. Pair of non-filled stimuli with identical reconstructions, making them identical. (d): Examples of petals (i) and watercolor (ii) reconstructions varying in intensity.

Figure 5c). Finally, we tested whether filled/non-filled pairs made with our reconstructions were effective with a validation task, in which subjects judged whether the filled and corresponding non-filled images were the same or different (see section A.2.7 and Figure 5c).

A.2 Methods

A.2.1 Subjects

Eight subjects ages 18 and older and with normal or corrected-to-normal visual acuity participated in the study. All research followed ethical guidelines for research involving human subjects, conducted under the approval of the Chapman University Institutional Review Board. Every subject signed an informed consent form prior to participating.

A.2.2 Apparatus

Stimuli were presented on a 1920 x 1080 ASUS VG248QE display monitor with a refresh rate of 144 Hz using Python version 3.7.6 and the Psychopy package. The experiment was performed in a dark room. The subject’s head position stabilized with a chin rest to ensure the monitor was centered and viewed binocularly with a viewing distance of 60 cm.

A.2.3 Stimuli

Six stimuli were used in the experiment (see Figure 5a). All stimuli were created in Adobe Illustrator®, drawn in Psychopy, and colored in HSV (hue, saturation, and value/brightness) color space.

Stimuli from the petals category were always presented on a light grey background (HSV [0, 0, 0.96]). A single petal is a black (HSV [0, 0, 0]) line (length 1.588° visual angle) with a triangular tip (height 0.053° visual angle) at one end. The complete stimulus comprised 12 petals, each at a unique orientation from 0 to $\frac{11\pi}{6}$ in steps of $\frac{\pi}{6}$ radians such that parallel petals were placed on the same line with tips pointed at one another but separated by 0.70° visual angle. The petals non-filled stimulus (Figure 5a.ii) comprised the same petals with a disk of 0.70° in diameter at the center. The disc is meant to replace filling-in. The petals neutrally-bordered reconstruction stimulus (Figure 5a.iii) was a circle of 0.70° in diameter with a black border 0.2° thick.

Stimuli from the watercolor category were always presented over a white background (HSV [0, 0, 1]). The watercolor stimulus is an annulus with sinusoidally-wavy and bi-chromatic contours. To encourage filling-in, the bands of color making up the contours were contiguous (Devinck and Spillmann 2009), equal in width (Devinck, Peggy Gerardin, et al. 2014), and of high chromatic contrast (Devinck, Hardy, et al. 2006). The outer contour was 4° of visual angle in diameter with a 0.11° wide contour (0.055° for each color band) that undulated at a frequency of 42 cycles per revolution (cpr) and amplitude of 0.32° . The inner contour was 3.35° in diameter with a 0.074° wide contour (0.037° for each band) and undulated at a frequency of 34 cpr and 0.22° in amplitude (Gerardin, Devinck, et al. 2014). The watercolor non-filled stimulus (Figure 5a.v) was an annulus with an inner diameter of 1.74° and an outer diameter of 3.6° , with an initial watercolor stimulus over it. The watercolor neutrally-bordered recon-

struction (Figure 5a.vi) was also an annulus with an inner diameter of 1.74° and an outer diameter of 3.6° but with black contours 0.086° in thickness.

The filled watercolor stimulus came in two kinds: purple fringe (HSV [290, 0.997, 0.6]) with orange inducer (HSV [35, 0.997, 1]) or yellow fringe (HSV [30, 0.98, 0.8]) with blue inducer (HSV [200, 0.997, 1]). These color combinations (Pinna, Brelstaff, and Spillmann 2001) and brightness (V) ratios, both between the contours as well as between each contour and the background, seem to provoke the strongest filling-in (Devinck, Delahunt, Hardy, Spillmann, et al. 2005; Devinck and Knoblauch 2012; Gerardin, Devinck, et al. 2014; Coia, Jones, et al. 2014; Gerardin, Dojat, et al. 2018). To optimize the watercolor effect within-subjects, each participant was presented with both versions at the start of the experiment and asked to choose the stimulus in which they perceived a more opaquely colored ring between the contours (and thus yielded the strongest filling-in). Their choice was used for the duration of the experiment.

The monitor was not color calibrated but we offer the minimum values of brightness and saturation that were perceptible against the petals and watercolor backgrounds, respectively, and also checked that our stimuli induced filling-in. For replication purposes, these elements should suffice and, if desired, more precise colors for watercolor stimuli can be found in the relevant literature (e.g., Gerardin, Devinck, et al. 2014).

A.2.4 Procedure

Subjects adapted to the screen for 2 minutes before starting the experiment. The experiment comprised two iterations (one per stimulus category, petals and watercolor) of three parts: attribute matching, MCS, and validation.

A.2.5 Attribute matching task

During attribute matching (see Figure 5b; Devinck, Delahunt, Hardy, Spillmann, et al. 2005), subjects adjusted inducer-bordered reconstructions within the original illusion (the reconstruction of a non-filled image; Figure 5b.i) and neutrally-bordered reconstructions with black borders (Figure 5b.ii) until they matched the illusory filled-in properties perceived in a filled image. At the end of each trial, subjects reported their satisfaction with the match (i.e., how well they were able to make the reconstruction appear identical to the illusory parts of the filled image) on a Likert scale from 1 (not at all) to 5 (completely).

Subjects were trained on the keyboard controls and then completed 4 main trials (2 inducer-bordered and 2 neutrally-bordered) for each stimulus category, randomly shuffled.

For petals matching, subjects adjusted brightness (V in HSV space) and could freely switch between three step sizes (0.1, 0.01, and 0.001) to optimize the match. For watercolor matching, subjects could change hue, saturation, and brightness, coordinates with the step sizes (30, 10, or 1, for hue and .01, 0.005, or 0.001 for both saturation and brightness). To reduce adaptation to the watercolor patterns, the watercolor stimuli were presented for 2-s intervals, with a 2-s inter-stimulus interval of the white background (Devinck, Delahunt, Hardy, Spillmann, et al. 2005; Devinck, Hardy, et al. 2006; Coia and Crognale 2018). There were no restrictions on match hue, saturation, and/or brightness during this task. Subjects could submit matches that were indistinguishable from the background (henceforth, background-indistinguishable). This ensured we did not bias subjects towards submitting background-distinguishable matches.

Four final reconstructions for each subject to use in filled/non-filled pairs were acquired from the attribute matching task. The first two came from inducer-bordered attribute matching trials. For petals, we will call this the petals IB-reconstruction and for watercolor, the watercolor IB-reconstruction. These reconstructions took either the mean brightness (for petals) of matches on inducer-bordered trials or the mean hue, saturation, and brightness (for watercolor). The second two came from neutrally-bordered matching trials, the petals NB-reconstruction and the watercolor NB-reconstruction. They also took the mean brightness (for petals) or hue, saturation, and brightness (for watercolor) values but from neutrally-bordered trial matches.

A.2.6 Method of constant stimulation (MCS)

Subjects viewed pairs of stimuli comprised of identical or different images and judged whether the images were the same or different (see Figure 5c). For each stimulus category (petals and watercolor), subjects completed 252 trials broken up into 168 signal trials (8 repetitions of 21 intensities) and 84 catch trials (4 repetitions of each intensity), randomly shuffled. Signal trial pairs (Figure 5c.i) were filled/non-filled pairs whose non-filled member had a reconstruction that varied in intensity. Catch trial pairs (Figure 5c.ii) comprised two non-filled images with reconstructions that varied in intensity but in tandem, making them

identical. Petals reconstructions varied in brightness and watercolor reconstructions varied in saturation.

The 21 intensities used were carefully selected during pilot testing ($n = 2$) to range from intensities that were indistinguishable from the background to those that were markedly visible against the background. The petals reconstruction varied in brightness from 0.96 (identical to the background) to 1.0 (highest possible) in steps of 0.002. 0.97 was the minimum brightness that made a grey patch 4° in visual angle distinguishable from the petals background for pilots. The watercolor reconstruction varied in saturation from 0.01 (background-indistinguishable) to 0.052 (clearly distinguishable) in steps of 0.002. The minimum saturation to render a blue patch distinguishable from the watercolor background for pilots was 0.015. Only saturation varied for the watercolor MCS because it was the most heavily-used coordinate during attribute matching piloting and thus seemed to be the most useful HSV coordinate for reconstructing filled-in watercolor. Hue and brightness were set within-subjects to the mean hue and brightness values across attribute matching trials and held constant. The perception-minimums for brightness (0.97, a deviation from the background in brightness of 0.01) and saturation (0.015) from piloting provide cutoffs—above which reconstruction intensities are background-distinguishable and under which all reconstruction intensities are background-indistinguishable.

A free-viewing procedure (Pinna 1987; Pinna, Brelstaff, and Spillmann 2001; Devinck, Delahunt, Hardy, Spillmann, et al. 2005; Devinck, Hardy, et al. 2006; Devinck and Spillmann 2009; Devinck and Knoblauch 2012; Gerardin, Dojat, et al. 2018) was used to encourage judgements based on foveal views of the stimuli. A pair of stimuli was presented simultaneously with only 10° of visual angle separating them so that they could be actively compared, making our procedure more strict with respect to phenomenological identity. The constraints were a 1.5 second viewing limit, which was implemented to provide ample¹⁵ viewing time for subjects to make informed judgements while regulating experiment duration, and a stabilized head position, to avoid viewing angle affecting the appearance of stimuli.

We expected subjects to get better at seeing differences between filled and non-filled images on MCS signal trials as reconstruction intensity increased.

¹⁵Longer than typical presentation times in watercolor filling-in studies e.g., Heydt and Pierson 2006; Cao, Yazdanbakhsh, and Mingolla 2011; Devinck and Knoblauch 2012.

However, we also anticipated a ‘sweet spot’, at which reconstruction intensity was relatively high while the stimuli still appeared identical. We defined the sweet spot as the highest reconstruction intensity for which a subject judged the images to be identical more than 80% of the time. From this, we acquired two more final reconstructions. The petals MCS-reconstruction was set to the brightness sweet spot during petals MCS and the watercolor MCS-reconstruction was set to the saturation sweet spot. The hue and brightness values for the watercolor MCS-reconstruction were set within-subjects to the mean hues and brightness’ from attribute matching main trials—the same values used during the watercolor MCS.

A.2.7 Validation

To test our tuned reconstructions and assess their efficacy in making filled/non-filled pairs, we used a same-different task (Figure 5c). Subjects viewed stimuli in pairs and judged whether the images were the same or different. 280 trials per stimulus category were broken into four groups of 70 trials, including 35 signal and 35 catch trials, randomly shuffled. Signal trial pairs were filled/non-filled pairs whose non-filled image took a previously tuned reconstruction (IB, NB, or MCS) or a control reconstruction, set to a very high intensity. Catch trial pairs were identical images, two non-filled images with identical reconstructions or two filled images.

Validation trials were split into groups based on the reconstruction used. There was an IB-reconstruction group, NB-reconstruction group, MCS-reconstruction, group, and control (CTRL) group. IB, NB, and MCS, signal trials comprised filled/non-filled pairs with previously tuned reconstruction. Their catch trials comprised two non-filled images with identical reconstructions. On CTRL signal trials, the reconstruction used was set to an extremely high intensity, making the images obviously different: dark-grey HSV [0, 0, .85] for petals and either a high-saturation blue HSV [200, 0.10, 1] or orange HSV [35, 0.10, 1] (depending on which watercolor stimulus type the subject was working with) for watercolor. CTRL group catch trial pairs were two filled images.

A.3 Results

A.3.1 Attribute matching task

We offer descriptive statistics for matches from the attribute matching task in Figure 6. Subjects took their time during petals ($M = 14.79$ s; $SD = 11$ s) and watercolor ($M = 38.76$ s; $SD = 25.6$ s) matching and were consistently satisfied with their matches ($M = 4.45$ out of 5; $SD = 0.80$). High satisfaction ratings (see Figure 7a) demonstrate that filled-in properties could be reproduced well with real ones—here, hue, saturation, and value coordinates. Watercolor reconstructions deviated most in saturation on inducer-bordered ($M = 0.008$; $SD = 0.012$) and neutrally-bordered ($M = 0.032$; $SD = 0.027$) matching trials, indicating that saturation is the most important HSV coordinate for reconstructing filled-in watercolor.

We confirm that the petals and watercolor stimuli induced filling-in for our subjects with their matches on neutrally-bordered trials. During these trials, participants were matching the illusion to a neutrally bordered reconstruction (that is, no illusory, filled-in properties should appear on the tunable stimulus). Subjects would only submit matches on these trials that were distinguishable from the background if they were filling-in properties in the original illusion. We thus ran one-sample t-tests comparing reconstruction intensities from neutrally-bordered matching trials for each stimulus category to a hypothetical mean of 0, which we would expect if the fixed, filled stimulus did not induce filling-in and subjects need not submit reconstructions different from the background. Reconstructions from petals neutrally-bordered trials ($M = 0.0273$, $SD = 0.021$) were significantly greater than 0, $t(7) = 3.407$, $p = 0.006$, $d = 1.29$, and so were watercolor reconstructions ($M = 0.031$, $SD = 0.024$), $t(7) = 3.512$, $p = 0.005$, $d = 1.327$).

Subject 6 is an outlier that skews the petals attribute matching results. On their first petals inducer-bordered matching trial, they submitted a match whose brightness was very dark (0.76) relative to the background (0.96). This pulls their average from the other subjects in Figure 7a.i. They gave a satisfaction rating of 2 on that trial and submitted a match with the max brightness (1) on the next inducer-bordered matching trial with a satisfaction on of 5, so this was likely a mistake. Without this subject, the group-level reconstruction intensities petals inducer-bordered trials ($M = 0.0208$, $SD = 0.014$) shift to larger

	Trial type	Petals		Watercolor (blue)		Watercolor (orange)	
		IB	NB	IB	NB	IB	NB
	Count	16	16	12	12	4	4
Hue (H)	Minimum	-	-	180.000	70.000	30.000	35.000
	Median	-	-	200.000	200.000	35.000	35.000
	Max	-	-	210.000	210.000	35.000	50.000
	Mean	-	-	194.167	184.167	33.750	38.750
	SD	-	-	9.962	37.285	2.500	7.500
Saturation (S)	Minimum	-	-	0.000	0.000	0.000	0.000
	Median	-	-	0.000	0.046	0.000	0.009
	Max	-	-	0.039	0.095	0.010	0.020
	Mean	-	-	0.010	0.039	0.003	0.010
	SD	-	-	0.013	0.027	0.005	0.011
Value (V)	Minimum	-0.205	-0.030	-0.010	-0.015	0.000	0.000
	Median	0.025	0.040	0.000	0.000	0.000	0.000
	Max	0.040	0.040	0.000	0.000	0.000	0.000
	Mean	0.008	0.027	-0.001	-0.003	0.000	0.000
	SD	0.059	0.024	0.003	0.005	0.000	0.000

*Note. All values are deviations from the background

Figure 6: Attribute matching task descriptive statistics. All trials are included (2 trials per subject per stimulus category). The background hue, saturation, and/or brightness, coordinates were subtracted from the reconstruction hue, saturation, and/or brightness coordinates, respectively, to calculate deviations from the background. For petals, the background brightness of 0.96 was subtracted from reconstruction brightness. For watercolor, the background hue (0), saturation (0), and brightness (1) were subtracted from reconstruction hue, saturation, and brightness, values.

values. To ensure that no significant relationships we found were the result of this outlier, we ran all analyses with them excluded. Significant relationships remained.

A.3.2 Method of constant stimulation (MCS)

MCS results differed for petals and watercolor. On signal trials (see Figure 7b.i and b.ii), the proportion of same responses did not vary on petals trials across intensities but did on watercolor trials. Subjects got better at seeing differences between filled and non-filled images as reconstruction intensity increased for watercolor but not for petals. Reconstruction intensity was not correlated with the proportion of same judgements during petals MCS ($r = -0.026$, $p = 0.738$) while watercolor reconstruction intensity was strongly, negatively correlated with proportion of same judgements ($r = -0.713$, $p < 0.001$).

A psychometric function is a mathematical function (typically a sigmoid) mapping a stimulus level (here, petals reconstruction brightness intensity and watercolor reconstruction saturation intensity) to a forced-choice response level (here, same or different), see Wichmann and Hill 2001. To model the relationship between the intensity of a reconstruction in a non-filled image and whether it looks identical to a filled image, we attempted to fit sigmoid functions to the petals and watercolor MCS signal trial data. A psychometric function could be

fit for watercolor MCS such that the majority of the variance was explained, $r^2 = 0.54$ (see Figure 7b.ii). However, no sigmoid or polynomial below five degrees with an $r^2 > 0.01$ could be fit for petals.

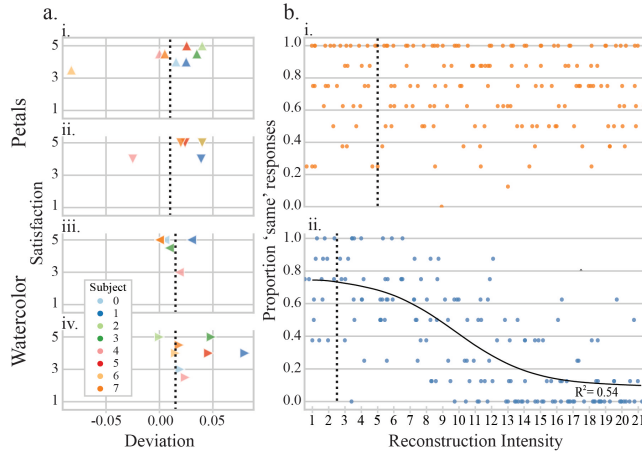


Figure 7: Attribute matching task and MCS results. The black dashed lines on every plot mark the minimum intensities perceptible against the background (for petals, brightness 0.01 and saturation 0.015 for watercolor). (a): Attribute matching task reconstruction deviations and corresponding satisfaction ratings. There is one point for each subject per plot, which is their mean reconstruction intensity and satisfaction rating for the trial type. (i): Petals inducer-bordered matching trial. (ii): Petals neutrally-bordered matching trial. (iii): Watercolor inducer-bordered matching trial. (iv): Watercolor neutrally-bordered matching trial. (b): MCS results. Proportion of same judgements on signal trials were calculated for each subject at each stimulus intensity, resulting in one data point per subject at each intensity in both plots. (i): Proportion of same responses for each intensities during petals MCS signal trials. The 21 brightness intensities ranged from 0.96 to 1 in steps of 0.002. (ii): Proportion of same judgements across intensities for watercolor MCS signal trials. A psychometric model (black sigmoid) is fit to the data with an $r^2 = 0.54$. The 21 intensities ranged from 0.01 to 0.052 in steps of 0.002.

A.3.3 Final reconstructions

Final reconstructions (IB, NB, and MCS) took distinct but similar intensities for each subject (see Figure 8). Most petals NB-reconstructions ($M = 0.027$, $SD = 0.0212$) and MCS-reconstructions ($M = 0.034$, $SD = 0.011$) were close to the maximum V possible of 1. Watercolor IB-reconstructions ($M = 0.008$, $SD = 0.01$) were close to 0, the lowest saturation possible. All watercolor IB-reconstructions fell below the minimally background-distinguishable saturation of 0.015, except for subjects 1 (S of 0.03) and 4 (S of 0.019). Conversely, all watercolor NB-reconstructions fell above the cutoff except for subject 2 (S of 0).

Low IB-reconstructions suggests subjects filled-in the fixed and tunable stimuli equally during the attribute matching task. High NB-reconstructions reflect filling-in. Background-distinguishable MCS-reconstructions indicate that our procedure worked well in finding relatively high intensity reconstructions for which filled and non-filled images still appear identical.

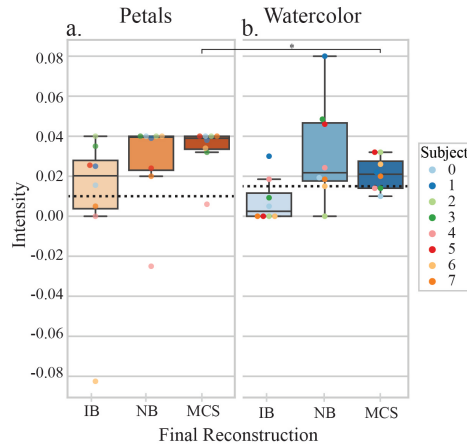


Figure 8: Final reconstructions. Box-and-whisker plots display three quartile values (25%, 50%, and 75%) for the distributions of final reconstructions with whiskers that extend to points that lie 1.5 times the interquartile range of the lower and upper quartiles. Values that fall outside of the whiskers are outliers. The black, dotted, horizontal lines show the minimum background-distinguishable brightness and saturations. (a): Petals final reconstructions. (b): Watercolor final reconstructions.

A.3.4 Validation

For each validation trial group (IB, NB, MCS, and CTRL), z -transformed proportions of false alarms (judging stimuli to be ‘different’ when they were the same) were subtracted from z -transformed proportions of hits (judging stimuli as ‘different’ when they were) and then used to calculate d' —a measure of discriminability—in accordance with the independent-observation decision rule (Macmillan and Creelman 2004; DeCarlo 2013; Lu and Doshier 2013).¹⁶ Under this rule, a d' of 0.16 indicates subjects were performing at chance-level (50% correct) in their judgements, a d' of 1 is equivalent to 57% correct, and a d' of

¹⁶The independent-observation decision rule is considered the optimal rule for deriving d' -primes in a same-different task. Tables in Macmillan and Creelman 2004 were used for z -transformations and finding corresponding d' -primes based on the independent-observation rule.

2 is 73% correct. Lower d-primes reflect that subjects could not see differences between filled and non-filled images while higher d-primes indicate that they could.

Subjects almost never judged non-filled images with the control reconstruction as identical to filled images (see Figure 9a and b). Both petals ($M = 5.288$, $SD = 1.368$) and watercolor ($M = 5.161$, $SD = 1.211$) control pairs were associated with high d-primes (see Figure 9c and d). Our control is thus useful for comparing the discriminability of filled/non-filled pairs made with our final reconstructions (IB, NB, and MCS) to a pair of obviously different images.

Filled/non-filled pairs using our final reconstructions had low d-primes ($M = 1.631$, $SD = 1.368$) on average. At the group level, filled/non-filled pairs made with petals final reconstructions worked. Most petals final reconstructions were background-distinguishable and the resulting pairs were judged to be identical for most subjects on most trials (see Figure 9a). D-primes for petals filled/non-filled pairs made with MCS-reconstructions did not differ significantly from chance-level d' (0.16), $t(7) = 1.811$, $p = 0.113$, but did when made with IB-reconstructions, $t(7) = 2.437$, $p = 0.045$, and NB-reconstructions, $t(7) = 4$, $p = 0.005$, trials. Even so, neither d-primes from IB, $t(7) = 0.868$, $p = 0.207$, nor from NB trials, $t(7) = 1.625$, $p = 0.074$, were significantly greater than 1. So, subjects were performing at or close to chance-level when viewing their final petals filled/non-filled pairs.

At the group-level, we were less successful in creating filled/non-filled pairs for watercolor. Only d-primes for filled/non-filled pairs made with IB-reconstructions did not differ significantly from chance (d' of 0.16), $t(7) = 2.022$, $p = 0.083$, and most of these reconstructions were background-indistinguishable. NB-reconstructions ($t(7) = 3.943$, $p = 0.006$) and MCS-reconstructions ($t(7) = 5.651$, $p = 0.001$) d-primes differed significantly from chance. Looking at individual subjects, however, two had NB and MCS reconstructions that were background-distinguishable and could be used to make filled/non-filled pairs that worked (had low d-primes). NB-reconstruction filled/non-filled pairs worked for subjects 0 (saturation = 0.019, $d' = 1.42$) and 6 (saturation = 0.015, $d' = 1.5$). MCS-reconstruction filled/non-filled pairs worked for subjects 6 (saturation = 0.26, $d' = 1.060$) and 7 (saturation = 0.02, $d' = 1.47$). Interestingly, subjects almost never exclusively judged filled/non-filled images to be different on watercolor validation signal trials (see Figure 9b). So, watercolor filled/non-

filled made with NB and MCS reconstructions worked for two subjects and for the others, they worked at least on some if not on most trials.

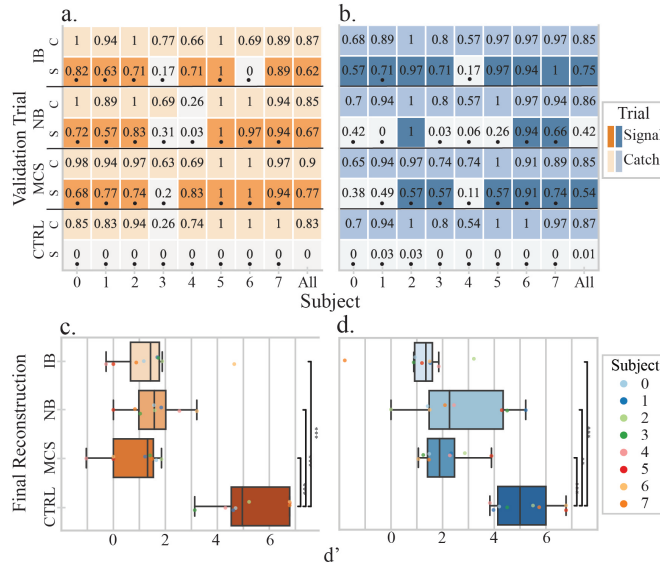


Figure 9: Validation task results. (a-b): ‘Same’ judgements during the validation task. The grand mean (across subjects) is included, called subject ‘All’ on the x-axis. Proportion of same judgements for each subject on each validation trial type. If the tile is colored, subjects judged the presented stimuli to be identical the majority of trials (more than 50% or a proportion ≥ 0.50). The black dots mark whether a subject’s final reconstruction for that trial type was background-indistinguishable (brightness ≥ 0.01 for petals and saturation ≥ 0.015 for watercolor). (c-d): Box-and-whisker plots of the distributions of d-primes for each final reconstruction type. Significant relationships are indicated by black bracket(s) and asterisk(s). (c): Petals d-primes and final reconstruction intensities. (d): Watercolor d-primes and final reconstruction intensities.

A.3.5 Conclusion

In sum, for some subjects and on many trials for those subjects, non-filled images with background-distinguishable reconstructions appeared identical to filled images. We were able to make non-filled images with background-distinguishable reconstructions that were indistinguishable from the original petals illusion for most subjects. The filled/non-filled pairs created for the watercolor stimulus only worked well (had both background-distinguishable reconstructions and low d-primes) for two subjects. However, the subjects for which the filled/non-filled pairs appear to do less well still judged non-filled images to be identical to filled images on many trials. Subjects thus judged non-filled images with background-

distinguishable reconstructions to be identical to filled images from some to most of the time. We conclude that filled/non-filled pairs are feasible, confirming our premise that images that induce different amounts of filling-in can look identical.

B Appendix: additional filled/non-filled pairs

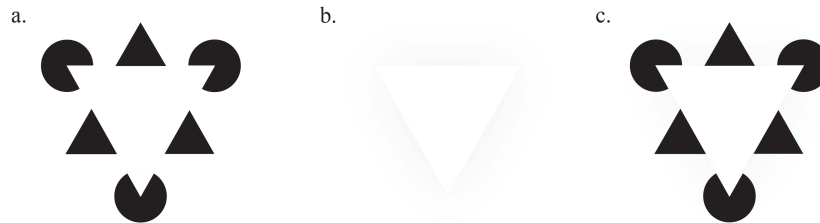


Figure 10: Kanizsa filled/non-filled pair. (a): Filled Kanizsa figure. A bright white triangle overlaying the black elements or inducers is filled-in. (b): Reconstruction. A white triangle with a gray extended shadow cast along its edges is made to match the illusory triangle perceived in (a). (c): Non-filled Kanizsa figure. This image adds (b) to (a) to create a non-filled Kanizsa figure. (a) and (c) are members of the the Kanizsa figure filled/non-filled pair.

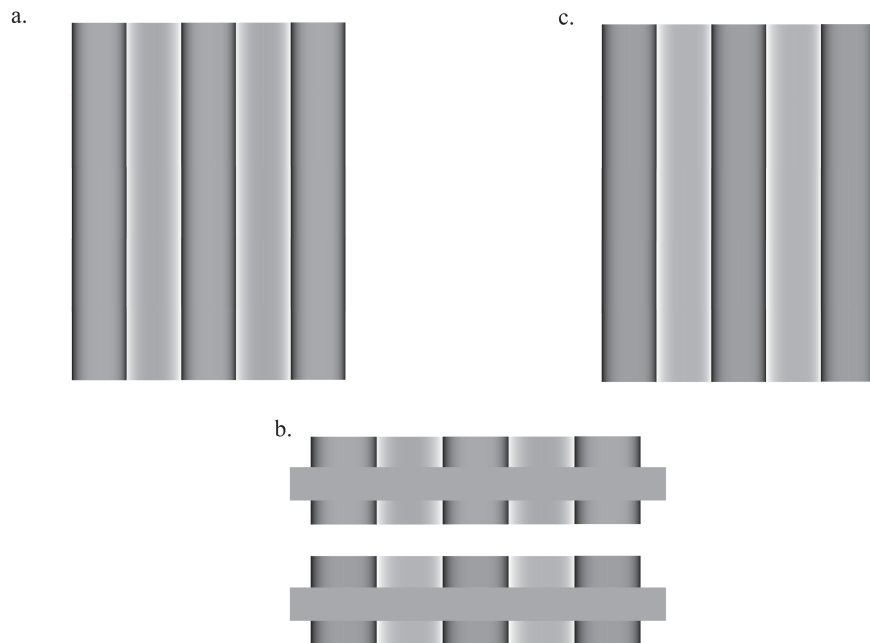


Figure 11: Panel Craik-O'Brien-Cornsweet illusion filled/non-filled pair. (a): Filled panel version of the COBC. The center of each panel appears darker or lighter if their edges are dark or light gradients, respectively. (b): Reconstruction. The pairs of panels on the top come from (a). The pairs of panels on the bottom are darkened or lightened if the gradients at their edges are dark or light, respectively. The luminance is adjusted such that the panels qualitatively match. However, solid-gray bars the shade of gray at the center of panels on the top overlay both rows of panels to demonstrate that panels on the top are identical at the center, while panels on the bottom differ in luminance. (c): Non-filled panel version of the COBC. This figure is made up of panels from the bottom of (b), with the center of each panel being actually darker or lighter if their edges are dark or light gradients, respectively. (a) and (c) constitute a filled/non-filled panel version of the COBC

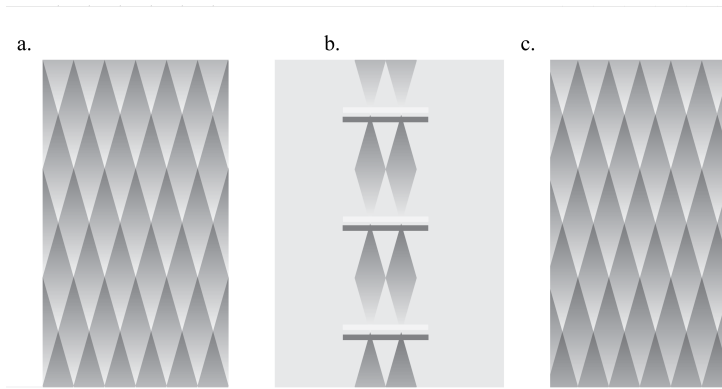


Figure 12: Filled/non-filled COBC pair. (a): Filled COBC illusion. (b): Reconstruction. The left column of diamonds come from image (a), while darker shades are added to the right column of diamonds such that diamonds darken as they descend. A grey bar is placed over each column of diamonds in hopes it will emphasize the differences in brightness between the columns. (c): Non-filled COBC. Diamonds from the left column in (b) are used to reproduce (a). Rather than darker shades being filled-in, real darker shades are present in the image. Images (a) and (c) illustrate how filled/non-filled pairs can be constructed.

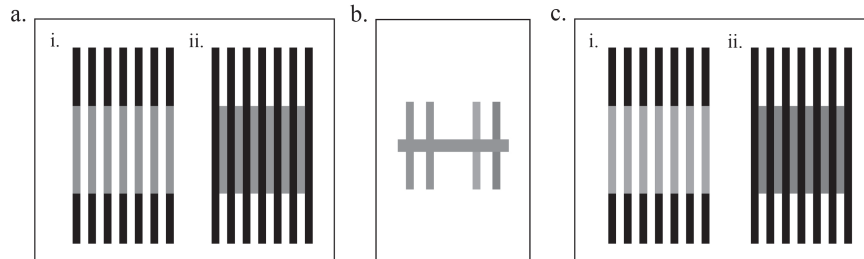


Figure 13: White's effect filled/non-filled pair. (a): Filled White's effect. The gray lines overlapping the black lines in (i) appear lighter than the gray lines between the black lines in (ii) due to filling-in. (b): Reconstruction. One gray line from (a.i) and another gray line from (a.ii) make up the pair of lines on the left. In order to reconstruct the illusory difference in luminance between the lines perceived in (a), two new gray lines are made, one lighter than those in (a) and one darker. A gray bar of the same luminance as the gray lines in (a) is placed over all four lines to demonstrate that pair of lines from (a) are identical, while the right pair of lines differ in luminance (the left is lighter than the overlapping bar and the right is darker than the bar). (c): Non-filled White's effect. The right pair of gray lines in (b) are used to create a non-filled version of (a). Lighter gray lines are placed over the black lines in (i) and dark gray lines are placed between black lines in (ii). (a) and (c) form the White's effect filled/non-filled pair.

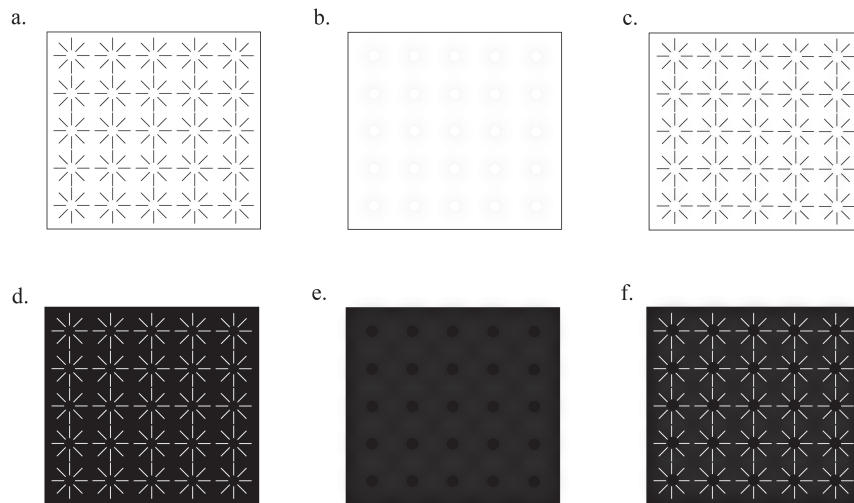


Figure 14: Ehrenstein figures filled/non-filled pair. (a): Filled Ehrenstein figures in white. Illusory bright white circles are perceived at the center of each Ehrenstein figure. (b): Reconstruction. Real white circles are placed in the same location in visual space as the illusory circles perceived in (a) and a light-gray shadow is cast along each circle's edges. (c): Non-filled Ehrenstein figures image. Adding (b) to (a) allows us to create a non-filled version of the white Ehrenstein figures. This image is designed to physically possess the illusory properties perceived in (a). (d): Filled Ehrenstein figures in black. Black circles are perceived at the center of each Ehrenstein figure. (e): Reconstruction. A set of black circles with gray shadows are created to match the feel and location of the illusory circles perceived in (d). (f): Non-filled Ehrenstein figures in black. In a similar manner to (c), (e) is added to (d) to create non-filled set of Ehrenstein figures in black. The filled/non-filled pairs in this image include (a)/(c) and (d)/(f).

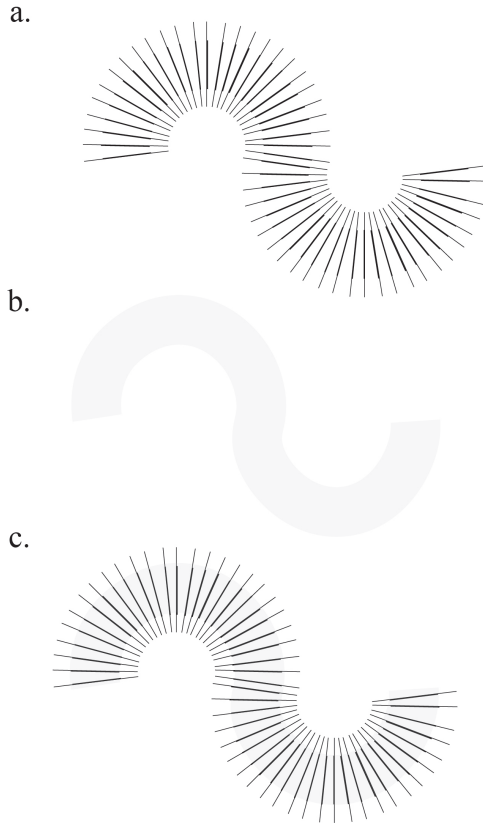


Figure 15: Achromatic color spreading filled/non-filled pair. (a): Filled achromatic color spreading (Wollschläger, Rodríguez, and Hoffman 2002). A thick black line is placed over each black line. An illusory spreading of gray color between the lines prompts the perception of an illusory overlaying gray figure. (b): Reconstruction. The illusory gray surface or figure perceived in (a) is isolated. (c): Non-filled version of achromatic color spreading. (b) is placed over (a) to create an image actually containing the illusory gray figure perceived in (a). (a) and (c) constitute a filled/non-filled pair.

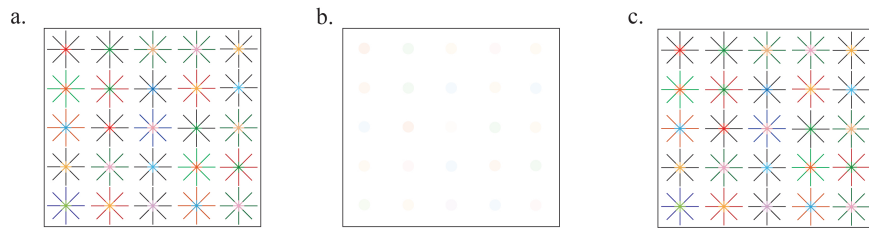


Figure 16: Neon color spreading with Ehrenstein figures filled/non-filled pair. (a): Filled neon color spreading Ehrenstein figures. Circles or discs taking the color of the lines at the center of each Ehrenstein figure are filled-in. (b): Reconstruction. The illusory circles perceived in (a) are reconstructed by filling real circles with the same color as the inducing lines, but a lower opacity. (c): Non-filled neon color spreading Ehrenstein figures. (b) is added to (a) to create the non-filled version. (a) and (c) form a filled/non-filled pair pair.

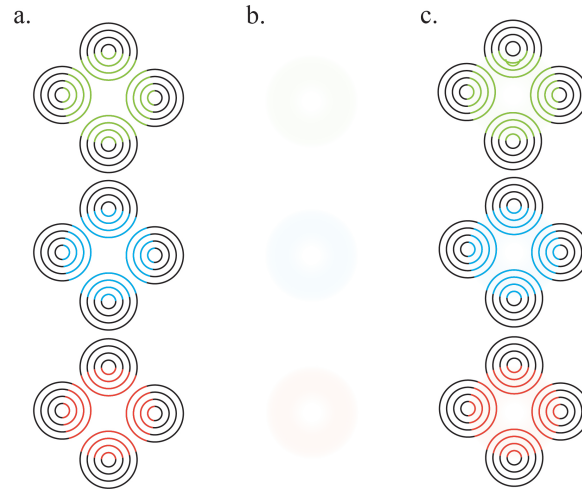


Figure 17: Neon color spreading filled/non-filled pair. (a): Filled neon color spreading image (Baingio Pinna and Grossberg 2005). Low-opacity circles overlapping each set of four black rings are perceived, taking the hue of the colored lines. (b): Reconstruction. Real circles are filled with the illusory colors of circles perceived in (a). (c): Non-filled neon color spreading image. This figure adds (b) to (a) to create an image with veridical properties that match the illusory properties perceived in (a). (a) and (c) are members of the neon color spreading filled/non-filled pair.

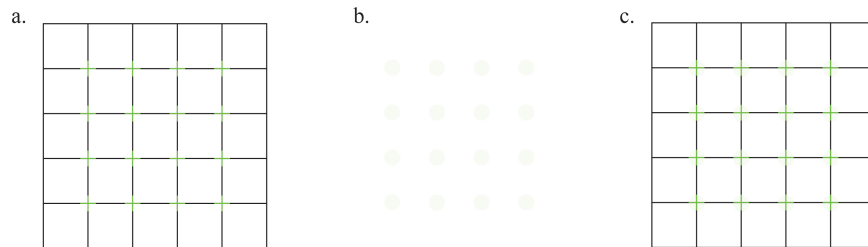


Figure 18: Pincushion grid filled/non-filled pair. (a): Filled pincushion grid image (Schachar 1976). Another kind of neon color spreading. Short colored lines are added to each intersection of black lines. Illusory circles of the same color as the lines at each intersection are perceived. (b): Reconstruction. Real circles are filled with the hue of the lines overlaying each intersection in (a). (c): Non-filled pincushion grid image. This figure adds (b) to (a) to create an image with veridical properties to match the illusory properties perceived in (a). (a) and (c) are members of the pincushion grid filled/non-filled pair.

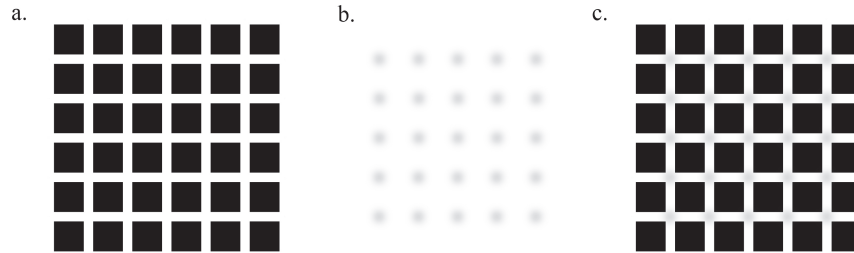


Figure 19: Hermann grid illusion filled/non-filled pair. (a): Filled Hermann grid illusion (Hermann 1870). Faint or blurred black circles appear and disappear between corners of each black square as the eyes explore the image. (b): Reconstruction. Veridical blurred and low-opacity black circles are reconstructed to match the illusory circles perceived in (a). These can be made to sporadically appear or disappear to match (a) more precisely. (c): Non-filled Hermann grid illusion. (b) is added to (a) to create a Hermann grid illusion that actually possess the blurred black circles between corners of each black square. (a) and (c) are members of the Hermann grid illusion filled/non-filled pair.

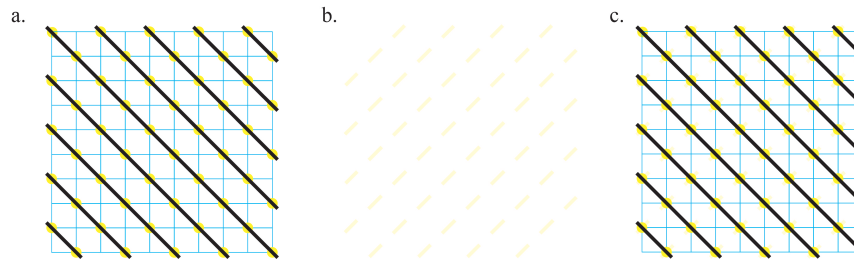


Figure 20: Colored-ray illusion filled/non-filled pair. (a): Filled colored-ray illusion (Kitaoka 2001). Illusory yellow dashes perpendicular to the black lines in the image are perceived. (b): Reconstruction. Dashed yellow lines perceived in (a) are isolated. (c): Non-filled colored-ray illusion. This figure adds (b) to (a) to create an image with veridical properties to match the illusory properties perceived in (a). (a) and (c) are members of the colored-ray illusion filled/non-filled pair.

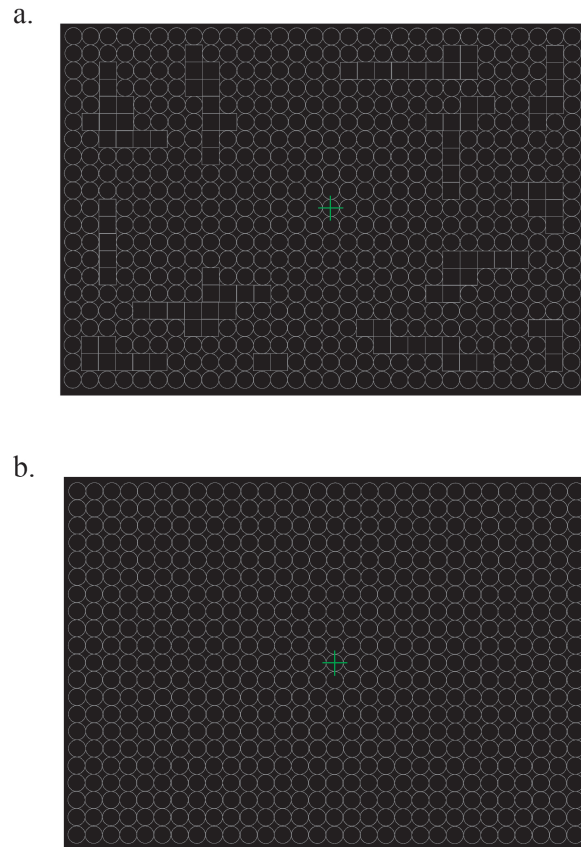


Figure 21: Uniformity illusion filled/non-filled pair. (a): Filled uniformity illusion Otten et al. (2016). Upon steady fixation at the center of the image for several seconds, the squares in the periphery become circles and a uniformly patterned surface is perceived across the image. (b): Non-filled uniformity illusion. A uniform surface of circles to match the illusory uniform surface perceived in (a). (a) and (b) are members of the uniformity illusion filled/non-filled pair.

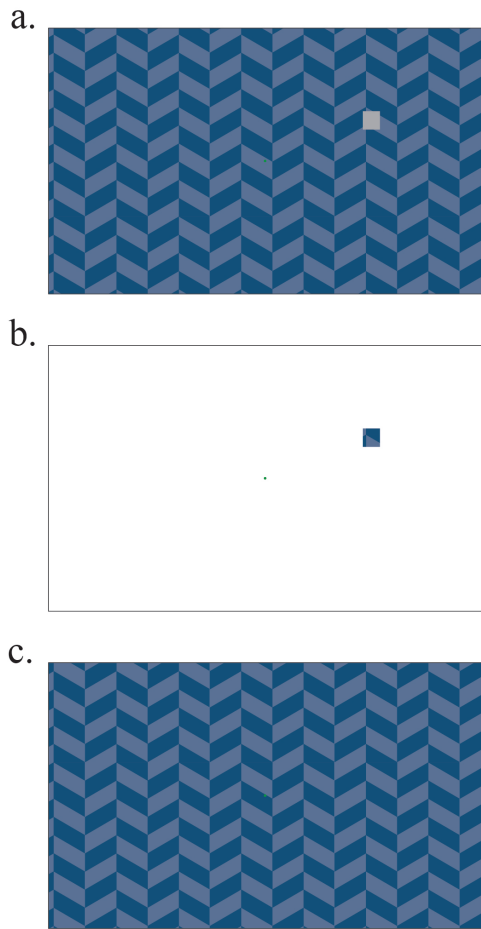


Figure 22: Artificial scotoma stimuli filled/non-filled pair. (a): Filled artificial scotoma stimuli Ramachandran (1992) and Weerd et al. (1995). Upon steady fixation at the center green dot at the center of the stimuli for several seconds, the gray square in the periphery fills in with the surrounding pattern. (b): Reconstruction. A gray square filled with the pattern in (a) is created and placed in the same spatial location as the gray square is in (a). (c): Non-filled artificial scotoma stimuli. (b) is added to (a) to make a completely uniform surface that does not induce filling-in. (a) and (c) make up an artificial scotoma stimuli filled/non-filled pair.

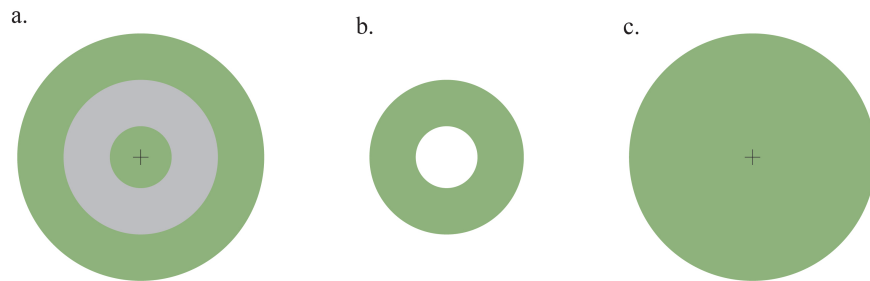


Figure 23: Hamburger stimuli filled/non-filled pair. (a): Filled Hamburger image (Hamburger et al. 2006). Upon steady fixation on the cross at the center of the rings, the middle gray ring fades into a green matching the center disc and the outer ring of the image. (b): Reconstruction. A ring in the same spatial location as the gray ring in (a) is colored the same green as the inner disc and outer ring in (a). (c): Non-filled Hamburger stimuli. (b) is added to (a) to create a circle uniformly filled with the green that (a) uniformly fills. (a) and (c) make up the artificial scotoma filled/non-filled pair.

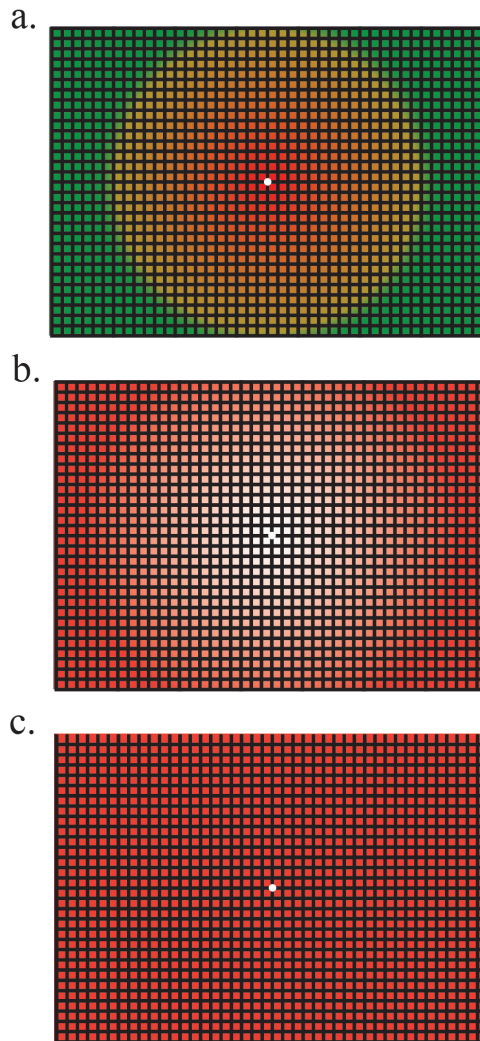


Figure 24: Discrete color filling-in filled/non-filled pair. (a): Discrete color filling-in stimulus (Kanai et al. 2006). When the center of the stimulus is gazed at steadily, the color red near fixation appears to fill the peripheral squares, replacing the green (called 'filling-out'). This is an example of filling-in in the face of strong luminance gaps (black lines). (b): Reconstruction. The illusory spread of red across the black squares in the periphery is isolated. (c): Non-filled version of the discrete color filling-in stimulus. (b) is placed over (a) to create a uniformly colored surface with only red squares. This stimulus is designed to match the percept that (a) converges to after steady fixation. (a) and (c) constitute a filled/non-filled pair for the discrete color filling-in stimulus.