Philosophy and
Phenomenological Research

Philosophy and Phenomenological Research
Vol. XCIII No. 2, September 2016
doi: 10.1111/phpr.12258

The Anna Karenina Principle and Skepticism about Unconscious Perception

NED BLOCK
New York University

All conscious perceptions are alike but each unconscious perception is unconscious in its own way. This "Anna Karenina" (Tolstoy, 1901, first sentence) principle concerning unconscious perception (Block, 2011) holds because conscious perception is a matter of oscillating feed-forward and feed-back signals that are uniform in the essential features of their successes but variegated in their failures. Breitmeyer (2015) describes 24 methods of producing unconscious perception that work by interfering with the oscillations in different ways.

Why do I start with the Anna Karenina principle in responding to Phillips' critique? Phillips criticizes inferences from studies of two brain damage syndromes, visuo-spatial neglect and blindsight. His criticisms must be understood in the context of his wider campaign against unconscious perception (Block & Phillips, forthcoming; Phillips, 2016 (this issue); forthcoming). Although there is no scientific certainty about any of the paradigms of unconscious perception, the critiques are an ad hoc group with no unity—except the superficial unity of "not perception" or "not unconscious". But what is the likelihood that each of the 24 paradigms is subject to its own fatal flaw? All of the 24 paradigms have passed the test of peer review, and in many cases have been subject to many years of intense scrutiny and subsequent refinement of methods by former critics (Klauer, Eder, Greenwald, & Abrams, 2007). Without some unified reason for skepticism, the plausibility that something different is wrong with each of the 24 methods is not high.

All experimental paradigms in psychology have weaknesses (and very often, also strengths), but studies of unconscious perception are problematic in a special way. We do not have a scientific account of what consciousness

is, and without such an account, all cases of unconscious perception lack scientific proof. Of course the same is true of conscious perception: we cannot prove scientifically that you are having conscious visual experiences while reading this. This is not to say, however, that there is no science of conscious and unconscious perception.

We have many theories of what consciousness is in the brain and none of those that are taken seriously by substantial numbers of working neuroscientists apply to bees or spiders. (You can see what working scientists think of panpsychism here: (Block et al., 2014)). For example there is no evidence of anything approximating a "global neuronal workspace" in arthropods. So we have some—far from decisive—scientific reason to believe that spiders and bees have no conscious states. Yet we know that both bees and spiders have sophisticated perceptions that involve objective perceptual representations of the environment. For example the jumping spider visually identifies, tracks and ambushes its prey (Burge, 2010).

I have never regarded the two syndromes Phillips discusses as among the most plausible cases of unconscious perception. As Phillips mentions, I have speculated that in a variant of neglect, there might be phenomenal consciousness without access consciousness. Why then did I introduce neglect as providing dramatic evidence of unconscious perception? The paper that Phillips criticizes was concerned with the issue of conscious seeing without attention, and not devoted to a discussion of conscious and unconscious perception. I had severe word limits (as I do here), but (as here) there was no explicit limit for captions. The argument discussed by Phillips appears entirely in the captions to figures 8 and 9 of that article. Although visuo-spatial neglect does not provide the best case for unconscious perception, it does have the benefit of being presentable very briefly and—if it really is a case of unconscious perception—it allows for a neat case of a single percept that combines conscious perception and unconscious perception. In a longer treatment I would have mentioned some of the problems in interpreting clinical syndromes.

Phillips' focus in this article is on whether the allegedly unconscious states in neglect and blindsight are really weakly conscious and whether the subjects' claims not to see the items on the left (in the case of left sided neglect) and the items in the blind field (in the case of blindsight) are really cases of a "conservative response criterion", in effect, the subjects are reluctant to say they see something unless they are very very sure they have seen it. It is always difficult to know what to think about what brain-damaged subjects tell us about the presence or absence of conscious experience. There is always a worry that one of the things affected by the brain damage is the cognitive processing underlying the subjects' reports.

However, the 24 ways of producing unconscious perception referenced above concern only neurotypical subjects. Of course none of the 24 methods is without controversy. One that has something going for it is

"continuous flash suppression" or CFS. Different stimuli are presented to each eye. One eye gets a "Mondrian", a variegated colorful stimulus in which the colors change 10 times per second. The Mondrian fills the whole visual field for a number of seconds. That is, if there is a face or a house presented to the non-Mondrian eye, for a number of seconds (or even minutes) subjects will not be aware of the face or house (though in some circumstances they can be aware of blotches of color from the face or house). When this method is used at its best, subjects are at chance in making a choice between, say, a face and a house projected to the eye that does not get the Mondrian and are at chance in saying whether anything at all is projected to that eye. A further line of evidence for lack of awareness used in some studies is: there is no correlation between the subjects' confidence in whether they have seen a face or a house and what they have actually seen—i.e. confidence does not predict accuracy (Raio, Carmel, Carrasco, & Phelps, 2012). Not only do they have no knowledge of what they are seeing, when they do think they might know what they are seeing, actually they are just guessing. This lack of correlation rules out the most obvious forms of a "criterion" effect of the sort that Phillips discusses. And when this method is used at its best, subjects insist they are seeing nothing other than the Mondrian. Indeed, I insist I am seeing nothing other than the Mondrian for at least a few seconds—and as long as several minutes—as have many students in classes where I have demonstrated this phenomenon (with presentation methods far cruder than those used in a laboratory).

But is what is unconscious in these cases really seeing—as opposed to the unconscious sub-personal processing underlying seeing? In many cases, unconscious perceptual processing reflects personal level values and understanding. For example, subjects' attention was attracted or repelled by unconsciously processed nudes in a way appropriate to the subjects' self-reported gender preferences (Jiang, Costello, Fang, Huang, & He, 2006). For example, the attention of most of the heterosexual males was attracted by female nudes and repelled by male nudes. Heterosexual females (and homosexual males) were more likely to have their attention attracted by both (except for two individuals who were repelled by both).

What about my claim of perceptions that are partly conscious and partly unconscious? (This is relevant to whether conscious and unconscious perception, despite their differences, are the same kind of state.) Such perceptions can occur in CFS when subjects perceive a combination of a face seen unconsciously and patches of color from the face seen consciously (Mudrik, Gelbard-Sagiv, Faivre, & Koch, 2013; Mudrik & Koch, 2013).

I will now shift to a better example of a percept that combines conscious and unconscious seeing. I choose this in part because it is very different from the one just described and the ones discussed by Phillips, illustrating the Anna Karenina Principle. The example of CFS just discussed depends

on competition between the two eyes. But that competition can be subverted by brief—under 100 ms—low contrast presentations of different stimuli to the two eyes, where the presentations are repeated with gaps of over 100 ms. In that circumstance, the different stimuli to the two eyes are fused into a single conscious percept. Remarkably, fused complementary colors—in this case, red and green—yield a conscious percept of a different color—in this case, yellow. And when the fused items are red and green faces on complementary backgrounds, what the subject consciously sees is a uniform yellow field (Moutoussis & Zeki, 2002).

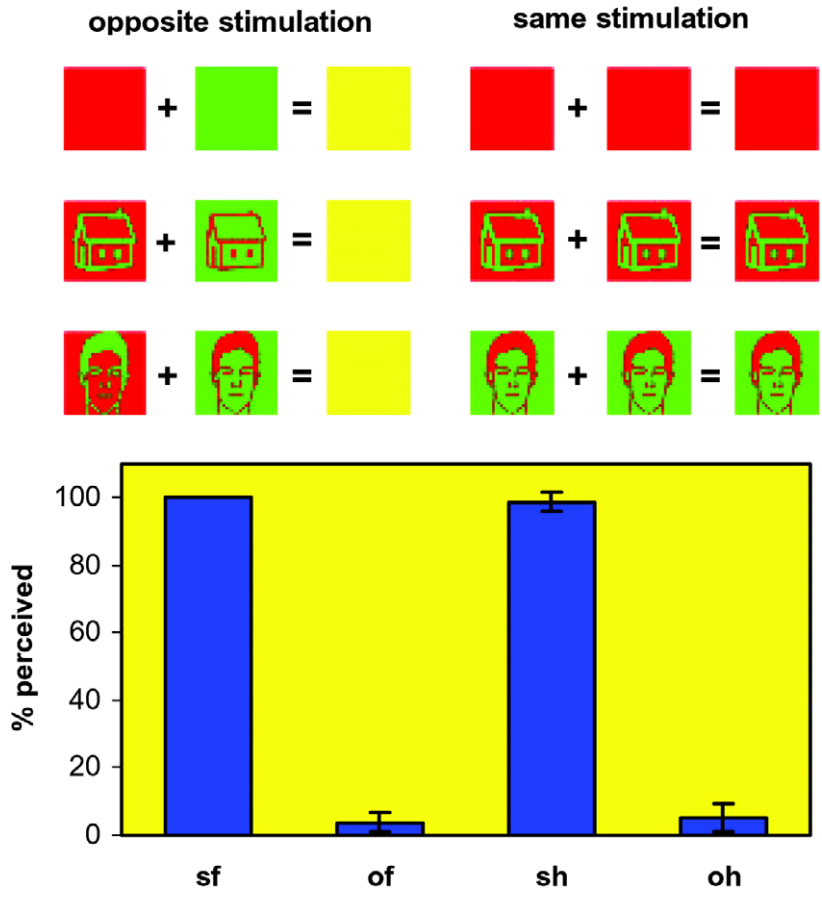

Figure 1. Diagram from (Moutoussis & Zeki, 2002). If you are reading this in a grayscale form you can see a color version that is not behind a paywall at http://www.pnas.org/content/99/14/9527.figures-only or at http://www.nyu.edu/gsas/dept/philo/faculty/block/figures/Zeki.png. The copyright for this image is held by the original publisher, PNAS.

Given the controversy over replicability of psychological experiments, I should mention that this is a robust phenomenon that has been replicated a number of times with somewhat different methods (Fahrenfort et al., 2012; Fogelson, Kohler, Miller, Granger, & Tse, 2014; Schurger, Pereira, Treisman, & Cohen, 2010; Schurger, Sarigiannidis, Naccache, Sitt, & Dehaene, 2015).

Brief repetitive presentations of stimuli in a binocular fusion experiment are diagrammed. The top left row shows a presentation of a red square to one eye and a green square to the other eye. When the squares are at the same retinal locations, the effect on conscious perception is that the subject sees a uniform yellow square. If, as on the top right, both eyes get red squares, then the conscious percept is as of a red square. Moving to the second row on the left: if one eye gets a green house on a red background and the other eye gets a red house on a green background, then the fused conscious percept is as of a yellow square. And similarly for faces in the bottom row. The phenomenon is called dichoptic color masking. For it to work, the contrasts must be low, the borders of figures slightly blurred and the luminances of the two colors about the same (isoluminant stimuli). The behavioral data are illustrated in the bar graph. Subjects had to choose between whether they saw a face, a house or a uniform square. Sf = same faces, of = opposite faces, sh = same houses, oh = opposite houses. When presented with the same faces in both eyes, subjects saw that face consciously. When subjects were presented with opposite faces, in 95% of the cases they reported a uniform field. The authors suggest that the exceptions were probably due to imperfect fusion, due to—for example—eye movements, resulting in an imperfect alignment of the two stimuli. They also did the same experiment with only two options, face and house, to see if subjects could detect—consciously or unconsciously—which stimuli they were getting. Subjects were 52.7% correct (50% would be chance performance), which the authors ascribe to imperfect fusion. Phillips claims that the high accuracy (often in the mid-90 percents) in the choices made by blindsight patients reflects weak conscious perception of the two options. However there is little support for such a devil's advocate position concerning the perception of the face or the house, especially given these very low numbers and the plausibility of occasional imperfect fusion.

In a variant of this technique, Fogelson, et.al. (2014) asked each subject to adjust the luminances of red and green stimuli until a fused color was achieved—dark greenish yellow given their stimuli. They used this individual adjustment to tailor the luminances of the stimuli to each subject. This technique makes it very plausible that the conscious perception was indeed of the fused color. In another variant, Schurger, et al. (2010) asked subjects to guess whether the invisible figure was a face or a house and then to wager for monetary rewards on whether their guesses were right. Subjects'

bets carried no information about what they were seeing, again suggesting that the subjects had no idea at all and so the stimuli were not consciously perceived. Contrary to Phillips, these items constitute scientific evidence of unconscious perception.

Faces and houses are known to activate specific brain regions, especially the "fusiform face area" in the case of faces and less reliably, the "parahippocampal place area" in the case of houses (Tong, Nakayama, Vaughan, & Kanwisher, 1998). Moutoussis and Zeki found that "much the same" brain areas were activated in both conscious and unconscious perception, albeit to a lesser degree in unconscious perception. But again we must ask whether these activations are indicative of perception—as opposed to the sub-personal processes underling perception. Low level visual representations, e.g. edges and textures, are the products of sensory transduction and are causally involved in the production of other (i.e. high level) visual representations. Low level properties are registered in early vision and do not normally figure in personal level cognition. However, high level activations—such as face area activations—are more likely to make contact with personal level understanding, just as the gender activations in the (Jiang et al., 2006) study made contact with personal level values.

The upshot is that in the case of complementary faces presented to the two eyes, we have a perceptual state that in its conscious aspect represents a uniform yellow square and in its unconscious aspect represents two complementary faces, i.e. a red face on a green background and in addition, a green face on a red background. I conclude that even if Phillips is right about the two brain damage syndromes, visuo-spatial neglect and blindsight, the significance of his point is limited because the points are unlikely to apply to unconscious perception in people who do not have brain damage and because no case has been made for similar flaws in most of these paradigms.

I will close with some additional remarks about the Anna Karenina principle.

If Tolstoy is right that every unhappy family is unhappy in a different way, a feature that is especially important in making one family happy might be present but undermined in its effectiveness in some unhappy families. So we should not expect a single magical ingredient in happy families that is never present in unhappy families. The happiness analogy goes only so far in application to consciousness since the plausibility of a functional or even behavioral analysis is much greater for happiness than for consciousness. Even if there is a physico-chemical essence to consciousness that is absent in every case of unconsciousness, it is very unlikely that there is any such essence to happiness. But the happiness analogy may apply to abilities that are allegedly "signatures" of consciousness.

An ongoing search for abilities that are only present in conscious states has repeatedly found some version in unconscious states. For example, it was once thought that the ability to inhibit a response required consciousness. However, Simon van Gaal and colleagues showed in a series of experiments that an unconsciously perceived "no go" signal made a response to a subsequent target less likely and also increased response times when the response was not inhibited (van Gaal, Ridderinkhof, Scholte, & Lamme, 2010).

Discussion of unconscious perception by both philosophers and scientists would be improved by an appreciation of the Anna Karenina principle.

Bibliography

Block, N. 2011. The Anna Karenina theory of the unconscious. *Neuropsychoanalysis* 13(1), 34–37.

——, Carmel, D., Fleming, S., Kentridge, R., Koch, C., Lamme, V. Rosenthal, D. 2014. Consciousness science: real progress and lingering misconceptions. *Trends in Cognitve Sciences* 18(11), 556–567.

——, & Phillips, I. (forthcoming). Unconscious seeing—a debate. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception*. New York: Routledge.

Breitmeyer, B. (2015). Psychophysical ''blinding'' methods reveal a functional hierarchy of unconscious visual processing *Consciousness and Cognition*. doi: http://dx.doi.org/10.1016/j.concog.2015.01.012

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

Fahrenfort, J., Snijders, T., Heinen, K., van Gaal, S., Scholte, H., & Lamme, V. 2012. Neuronal integration in visual cortex elevates face category tuning to conscious face perception. *Proceedings of the National Academy of Sciences* 109(52), 21504–21509.

Fogelson, S., Kohler, P., Miller, K., Granger, R., & Tse, P. 2014. Unconscious neural processing differs with method used to render stimuli invisible. *Frontiers in Psychology* 5(601), 1–11.

Jiang, Y. V., Costello, P., Fang, F., Huang, M., & He, S. 2006. A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences* 103(45), 17048–17052.

Klauer, K., Eder, A., Greenwald, A., & Abrams, R. 2007. Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition* 16, 63–83.

Moutoussis, K., & Zeki, S. 2002. The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences* 99(14): 9527–9532.

Mudrik, L., Gelbard-Sagiv, H., Faivre, N., & Koch, C. (2013). *Knowing where without knowing what: partial awareness and high level*

processing in continuous flash suppression. Paper presented at the Vision Science Society conference, Naples, Florida.

——, & Koch, C. 2013. Differential processing of invisible congruent and incongruent scenes: A case for unconscious integration. *Journal of Vision* 13(13): 1–14.

Phillips, I. (2016). Consciousness and criterion: on Block's case for unconscious seeing. *Philosophy and Phenomenological Research*.

—— (forthcoming). Unconscious perception reconsidered.

Raio, C., Carmel, D., Carrasco, M., & Phelps, E. A. 2012. Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology* 22(12), 477–479.

Schurger, A., Pereira, F., Treisman, A., & Cohen, J. 2010. Reproducibility distinguishes conscious from nonconscious neural representations. *Science* 327(97), 97–99.

——, Sarigiannidis, I., Naccache, L., Sitt, J., & Dehaene, S. 2015. Cortical activity is more stable when sensory stimuli are consciously perceived. *Proceedings of the National Academy of Sciences* 112(16), E2083–E2092.

Tolstoy, L. 1901. *Anna Karenina*. New York: Random House.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. 1998. Binocular Rivalry and Visual Awareness in Human Extrastriate Cortex. *Neuron* 21 (4), 753–759.

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. 2010. Unconscious Activation of the Prefrontal No-Go Network. *The Journal of Neuroscience* 30(11), 4143–4150.