Psychological Review

© 2021 American Psychological Association
ISSN: 0033-295X

2021, Vol. 128, No. 3, 592–595
https://doi.org/10.1037/rev0000277

REPLY

Bias and Blindsight: A Reply to Michel and Lau (2021)

Ian Phillips
William H. Miller III Department of Philosophy and Department of Psychological and Brain Sciences,
Johns Hopkins University

According to the textbooks, blindsight is a neuropsychological condition characterized by preserved capacities for voluntary visual discrimination unaccompanied by visual awareness. So construed, blindsight precipitated a revolution in theorizing consciousness. In Phillips (2021), I argued that the textbooks are wrong and the revolution ill-founded. Blindsight is exclusively a matter of conscious, albeit qualitatively degraded, vision which appears unconscious because of conservative response bias. Michel and Lau (2021) object: first, that residual awareness in blindsight cannot account for patients' impressive, feature-specific discriminatory abilities; and second, that performance matching makes response-bias explanations of unreported awareness implausible. They then offer a positive picture of blindsight as a specific deficit of detection, locating this idea within a framework which distinguishes perceptual from response bias. Here, I explain why neither objection convinces. I then argue that Michel and Lau give us no good reason to prefer their approach to our simpler, conscious-vision-only alternative.

Keywords: blindsight, awareness, signal detection theory, response bias, perceptual bias

According to the textbooks, blindsight is a neuropsychological condition characterized by preserved capacities for voluntary visual discrimination unaccompanied by visual awareness. So construed, blindsight precipitated a revolution in theorizing consciousness. In Phillips (2021), I argued that the textbooks are wrong and the revolution ill-founded. Blindsight is exclusively a matter of conscious, albeit qualitatively degraded, vision which appears unconscious because of conservative response bias. Michel & Lau (2021) object: first, that residual awareness in blindsight cannot account for patients' impressive, feature-specific discriminatory abilities; and second, that performance matching makes response-bias explanations of unreported awareness implausible. They then offer a positive picture of blindsight as a specific deficit of detection, locating this idea within a framework which distinguishes perceptual from response bias. Here, I explain why neither objection convinces. I then argue that Michel and Lau give us no good reason to prefer their approach to our simpler, conscious-vision-only alternative.

Feature-Specific Discrimination in Blindsight

Michel and Lau rightly note that the critical issue is not whether blindsight patients have "blind" field experiences, but whether residual performance is exclusively based on such experiences. In defending this claim, I emphasized that awareness in blindsight is qualitatively degraded, perhaps sometimes simply a matter of "feature-agnostic" salience. Michel and Lau counter that such awareness "cannot account for [the] ability to discriminate between specific visual features." My contention was never that all awareness in blindsight is feature-agnostic. Plausibly, some patients retain feature-specific awareness, perhaps of first-order motion, orientation, and luminance. Nonetheless, a pure salience hypothesis is consistent with highly accurate, feature-specific discrimination. All that is required is an exploitable correlation between salience and feature dimension, such as that which DB and GY both report between color and "strength of feeling" (Alexander & Cowey, 2010, p. 524, Weiskrantz et al., 1974, p. 721).

Michel and Lau also dispute the correlation between performance and awareness on the basis that Mazzi et al.'s (2016) patient, SL, reported awareness in the absence of performance. To them, "this suggests that blindsight patients' occasional reports of awareness are often unrelated to the specific features that they perceive." However, Mazzi et al.'s results show only that SL is sometimes dimly aware of stimuli while lacking awareness of specific features and thus at chance in discriminating them. This is unsurprising and does not undermine the significance of their finding that whenever there is performance, awareness is reported when a suitably sensitive scale is used. True, we cannot assume that this awareness is feature-specific. Critically, however, we do not find cases of performance unaccompanied by awareness. This is discriminating, positive evidence that performance has an exclusively experiential basis.

Performance Matching

Following Weiskrantz, Michel and Lau contend that the "discrepancy between awareness and performance [in blindsight] is

Ian Phillips https://orcid.org/0000-0003-2932-8045
Correspondence concerning this article should be addressed to Ian Phillips, William H. Miller III Department of Philosophy, Johns Hopkins University, Gilman Hall, 3400 North Charles Street, Baltimore, MD 21218, United States. Email: ianbphillips@jhu.edu

most evident in cases of performance matching." They focus on Persaud et al.'s (2011) finding that, even with discriminative sensitivity equated, GY reported awareness much more often in his intact field. I argued that this difference reflects differences in criterion setting (liberal in his intact field, conservative in his "blind" field) due to differences in criterion contents, perhaps compounded by field-specific expectations.

Michel and Lau object that GY was equally willing to wager on two-alternative forced-choice (2afc) discriminations in his "blind" field. However, as previously emphasized, GY's wagering is not a secure basis for inference. First, since the alternative to betting was a 50/50-coin flip, the task has a (weakly) dominant strategy: always bet on one's 2afc response (Clifford et al., 2008; Konstantinidis & Shanks, 2014). Consequently, and *pace* Michel and Lau, GY's betting does not imply knowledge or confidence in his overall performance. Second, even if his "blind" field wagers do reflect a lack of trial-by-trial metacognitive sensitivity, this could be for numerous reasons aside from lack of first-order awareness (Maniscalco & Lau, 2012; Seth, 2008).

Michel and Lau further note that changes in reported awareness occur *within* the "blind" field despite performance matching. This too can be explained in terms of variations in criterion contents across the "blind" field (or with stimulus parameters), again perhaps compounded by more specific expectations. Michel and Lau argue that neurotypical observers are slow to shift criteria. But great care is needed in generalizing to blindsight. There are large and nonuniform qualitative differences (Weiskrantz, 2009); we cannot assume homogeneous criterion contents; and criterion setting is notoriously unusual (Azzopardi & Cowey, 2001).

Michel and Lau's objections thus do not confound the hypothesis that blindsight is exclusively a matter of qualitatively degraded conscious vision. Do they offer a preferable positive account?

Michel and Lau's Account

Michel and Lau propose (a) that blindsight involves an impairment of *detection* despite preserved *discrimination*, and (b) that understanding blindsight requires distinguishing perceptual and response criteria.

Regarding (a), Michel and Lau focus on Azzopardi and Cowey's (1997, 1998) finding that the traditional theoretical relation between discrimination and detection (i.e., $d'_{2afc} = \sqrt{2}d'_{yn}$) apparently breaks down with static stimuli in blindsight. As discussed, Azzopardi and Cowey (2001) show that this apparent violation can be understood as an artifact of criterion instability ("jitter") in yes/no (yn) detection. Michel and Lau concur. Nonetheless, they claim that the reason for jitter with static but not moving stimuli is that static stimuli are not consciously experienced. However, we can equally explain the contrast via the well-established fact that motion sensitivity is relatively spared in blindsight (Azzopardi & Hock, 2011). Given this, it may be of greater functional use in everyday life (cf., Riddoch, 1917) and generate a larger pool of projectable traces for criterion stabilization. Motion experience may also be uniquely comparable across fields (Stoerig & Barth, 2001), allowing the use of traces from the intact field. Since Michel and Lau accept that "this is a plausible hypothesis," there would seem no need to hypothesize that static stimuli are perceived unconsciously.

What of (b), the claim that understanding blindsight requires distinguishing perceptual and response criteria? The idea here is that conscious detection of a stimulus requires its reaching a subject's *perceptual* criterion which may or may not coincide with their *response* criterion. Blindsight is genuinely blind because stimuli typically fail to reach the patient's conservative perceptual criterion. Two key questions arise here. Should we generally distinguish between perceptual and response criteria? If so, does the distinction illuminate blindsight?

Should We Distinguish Perceptual and Response Criteria?

Michel and Lau provide a very helpful review of numerous studies suggesting that we should distinguish perceptual and response criteria. Each deserves careful consideration. Here I restrict myself to some brief remarks.

First, many such studies can instead be understood as distinguishing different kinds of response bias. For example, Gallagher et al. (2019) propose that confidence judgments distinguish perceptual and response biases, intuiting that perceptual biases will affect both perceptual decisions and subjective confidence judgments, whereas response biases will exclusively affect decisions. They test this using a motion direction discrimination task, showing that adaptation to one direction affects both perceptual decisions and confidence judgments, whereas an explicit instruction to favor one direction only affects decisions. However, we can equally explain this pattern by distinguishing between spontaneous or robust response biases induced by adaptation, and deliberate or fragile response biases induced by instruction (cf., Gallagher et al., 2019, p. 9). Similarly, Fritsche et al.'s (2017) results are consistent with a distinction between short-lived, local response biases and longer-lived, nonlocal biases.

Second, many of the reviewed studies show only that a pattern of performance cannot be explained by some specific response bias. This falls short of showing that performance is not due to any such bias. For instance, Iemi and Busch (2018) compare the effects of prestimulus excitability on *2afc* detection and discrimination. They take their results to show that high-prestimulus excitability induces a specifically perceptual bias. Yet, strictly, what they show is that the effect of excitability cannot be understood in terms of a simple *interval* bias because it affects *2afc* detection but not *2afc* discrimination. This is quite consistent with excitability inducing a response bias toward judging that a given interval contains a stimulus which would only affect detection.

More broadly, the reviewed data require substantive and miscellaneous assumptions to evidence perceptual bias. For example, Crapse et al. (2018) induced criterion shifts by manipulating activity in monkey superior colliculus (SC). To infer a perceptual criterion shift requires assuming: (a) that SC activity is an implausible basis for response bias, and (b) that SC was the basis of the criterion shift. Notably, Crapse et al. themselves aver "that other brain areas likely cooperate with the SC to provide a signal of decision criteria, most notably . . . prefrontal cortex" (p. 191).

Michel and Lau also theoretically motivate distinguishing perceptual and response bias by arguing that spontaneous neural activity in sensory neurons is not inevitably accompanied by faint hallucinations. However, this does not require postulating a perceptual criterion. Instead, as Michel and Lau concede, we can deny that spontaneous neural activity inevitably registers as subjective likelihood. What really matters, they insist, is that in "malfunctioning conditions" such as blindsight, "high levels of otherwise functionally significant sensory activities may . . . lead to zero subjective stimulus likelihood." By "functionally significant," here, Michel and Lau must mean: available for use in discrimination—otherwise performance in blindsight will be left unaccounted for. Yet, it is surely puzzling that a subject might successfully discriminate between stimulus intervals despite registering both as having zero likelihood of containing anything.

Perceptual Criteria in Blindsight

Suppose we are, nonetheless, convinced that a distinction should be drawn between perceptual and response criteria. Is such a distinction required to explain the patterns of performance and (un)reported awareness characteristic of blindsight?

Consider, first, reported awareness. Michel and Lau suggest that subjects typically anchor their response criterion to their perceptual criterion. Assuming that blindsighted subjects have a stubbornly conservative perceptual criterion, this explains conservative responding and failures to report awareness. However, we can equally well explain why patients adopt a conservative response criterion by appeal to the qualitatively degraded nature of their residual vision (in combination with various motivational and individual factors). Nor does invoking a perceptual criterion illuminate the complex patterns of "blind" field criterion setting discussed above. More positively, and as discussed in detail, our response criterion-based account makes distinctive predictions borne out by the data. In particular, it predicts reported awareness under variations in instruction, response options, and motivation; and correlations between performance and awareness when sensitive (e.g., nonbinary) measures of awareness are used.[1]

What about criterion instability? Here, Michel and Lau suggest that because the perceptual criterion is extreme in blindsight, when patients are "required to adopt a less conservative response criterion in experiments, they cannot . . . anchor the response criterion to the perceptual criterion," creating instability. This is an intriguing proposal.[2] But, again, Michel and Lau give no reason to think that it represents an explanatory improvement over our own conscious-vision-only approach on which instability is due to the nonfunctional nature of residual vision in conjunction with general mechanisms of criterion stabilization. Such reasons are surely needed before we interpret blindsight as a revolutionary condition which overturns our traditional understanding of the interrelations between perception, action, and awareness.

Stepping Back

Michel and Lau remark that blindsight "opened up the possibility that awareness and visual sensitivity may be dissociated" and "inspired decades of studies using many other methods, confirming one and another in different ways." We must be "careful not to throw the baby out with the bathwater" (cf., Finkbeiner & Coltheart, 2014; Peters et al., 2016).

Certainly, we should not rush to judgment. Individual cases may not generalize (though the relevance of the small lesion in GY's right parietal cortex, which Michel and Lau flag, is obscure). There is also a vast amount to learn from blindsight independent of issues concerning awareness. However, we cannot simply assume the existence of a baby because the water is murky. Attempts to demonstrate unconscious perception have a long history of apparent success evaporating upon careful examination (e.g., Balsdon & Clifford, 2018; Eriksen, 1960; Holender, 1986; Newell & Shanks, 2014). Accordingly, Michel and Lau's contribution to the careful examination of blindsight is most welcome. Nonetheless, I remain unconvinced that they have identified a genuine baby.

References

Alexander, I., & Cowey, A. (2010). Edges, colour and awareness in blindsight. *Consciousness and Cognition*, *19*, 520–533. https://doi.org/10.1016/j.concog.2010.01.008

Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, near-threshold vision? *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 14190–14194. https://doi.org/10.1073/pnas.94.25.14190

Azzopardi, P., & Cowey, A. (1998). Blindsight and visual awareness. *Consciousness and Cognition*, *7*, 292–311. https://doi.org/10.1006/ccog.1998.0358

Azzopardi, P., & Cowey, A. (2001). Why is blindsight blind? In B. de Gelder, E. H. F. De Haan, & C. A. Heywood (Eds.), *Out of mind: Varieties of unconscious processes* (pp. 3–19). Oxford University Press.

Azzopardi, P., & Hock, H. S. (2011). Illusory motion perception in blindsight. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 876–881. https://doi.org/10.1073/pnas.1005974108

Balsdon, T., & Clifford, C. W. G. (2018). Visual processing: Conscious until proven otherwise. *Royal Society Open Science*, *5*, 1–16. https://doi.org/10.1098/rsos.171783

Clifford, C. W. G., Arabzadeh, E., & Harris, J. A. (2008). Getting technical about awareness. *Trends in Cognitive Sciences*, *12*, 54–58. https://doi.org/10.1016/j.tics.2007.11.009

Crapse, T. B., Lau, H., & Basso, M. A. (2018). A role for the superior colliculus in decision criteria. *Neuron*, *97*, 181–194. https://doi.org/10.1016/j.neuron.2017.12.006

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, *67*, 279–300. https://doi.org/10.1037/h0041622

Finkbeiner, M., & Coltheart, M. (2014). Dismissing subliminal perception because of its famous problems is classic "baby with the bathwater." *Behavioral and Brain Sciences*, *37*, Article e27. https://doi.org/10.1017/S0140525X13000708

Foley, R. (2015). The case for characterising Type-2 blindsight as a genuinely visual phenomenon. *Consciousness and Cognition*, *32*, 56–67. https://doi.org/10.1016/j.concog.2014.09.005

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, *27*, 590–595. https://doi.org/10.1016/j.cub.2017.01.006

Gallagher, R., Suddendorf, T., & Arnold, D. (2019). Confidence as a diagnostic tool for perceptual aftereffects. *Scientific Reports*, *9*, 1–12. https://doi.org/10.1038/s41598-019-43170-1

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1–23. https://doi.org/10.1017/S0140525X00021269

Iemi, L., & Busch, N. A. (2018). Moment-to-moment fluctuations in neuronal excitability bias subjective perception rather than strategic

[1] Regarding the former, Michel and Lau embrace the idea that reported awareness might reflect the sensorimotor results of sensory processing. This is hard to square with the distinctively visual language found in many such reports (Foley, 2015).

[2] It is a nice question how it applies to Azzopardi and Cowey's studies where GY was not required to adopt a less conservative criterion, and indeed remained "very biased" in Experiment 1 (Azzopardi & Cowey, 1998, p. 297).

decision-making. *eNeuro*, *5*, 1–13. https://doi.org/10.1523/ENEURO.0430-17.2018

Konstantinidis, E., & Shanks, D. R. (2014). Don't bet on it! Wagering as a measure of awareness in decision making under uncertainty. *Journal of Experimental Psychology: General*, *143*, 2111–2134. https://doi.org/10.1037/a0037977

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*, 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Mazzi, C., Bagattini, C., & Savazzi, S. (2016). Blind-Sight vs. degraded sight: Different measures tell a different story. *Frontiers in Psychology*, *7*, Article e901. https://doi.org/10.3389/fpsyg.2016.00901

Michel, M., & Lau, H. (2021). Is blindsight possible under signal detection theory? Comment on Phillips (2021). *Psychological Review*, *128*(3), 585–591. https://doi.org/10.1037/rev0000266

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *37*, 1–19. https://doi.org/10.1017/S0140525X12003214

Persaud, N., Davidson, M., Maniscalco, B., Mobbs, D., Passingham, R. E., Cowey, A., & Lau, H. (2011). Awareness-related activity in prefrontal and parietal cortices in blindsight reflects more than superior visual performance. *NeuroImage*, *58*, 605–611. https://doi.org/10.1016/j.neuroimage.2011.06.081

Peters, M. A. K., Ro, T., & Lau, H. (2016). Who's afraid of response bias? *Neuroscience of Consciousness*, *1*, 1–8. https://doi.org/10.1093/nc/niw001

Phillips, I. (2021). Blindsight is qualitatively degraded conscious vision. *Psychological Review*, *128*(3), 558–584. https://doi.org/10.1037/rev0000254

Riddoch, G. (1917). Dissociation of visual perceptions due to occipital injuries, with especial reference to appreciation of movement. *Brain: A Journal of Neurology*, *40*, 15–57. https://doi.org/10.1093/brain/40.1.15

Seth, A. K. (2008). Post-decision wagering measures metacognitive content, not sensory consciousness. *Consciousness and Cognition*, *17*, 981–983. https://doi.org/10.1016/j.concog.2007.05.008

Stoerig, P., & Barth, E. (2001). Low-Level phenomenal vision despite unilateral destruction of primary visual cortex. *Consciousness and Cognition*, *10*, 574–587. https://doi.org/10.1006/ccog.2001.0526

Weiskrantz, L. (2009). Is blindsight just degraded normal vision? *Experimental Brain Research*, *192*, 413–416. https://doi.org/10.1007/s00221-008-1388-7

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain: A Journal of Neurology*, *97*, 709–728. https://doi.org/10.1093/brain/97.1.709

Received October 5, 2020
Revision received November 23, 2020
Accepted November 23, 2020 ■