

Look Closer: A Point-by-Point Reply Regarding ‘Moral Perception’ (Or: Do We Really Think That Hitting Small Children Is Okay?)

Chaz Firestone & Brian J. Scholl

A primary focus in our laboratory is the study of how perception connects up with the rest of the mind, and in particular how visual processing itself traffics in factors that we typically associate with higher-level cognition — e.g. physical causality (e.g. Scholl & Nakayama, 2002), animacy and intentionality (e.g. Gao et al., 2010), and causal history (e.g. Chen & Scholl, in press). As a result, we were deeply interested in the notion of ‘moral perception’ — the novel suggestion that “perception is preferentially attuned to moral content” — as introduced in a recent *TICS* forum piece by Gantman and Van Bavel (2015). We were then delighted to engage in a brief dialogue about this notion, as realized in our recent commentary (Firestone & Scholl, 2016) and in their subsequent response (Gantman & Van Bavel, 2016). Reacting to this dialogue, some colleagues asked us what we thought of their response, and so the purpose of this document is just to offer a few brief informal rejoinders.

We thought nearly every one of G&VB’s points missed the mark, and that the arguments for ‘moral perception’ continue to be based on misinterpretations, overinterpretations, and flawed evidence. Here’s why:¹

Our 1st Argument: Much of the trumpeted work doesn’t reflect morality.

The search for evidence of ‘moral perception’ ended up focusing on phenomena that may not be about morality in the first place. In one case, the original study had already explicitly ruled out morality empirically (Sherman et al., 2012, Experiment 3), though G&VB did not mention this. In another case (Anderson et al., 2011), the effect worked with immoral actions (e.g. hitting a child) but also (a) worked with actions that were socially negative but not necessarily *immoral* (e.g. defecating in public, or being in a car accident while borrowing your friend’s car), and (b) *failed* to work with morally good actions (e.g. helping the elderly) — suggesting that morality *per se* may not be driving this effect.

G&VB

“We fail to see how hitting a small child does not pertain to morality”.

Response: G&VB are missing the point.

Of course hitting a small child is immoral. (Pro tip: If ever you think someone is arguing that hitting small children is okay, consider whether you might be misinterpreting them!) Our point was that *in addition to* those clearly immoral

¹ G&VB made some related arguments when commenting on our forthcoming *BBS* target article, “Cognition Does Not Influence Perception: Evaluating the Evidence for ‘Top-Down’ Effects”. As a result, a few of our responses here are cribbed from the reply to our *BBS* commentators. Both the target article and the response are available here:

<http://perception.research.yale.edu/preprints/Firestone-Scholl-BBS-Full.pdf>

cases, “such effects *also* occurred with mere norm violations that do not invoke morality (e.g. indecent public behavior)” (Firestone & Scholl, 2016, emphasis added); so, for this and other reasons, we don’t know whether morality is really responsible. (Indeed, Anderson et al. themselves think this implicates “negative social affect” rather than morality.) In short, G&VB focus too much on hitting small children, and not enough on public defecation.

Additional Response: What happened to Sherman et al.?

In the initial forum article, the very first piece of evidence that was trumpeted as indicating ‘moral perception’ was the work by Sherman et al. — which G&VB interpreted as reflecting “moral emotions”. We pointed out that these authors had explicitly ruled this out — discovering that their results were driven by physical disgust, but *not* moral disgust (something that the *TICS* forum piece failed to mention). G&VB did not respond to this in their reply; perhaps they agree on this point?

Our 2nd Argument: Much of the trumpeted work doesn’t reflect perception.

The discovery that people attend to just outcomes in moral scenarios (Callan et al., 2013) tells us something about what we expect in moral scenarios, but it doesn’t tell us anything new about how *attention itself* works. We already knew that attention is drawn to what people expect.

G&VB

“Isn’t that interesting?”

Response: Sure, it’s interesting, but it still doesn’t support ‘moral perception’.

These findings are indeed interesting for certain purposes (which is why our *TICS* comment explicitly described them as “interesting”!). But they’re interesting in revealing *what our moral expectations are*, not in revealing new ways that attention works.

G&VB

“The goal of cognitive science is building process-oriented models”.

Response: Exactly!

This is exactly what motivated our studies. The initial allegations of ‘moral popout’ simply reported the effect and gave it a shiny new name — without any consideration of just *how* such effects could instead be produced by the underlying mechanics of familiar cognitive processes (such as memory). Our entire point is that when you consider such “process-oriented models”, you realize that this exciting new phenomenon is really just semantic priming:

Our 3rd Argument: ‘Moral popout’ doesn’t exist; it’s just semantic priming.

Moral words (e.g. “crime”, “punishment”) are easier to see than random non-moral words (e.g., “steel”, “ownership”) simply because moral words are semantically related (whereas random non-moral words are not), and related words prime each other. As a result, words from *any* arbitrary category —

including 'silly' categories such as fashion — will “pop out” when compared to unrelated words (Firestone & Scholl, 2015). There is no evidence that morality plays a role.

G&VB

“They argued that moral pop-out can be fully explained by semantic priming because ‘relatedness is the key factor in such effects’”.

Response: No, we argued this on the basis of our experiments.

The relatedness confound alone doesn't mean that moral popout must be fully explained by semantic priming; it merely means that semantic priming *could* be responsible, such that Gantman and Van Bavel's (2014) original conclusions wouldn't follow. (Note that G&VB never mentioned the 'relatedness' confound in their initial studies.) Nevertheless, we actively tested this alternative, and semantic relatedness *in fact* produced analogous “popout” effects. Our experimental results — not just the confound itself — are what suggest that 'moral popout' is really just semantic priming.

G&VB

Moral and non-moral words were “matched for length and frequency”.

Response: No: they were confounded in a self-serving way.

The moral words used by Gantman and Van Bavel (2014) were 7% shorter and 25% more frequent than the non-moral words were — two self-serving biases that have never been pointed out. (We're not sure where the line should be drawn between “matched” and “confounded in an unacknowledged and self-serving way” — but surely it's at least an order of magnitude smaller than 25%.) Our experiments (Firestone & Scholl, 2015) carefully controlled for these factors and are thus the only interpretable published data on “moral popout”.

G&VB

“The words were presented near the threshold for awareness”.

Response: We can't be sure, since the reported duration was impossible.

Nobody knows how long the words were presented onscreen in G&VB's (2014) experiments, since they reported a critical presentation duration of 40ms on a 60Hz monitor. This is physically impossible: such a monitor can only present words for multiples of 1/60 of a second — e.g. 17ms, 33ms, 50ms, etc. (For this reason, our replication used a 50ms duration.)

G&VB

“They [F&S] did not randomly assign participants to detect moral versus fashion and/or transportation words.”

Response: Look closer: Just how is this challenge supposed to work?

It is true that subjects were not “randomly assigned” to which experiment they participated in: first we discovered a silly 'fashion popout' effect, then we

generalized the effect to another arbitrary category (transportation), and then (in response to comments from reviewers, including G&VB) we replicated 'moral popout' too. So what is the challenge here? G&VB don't elaborate, so let's look closer. If we were claiming *differences* between the experiments, then non-random assignment could be problematic in a straightforward way (e.g., perhaps one group was more tired or stressed out). But we suggested that there is *no* evidence of any relevant differences among these popout effects, and our explanation is that the same underlying process — semantic priming — drives all the effects (with no evidence that morality plays any role). Can a lack of random assignment explain this apparent equivalence differently? Such an explanation would have to assume that the 'true' effect with morality in our experiments was much larger than for the other categories (due to the morality-specific boost) but that these particular subjects (i.e. Yale community members tested in one month rather than another) somehow deflated this previously undiscovered 'super-popout' down to ... exactly the same magnitude (of 4%) previously reported by Gantman and Van Bavel (2014)! Random assignment is a red herring here; it cannot save the day for 'moral popout'.

G&VB

"Curiously, however, Firestone and Scholl did not report the analogous means for their morality study..."

Response: Look closer: The means support our account.

Let's explore this in detail. Our claim is that words from a given category (whether fashion, transportation, or morality) prime each other and enhance lexical decision performance. The critical evidence for this is simply that words from such categories show a "popout" effect: the mere fact that fashion and transportation pop out too is why we think there's no evidence that morality *per se* plays any role. But, in addition to this, we can also look for evidence of semantic priming at the trial-by-trial level, to see whether category words seen on one trial enhance performance for category words on the next trial (even though our experiments were not designed for this analysis). To test this — say, for fashion — we first need to ask whether fashion words preceded by fashion words (e.g. "blouse" preceded by "dress") were recognized more accurately than fashion words preceded by non-fashion words (e.g. "blouse" preceded by "limit"). In fact, they were, by 4.9%. However, this number is meaningless on its own: maybe fashion words are exciting or arousing (after all, they included the word "bikini"!) and would increase performance for *any* subsequent word (in which case the true priming effect could be nil); or, maybe fashion words are distracting and would, all else equal, *impair* performance on subsequent words (in which case the true repetition priming effect could be even larger than 4.9%). So, for this to be interpretable at all, we also need to compare performance on *non-fashion* words preceded by fashion words (e.g., "limit" preceded by "blouse") to *non-fashion* words preceded by non-fashion words (e.g., "limit" preceded by "diesel"). And indeed, fashion words in this case confer a 3.9% *disadvantage* for non-fashion-word recognition accuracy. So, fashion words (rather than non-fashion words) boost recognition of fashion words (rather than non-fashion

words) by $[(4.9\%) - (-3.9\%)] = 8.8\%$. For transportation, that value is $[(5.7\%) - (-0.2\%)] = 5.9\%$. And for morality, that value is $[(1.1\%) - (-3.4\%)] = 4.5\%$. And we also reported a valenced-matched version of this morality analysis, which produces $[(2.5\%) - (-2.7\%)] = 5.2\%$. These values did not differ significantly — though, as might not be surprising, these three values of 8.8% (fashion), 5.9% (transportation), and 5.2% (morality) are rank-ordered just as are the magnitudes of the 'popout' effects themselves: 8.6% (fashion), 4.3% (transportation), and 3.9% (morality). These are the means G&VB asked for, and they plainly support our account.² In sum, looking closely and carefully at the data supports the claim that there is no qualitative difference between 'moral popout' and 'fashion popout', and that both can be entirely explained by semantic priming.

G&VB

"...despite the fact that it was central to their hypothesis."

Response: This analysis was never "central" (though it does support us).

Our hypothesis is that semantic priming explains "popout" effects. To demonstrate this, we showed that "popout" effects emerge for arbitrary categories in the exact same way as in Gantman and Van Bavel (2014). Whether these effects also manifest at the trial-by-level is a completely separate matter that is not at all required by our hypothesis. Nevertheless, discovering such trial-level effects is *further evidence* for our hypothesis (as reported in detail above), even though our study was not designed to test for such effects. We have reported all of these means, which continue to suggest that "moral popout" is really just semantic priming.

G&VB

"[T]he burden of proof is on Firestone and Scholl to design and fully report a study that demonstrates how fashion and/or transportation pop-out is 'entirely analogous' to moral pop-out."

Response: This seems backward, and counter to general scientific practices.

G&VB are claiming that there is a revolutionary new phenomenon of "moral popout". We noted a stark and unrecognized confound in those studies — implicating an alternative factor (semantic priming) that has been well understood for decades — and then showed empirically that this confound alone produces similar effects. Surely this means that the burden of proof is now on G&VB to show that there is more to "moral popout" than semantic priming!

² You may notice that, relative to fashion and transportation, moral words preceded by moral words weren't recognized much more accurately than moral words preceded by non-moral words (either 1.1% or 2.5% [depending on whether valence is taken into account] compared to 4.9% for fashion, and 5.7% for transportation [though these differences are not statistically significant]). But as with fashion, *this value is entirely uninterpretable on its own*: perhaps reading morally relevant words like "murder" and "violence" is startling and impairs performance on the next trial in a way that counteracts the moral-moral repetition priming advantage. That's why we also need to look at the effect of moral words on subsequently presented *non-moral* words — and indeed we see that non-moral words are recognized less accurately when preceded by moral words than by non-moral words (-3.4% or -2.7%).

(This is the same reason why the burden of proof is generally held to be on proponents of alternative medicines to produce compelling evidence for their claims — and not on their critics to design and fully report studies that refute them.)

G&VB

“Perhaps the only thing we can agree on is that more research is needed.”

Response: Amen!

This was the entire motivation for our studies. We think more — and more careful — research is needed before claiming that there is ‘moral popout’; before telling millions of lay readers that “moral concerns can alter a person’s perceptual experience” (<http://nyti.ms/1efPqcZ>); and before claiming that “moral perception” is a trend in cognitive science that has “important consequences for policy”. To proclaim such lofty conclusions in the absence of this evidence seems premature, and even a form of uncaring overreach. In any case, we continue to find ‘moral perception’ to be a fascinating (if unlikely) possibility, and we hope that further research on the underlying mechanics of these effects will help the empirical basis for this idea to catch up to the bold pronouncements that are made in its name.

References

- Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The visual impact of gossip. *Science*, 332, 1446–1448.
- Callan, M. J., Ferguson, H. J., & Bindemann, M. (2013). Eye movements to audiovisual scenes reveal expectations of a just world. *Journal of Experimental Psychology: General*, 142, 34–40.
- Chen, Y. -C., & Scholl, B. J. (in press). The perception of history: Seeing causal history in static shapes induces illusory motion perception. *Psychological Science*.
- Firestone, C., & Scholl, B. J. (in press). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences*.
- Firestone, C., & Scholl, B. J. (2015). Enhanced visual awareness for morality and pajamas? Perception vs. memory in “top-down” effects. *Cognition*, 136, 409–416.
- Firestone, C., & Scholl, B. J. (2016). ‘Moral perception’ reflects neither morality nor perception. *Trends in Cognitive Sciences*, 20, 75–76.
- Gantman, A. P., & Van Bavel, J. J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition*, 132, 22–29.
- Gantman, A. P., & Van Bavel, J. J. (2015). Moral perception. *Trends in Cognitive Sciences*, 19, 631–633.
- Gantman, A. P., & Van Bavel, J. J. (2016). See for yourself: Perception is attuned to morality. *Trends in Cognitive Sciences*, 20, 76–77.
- Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, 21, 1845–1853.
- Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science*, 13, 493–498.
- Sherman, G. D., Haidt, J., & Clore, G. L. (2012). The faintest speck of dirt: Disgust enhances the detection of impurity. *Psychological Science*, 23, 1506–1514.