Exposure to Justice Diminishes Moral Perception

Ana P. Gantman
Princeton University

Jay J. Van Bavel
New York University

Evidence suggests that people have a lower threshold for the conscious awareness of moral words. Given the potential motivational relevance of moral concerns, the authors hypothesized and found that motivational relevance of moral stimuli enhanced the detection of moral words. People who saw a CrimeStoppers advertisement in which a majority (vs. minority) of wanted murderers had been brought to justice exhibited reduced detection of moral words (Experiment 1). Similarly, people who read that an assailant was arrested (vs. escaped punishment) exhibited reduced detection of moral words (Experiment 2). In both experiments, the effect of justice motives on moral word detection was specific to words presented near (vs. distant) to the threshold for perceptual awareness. These findings suggest that satiating (vs. activating) justice motives can reduce the frequency with which moral (vs. non-moral) words reach perceptual awareness. Implications for models of moral psychology, particularly the role of perception in morality, are discussed.

Keywords: morality, motivation, justice, perception, detection

Lexical Properties That Affect Word Recognition

Any contextual effects on word detection are in addition to aspects of the lexical content that determine how easily a word is recognized. These include how frequently the word is used in the lexicon, how long it is, and whether the word is valenced, elicits arousal, or contains emotional content (Adelman & Estes, 2013). These factors of fluency can, in turn, determine valuation, as well as judgments of veracity and liking (for a review, see Alter & Oppenheimer, 2009). Much research on visual word recognition—one aspect of visual perception—pertains to how quickly a word is recognized in a lexical-decision task or vocalized in a pronunciation task. Many factors influence visual word recognition, including word length and frequency (i.e., how often a word is used; Brysbaert & New, 2009; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Word frequency also explains variation in word recognition accuracy, such that frequent words are recognized more quickly and accurately than less frequent words (Balota et al., 2004; Brysbaert & New, 2009; Yap & Balota, 2009). Words are also recognized more quickly when they have fewer syllables (Ferrand & New, 2003) and fewer orthographic neighbors (e.g., when one letter-change creates a new acceptable word; Andrews, 1997). Despite the inclusion of these lexical factors, a large amount of the variance in word recognition still remains unexplained (Adelman, Marquis, Sabatos-DeVito, & Estes, 2013).

It has been argued that social and emotional factors, such as valence (whether a word is positive or negative) and arousal (whether a word leads to excitation or relaxation) play an important role in word recognition (Kuperman, Estes, Brysbaert, & Warriner, 2014). There is reason to believe that people are able to differentially detect significant versus mundane stimuli because the visual system is closely integrated with other parts of the brain (Lim, Padmala, & Pessoa, 2009; Pessoa, 2015; Gilbert & Li, 2013). Indeed, emotional words appear to have a processing advantage because they are motivationally significant and recruit...
attention (Anderson, 2003; Anderson & Phelps, 2001; Egner & Hirsch, 2005; Summerfield & Egner, 2009), potentially reaching visual awareness earlier than their neutral counterparts. Specifically, emotions are motivationally relevant because they are organized around the appetitive (toward survival-promoting positive stimuli) and defensive (away from threatening or negative stimuli) systems (Kousa, Vinson, & Vigliocco, 2009; Lang, Bradley, & Cuthbert, 1990). Despite initial evidence that emotion words recruit attention that slows responding (Wentura, Rothermund, & Bak, 2000), later work found that when correcting for emotional words’ lower frequency, increased letters, and fewer orthographic neighbors, emotion words are responded to more quickly than neutral words (Kousa et al., 2009). Further, the facilitation of emotion word recognition appears relatively early in processing, reflecting preconscious processing (Kousa et al., 2009; Gaillard et al., 2006; Zeelenberg, Wagenmakers, & Rotteveel, 2006). People appear sensitive to both positive and negative words (vs. neutral) as early as 100 ms after word onset (Kissler, Herbert, Wingler, & Junghofer, 2009; Scott, O’Donnell, Leuthold, & Sereno, 2009; but see Nasrallah, Carmel, & Lavie, 2009). Most importantly, emotion words appear to have a lower threshold for visual awareness than neutral words (Gaillard et al., 2006).

Moral Word Detection

A growing body of research suggests that moral relevance can alter visual perception by lowering the threshold for the consciousness awareness of words (see Gantman & Van Bavel, 2015). Morality may be chronically motivationally salient as moral concerns fulfills multiple core motives, such as need to belong and maintain social groups (Haidt & Graham, 2009), need for justice (Lerner & Miller, 1978), and need for control (Kay, Gaucher, McGregor, & Nash, 2010). In short, the ability to recognize moral situations and act appropriately is critical to one’s survival in social groups and may be essential for securing access to needed physical and psychological resources afforded by group members (Gantman & Van Bavel, 2015). As such, we hypothesized that the motivational relevance of moral stimuli could enhance detection of moral words.

To test whether moral words were detected with greater frequency than non-moral words (i.e., “pop out”), we asked people to identify whether letter strings comprised words or nonwords when presented for a few dozen milliseconds. As predicted, people detected moral words (e.g., kill, should, just) more frequently than non-moral words (e.g., die, could, even). Not only were the moral and non-moral words matched for length, frequency in the language, and semantic content, a meta-analysis of three experiments revealed that the moral pop-out effect was not due to differences in valence, intensity, extremity, or reported arousal (Gantman & Van Bavel, 2014). In other words, moral words were correctly categorized as words more frequently than matched non-moral words when presented ambiguously—a phenomenon termed the moral pop-out effect.

The influence of top-down, recurrent feedback from higher-order systems is especially useful for object detection under conditions of ambiguity (Wyatte, Jilk, & O’Reilly, 2014). When visual input is varied in terms of strength or how much information is present to the visual system, (here, in terms of how long it is presented on screen) it is possible to determine what input strength (i.e., stimulus duration) is necessary for people to become conscious of a stimulus and generate a correct behavioral response (Kouider & Dehaene, 2007). Indeed, the moral pop-out effect was only present when letter strings were presented ambiguously—near the threshold for visual awareness (approximately 40–60 ms; Gantman & Van Bavel, 2014). When words were presented too quickly, people could not see them. When words were presented too slowly, people could see almost all of them. In other words, the moral pop-out effect suggests that moral content required less perceptual input to elicit a correct response. Similarly, work on the detection of emotion words suggests that they require fewer processing prerequisites (Anderson, 2003; Gaillard et al., 2006; Kousa et al., 2009). As such, we speculated that moral words that were perceptually ambiguous would be detected more frequently due to their motivational relevance (Gantman & Van Bavel, 2014).

Motives Tune Perception

Goals appear to exert a top-down influence on perception, making stimuli “pop out” in the environment when they are motivationally relevant in a domain-general fashion. When a goal is activated it heightens the accessibility of goal-related constructs (Kruglanski et al., 2002). Conversely, when a goal is attained, postattainment decrements in accessibility can be observed (Eitam & Higgins, 2010; Förster, Liberman, & Friedman, 2007). When the accessibility of goal-related constructs is measured before and after goal attainment, accessibility is heightened before and then inhibited after the goal is met (Marsh, Hicks, & Bink, 1998). In the current article, we tested the influence of justice motives on the accessibility of moral words to the visual system (i.e., whether or not moral words selectively reach conscious awareness).

Extensive research has identified that people are sensitive to justice concerns, and violations of justice lead to a motivation to restore justice. Most people are sensitive to justice concerns (Schmitt, Baumert, Gollwitzer, & Maes, 2010), refer to justice as one of their primary moral concerns (Graham, Haidt, & Nosek, 2009), and want to believe in a just world (Lerner & Miller, 1978). The need for justice has been characterized as an epistemic motive that possesses the same hallmarks of goal pursuit (Lerner & Miller, 1978), such that when the belief is threatened, just world-confirming information becomes more salient in the environment (Hafer, 2000; Kay & Jost, 2003). Just world needs also affect attention directly, as people who learned about a morally good or bad actor directed their eyes toward morally good or bad outcomes that would befall that same actor (Callan, Ferguson, & Bindemann, 2013). Similarly, following exposure to an unjust situation, people high in justice-sensitivity paid greater attention to justice-related words, interpreted an ambiguous situation as a justice violation, and show better recall for unjust information (Baumert, Gollwitzer, Staubach, & Schmitt, 2011). In sum, justice needs appear to increase the accessibility of justice-related words and tune attention toward justice-related stimuli.

Present Research

The current research aims to test whether moral motives can shape what we consciously see, and in so doing, to offer one
explanation for the moral pop-out effect (Gantman & Van Bavel, 2014). We conducted two experiments to investigate whether activating versus satiating justice motives would alter the detection of perceptually ambiguous moral words. To test this question, we used a modified lexical-decision task designed to vary the visibility of letter strings and measure the likelihood of a correct behavioral response as a function of three components: (a) perceptual ambiguity (i.e., the amount of information available to the visual system), (b) moral (vs. non-moral) content, and (c) motivational relevance (justice needs active vs. satiated). We hypothesized that satiating (vs. activating) justice needs would diminish the detection of perceptually ambiguous moral (vs. non-moral) words.

These experiments served three primary purposes. First, we sought to examine the importance of the motivational context in determining how and when a perceptually ambiguous lexical stimulus might be detected. Second, these experiments provided the first direct test of whether the moral pop-out effect is partly due to the motivational relevance of moral stimuli. Third, we developed ecological valid manipulations of justice motives (e.g., CrimeStoppers website, a newspaper article), to mimic how common experiences of learning about justice might influence word detection. More broadly, we sought to help bridge the relationship between morality and perception (Gantman & Van Bavel, 2015).

**Experiment 1**

CrimeStoppers is a program that advertises criminal activity to the community and allows anonymous individuals to report criminal activity. In the United States, CrimeStoppers has been responsible for over half a million arrests and several billion dollars in recovered property (www.crimestoppersusa.com). It may also serve as a regular reminder of the justice or injustice in a given community and many analogues are frequently presented online, on posters, on billboards, and in newspapers. In Experiment 1, we presented people with a CrimeStoppers advertisement in which a majority (vs. minority) of wanted murderers had been brought to justice to satiate (vs. activate) their need for justice. We examined the influence of justice needs on the subsequent detection of perceptually ambiguous moral (vs. non-moral) words. To enhance ecological validity, the CrimeStoppers ads were adapted from a real national initiative.

**Method**

**Participants and design.** Seventy-nine undergraduate students at New York University participated for partial course credit.1

**Materials and procedure.** Participants arrived in the lab and were told that the experiment was about the relationship between keeping up with current events and language skills. Participants were randomly assigned to view one of two nearly identical images for 1 min prior to completing the lexical-decision task. They were told to pay close attention as they would answer questions about it during the experiment. In both conditions participants saw an image based on real images from the CrimeStoppers program. In each image, there was an array of 11 male faces, altered to be balanced in terms of ethnicity, and all bearing a neutral expression. In the unjust condition, two of the 11 faces had the word arrested written diagonally over them in red. In the just condition, eight of the 11 faces had the word arrested written in red over their faces. The images were otherwise identical (see Figure 1). Participants completed the lexical-decision task with all letter strings presented for 40 ms, chosen a priori as ambiguous, around the threshold for visual awareness in previous experiments (see Gantman & Van Bavel, 2014).

**Lexical decision task.** The lexical-decision task was adapted from Gantman & Van Bavel (2014) and administered in DirectRT on a Dell Optiplex 760 with a 100 Hz refresh rate. Participants completed the study alone in a dimly lit room and sat approximately 16 in. from the monitor. Stimuli appeared in white letters on a black background, size 24 font in the center of the computer monitor. The experiment began with a brief tutorial with five trials of non-moral words and nonwords (apple, speic, building, kroaf, parrot) at decreasing stimulus durations (500, 300, 100, 80, and 60 ms) to allow participants to learn the task. On every trial, participants saw a fixation cross in the center of the screen for 100, 200, or 300 ms (randomized to prevent participants from feeling lulled by a repetitive rhythm). The fixation cross was followed by the stimulus letter string presented in the center of the screen for 40 ms. Finally, there was a 200 ms backward mask of ampersands that corresponded to the number of letters in the word (e.g., ‘useful’ was followed by ‘&&&&&&&&’). The screen was black until participants responded (see Figure 2). There were 82 moral/non-moral words included, which we had previously pretested by asking a separate student sample how relevant to morality each word was on a 5-point Likert scale ranging from 1 (not at all moral) to 5 (very moral) (see Gantman & Van Bavel, 2014) and 81 nonwords presented in random order. Moral words were rated as significantly more moral (M = 3.84, SD = 0.50) than non-moral words (M = 2.03, SD = 0.49), t(18) = 16.36, p < .001, η² = .94. Moral words were selected from multiple subdomains of morality. There were words related to moral mental states (e.g., right, wrong, evil, responsible, innocent), justice (e.g., just, justice, law, crime and punishment), and religious notions (e.g., god, devil, sin, religion, confess).2 Although all words were matched for frequency in the English language and word length, we have previously measured three dimensions—extremity, emotional arousal, and valence for our moral and non-moral word list—and found that moral words were rated as more emotionally arousing, and extreme (we compared absolute values of valence scores). Differences in moral versus non-moral word detection occurred over and above these dimensions when we had previously included them in statistical analyses (see Gantman & Van Bavel, 2014). All materials (including full moral and non-moral word lists) are publically available, and data for all experiments will be made available upon request online via the Open Science Framework following publication (osf.io/jmq4v).

---

1 It was determined a priori to run this experiment until the end of the semester. After one semester the sample was deemed too small so a second semester was added. We added 40 observations after analyzing the first 39.

2 The basic moral pop-out effect has been successfully replicated by an independent lab using an expanded word list (Firestone & Scholl, 2015), which also includes more general words for moral wrongdoing (e.g., atrocity, abomination) and particular moral good or bad actions or mental states (e.g., liar, lust, chaste, shame, duty, felony, adultery, faith).
Results and Discussion

Analytic strategy for lexical decision-task. Given the categorical dependent measure and mixed design, we used generalized estimating equations to estimate our regression parameters instead of ordinary least-squares regression (Zeger & Liang, 1986). This allowed us to take learning effects and other forms of interdependence among participants’ responses into account (see also Gantman & Van Bavel, 2014). Because our stimuli were presented in random order, an exchangeable correlation matrix was specified for all models (Ballinger, 2004). For analyses using generalized estimating equations models, we report unstandardized regression coefficients ($B$), standard errors ($SE$) and Wald Z’s (for a similar analytic strategy, see Stern, West, Jost, & Rule, 2013; Freeman, Johnson, Ambady, & Rule, 2010). To provide further information about effect size, 95% confidence intervals on $B$ values are also reported.

Moral pop-out effect. Following previous work (Gantman & Van Bavel, 2014) we decided a priori to use ~40 ms as a stimulus duration that would be perceptually ambiguous (i.e., close to the threshold for perceptual awareness). We found that accuracy was at 60% ($SE = .6\%$), which is below the 75% mark we had hypothesized. In other words, in this sample, people tended to have a higher threshold for visual awareness, leading to general underperformance. This creates a conservative test of our hypothesis that moral words would be detected more frequently than non-moral words. Replicating previous research on the moral pop-out effect (Gantman & Van Bavel, 2014), moral words ($M = 63\%, SE = 1\%$) were detected more frequently than non-moral words ($M = 58\%, SE = 1\%$), $B = -.12$, $SE = .02$, 95% confidence interval (CI) $[-.15, -.06]$, $p < .001$, $z = 4.72$. Overall, we replicated previous findings of the moral pop-out effect—moral words were detected more frequently than matched non-moral words.

Exposure to injustice. To investigate whether activating versus satiating justice motives moderates the moral pop-out effect,
we included the between subjects condition in the model (-1 = just world, 1 = unjust world). As predicted, the moral pop effect was qualified by a significant interaction between justice condition and word type, $B = -0.05$, $SE = .02$, 95% CI [-0.09, -0.0001], $p = .05$, $z = 1.94$. In the unjust condition, we found a significant simple main effect for moral versus non-moral words, $B = -.15$, $SE = .03$, 95% CI [-.21, -.09], $p < .001$, $z = 4.78$. Moral words were detected more frequently ($M = 61\%$, $SE = 1\%$) than non-moral words ($M = 53\%$, accuracy, $SE = 1\%$), indicating that for those exposed to injustice, there is a large moral pop-out effect. In the just condition the moral pop-out effect was diminished, $B = -.06$, $SE = .03$, 95% CI [-.13, -.006], $p = .07$, $z = 1.80$. Moral words were detected more frequently ($M = 65\%$, $SE = 1\%$) than non-moral words ($M = 62\%$, $SE = 1\%$). People exposed to injustice showed a larger moral pop-out effect than those whose justice needs were satiated, when letter strings were perceptually ambiguous (see Figure 3).3

Fear versus motivation. To further examine the role of the justice motive, we tried to rule out the alternative possibility that our injustice condition simply induced more fear (e.g., “there’s a murderer on the loose”) than our just condition. It was theoretically possible that fear could have broadly enhanced perceptual intake (Susskind et al., 2008). However, there was no main effect of justice condition: learning that the majority of criminals have been caught or not did not enhance detection of words in general ($p = .16$). Moreover, the significant interaction between justice condition and word type reported above suggests that activating versus satiating just world needs selectively enhances versus diminishes the detection of moral words. Taken together, the data were inconsistent with the selective detection of moral words in general ($M = 61\%$, $SD = 22\%$) were correctly detected more frequently than non-moral words ($M = 54\%$, $SD = 22\%$), $t(37) = 4.80$, $p < .001$, Cohen’s $d = .33$. In the just condition, moral words ($M = 64\%$, $SD = 22\%$) were correctly detected more frequently than non-moral words, ($M = 61\%$, $SD = 22\%$), $t(40) = 2.04$, $p = .05$, $d = .15$. Thus, the moral pop-out effect was larger in the unjust condition than the just condition.

Experiment 2

In Experiment 2, we sought to replicate and extend the results of Experiment 1 in several important ways. Given that we found that words were detected with only 60% frequency at 40 ms durations—less than we had hypothesized—we decided to include a larger range of stimulus durations. This allowed us to rigorously test whether motivation would shape the detection of moral words only when the letter strings were perceptually ambiguous. Critically, if moral motives shape the detection of moral words, we should find that moral motives selectively increase the detection of moral lexical content only when stimuli are presented close to the threshold for perceptual awareness. Accordingly, letter strings were presented from 20 to 100 ms at ~10 ms intervals. We also manipulated justice concerns with two nearly identical false New York Times articles. This offered additional assurance that justice concerns—rather than something specific to the Crimestoppers manipulation—accounted for the change in moral word detection. It also offered additional evidence of ecological validity since millions of people learn about justice or injustice by reading the newspaper.

Methods

Participants and design. Eighty-five undergraduate students at New York University participated for partial course credit.4

Materials and procedure. Participants arrived at the lab and were told that the experiment was about the relationship

---

3To help communicate the nature of the interaction, we created a separate average accuracy score for moral and non-moral words. Using a paired samples $t$ test, we analyzed the differences in moral versus non-moral word detection in each of our between-subjects conditions. In the unjust condition, moral words ($M = 61\%$, $SD = 22\%$) were correctly detected more frequently than non-moral words ($M = 54\%$, $SD = 22\%$), $t(37) = 4.80$, $p < .001$, Cohen’s $d = .33$. In the just condition, moral words ($M = 64\%$, $SD = 22\%$) were correctly detected more frequently than non-moral words, ($M = 61\%$, $SD = 22\%$), $t(40) = 2.04$, $p = .05$, $d = .15$. Thus, the moral pop-out effect was larger in the unjust condition than the just condition.

4It was determined a priori to run this study until the end of the semester with a target of ~40 participants per condition. All subjects were included for analysis.
between keeping up with current events and language skills. The concept of morality was never mentioned. Participants were randomly assigned to read one of two short news articles prior to completing the lexical-decision task. They were told to pay close attention as they would answer questions about the article during the experiment. In both conditions, participants read a fictitious *New York Times* article (see Figure 4). The article detailed the story of a homeless man who runs to the rescue of a woman being mugged at knifepoint. When he chases down the perpetrator, the man stabs him fatally and onlookers pass his body for hours. In the unjust condition, the assailant is never caught. In the just condition, police catch the man responsible and hold him in prison without bail. The articles in each condition were identical until the final line. After reading the article participants completed the lexical-decision task described in Experiment 1. To critically test whether moral words are more frequently detected than non-moral words only when presented ambiguously, we decided to include the full range of stimulus durations, with words presented from 20 to 100 ms at 10 ms intervals, presented randomly.

**Results and Discussion**

*Overall visibility curve.* As expected, participants overall accuracy increased as the letter strings were presented on the screen for longer durations. At short durations (20–30 ms), participants detected words with 36% accuracy ($SE = 1\%$), whereas at moderate durations (40–60 ms), participants detected words at 71% accuracy ($SE = 1\%$), and at long durations (70–100 ms), participants detected words with 90% accuracy ($SE = 1\%$). In short, words presented for moderate durations were ambiguous—detected with accuracy halfway between chance (50%) and perfect accuracy (100%). We ran a logistic regression to fit the log odds of word categorization (word or nonword) on stimulus duration treated as a continuous variable. We found a significant effect of stimulus duration on accuracy, $B = 0.05$, $SE = .002$, Wald $\chi^2 = 1052.23$, $p < .001$, $r = .43$. We had selected 40–60 ms a priori as moderate durations because (a) overall accuracy rates hover around 75% (here 71%) which is close to halfway between 50% and 100% accuracy, and (b) we have previously identified these durations as perceptually ambiguous (Giantam & Van Bavel, 2014).

*Moral pop-out effect.* We have previously found that moral words were detected more frequently than non-moral words, especially when they were perceptually ambiguous (Giantam & Van Bavel, 2014). To formally test whether word detection was different at ambiguous (40–60 ms) durations versus nonambiguous (i.e., fast 20–30 ms and slow, 70–90 ms) durations, we regressed categorization accuracy (word, nonword) against word type (non-moral = -1, moral = 1) and stimulus ambiguity (nonambiguous = -1, ambiguous = 1). As predicted, we found a significant interaction effect between word type and ambiguity, $B = -2.02$, $SE = .19$, $95\% CI [-2.45, -1.59]$, $z = 2.08$. As predicted, we found a marginally significant three-way interaction effect, $B = -2.02$, $SE = .06$, $95\% CI [-1.93, -1.11]$, $p < .04$, $z = 2.02$. At ambiguous durations, for participants in the unjust condition (when the killer was still at large), there was a large simple main effect of moral versus non-moral words, $B = -2.08$, $SE = .06$, $95\% CI [-2.09, -1.99]$, $p < .001$, $z = 4.83$. Moral words were detected more frequently ($M = 79\%; SE = 2\%$) than non-moral words ($M = 69\%; SE = 2\%$). At ambiguous durations, for participants who read that the killer was caught, the simple main effect was diminished, $B = -1.12$, $SE = .06$, $95\% CI [-1.23, -.007]$, $p = .04$, $z = 2.08$. Moral words were detected more frequently ($M = 71\%; SE = 2\%$) than non-moral words ($M = 66\%; SE = 2\%$). In other words, people exposed to an unjust world showed a larger moral pop-out effect than those whose justice needs were satiated. When words are presented unambiguously, (i.e., for 20–30 ms or 70–100 ms) there is no significant interaction effect between just world condition and word type, $B = .02$, $SE = .03$, $95\% CI [-.05, .08]$, $p = .53$, $z = .53$ (see Figure 6).

*Fear versus motivation.* To further examine the role of the justice motive, we again tried to rule out the alternative possibility that our injustice condition simply induced more fear (e.g., “there’s a murderer on the loose”) than our just condition. It was theoretically possible that fear could have broadly enhanced perceptual intake (Susskind et al., 2008). However, there was no main effect of justice condition: learning that the killer is at large versus captured did not enhance detection of words in general ($p = .14$). Moreover, the significant interaction between justice condition and word type reported above suggests that activating versus satiating just world needs selectively enhances versus diminishes the detection of moral words. Taken together, the data were inconsistent

---

5 We have found in multiple experiments that at durations too fast to see there is a bias to select nonword. We suspect that this is because it feels more natural to say that an unseen stimulus is a nonword (since it was experienced as nothing) than a word.

6 To help communicate the nature of the interaction, we created a separate average accuracy score for moral and non-moral words presented at ambiguous durations. Using a paired samples *t*-test, we analyzed the differences in moral versus non-moral word detection in each of our between-subjects conditions. In the unjust condition, moral words ($M = 79\%; SD = 19\%$) were correctly detected more frequently than non-moral words ($M = 67\%; SD = 22\%$), $t(45) = 5.49$, $p < .001$, $d = .56$. In the just condition, moral words ($M = 71\%; SD = 17\%$) were correctly detected more frequently than non-moral words ($M = 65\%; SD = 20\%$), $t(38) = 2.32$, $p = .03$, $d = .31$. Thus, the moral pop-out effect was larger in the unjust condition than the just condition.
Figure 4. Justice motive manipulation for Experiment 2. The vignettes detail a homeless man who attempts to save a woman from assault. He is, in turn, attacked and killed by her assailant and left unaided and ignored by many passersby. The two vignettes are identical up until the final line which reveals either that the killer was brought to justice, having been captured and held in prison without bail (below) or that justice has not been served as the killer had not yet been found (above). See the online article for the color version of this figure.
with the notion that fear simply enhanced word detection and instead supported the motivational explanation.

We also again evaluated the possibility that fear leads to changes in moral word detection by increasing the accessibility of words with negative valence. When we entered valence into the model, we found that negative words were detected marginally more often than positive words ($p = .06$). However, adding fear to the model did not eliminate the moral pop-out effect ($p < .001$) or the interaction effect between justice condition and word type ($p = .04$) at ambiguous durations. No significant interaction effects between valence and moral versus non-moral words, or between valence and just condition were detected at ambiguous durations (all $p s < .56$). As such, increased fear did not fully explain the selective detection of moral words.

**General Discussion**

This article provides the first evidence that contextual social motives alter the detection of moral content. We previously theorized that moral words “pop-out” because they have chronic motivational value (Gantman & Bavel, 2016). In two experiments, we found that satiating (vs. activating) justice needs can diminish the moral pop-out effect. People who saw a CrimeStoppers ad in which a number of wanted criminals had been arrested (vs. not) were less likely to detect moral (vs. non-moral) words (Experiment 1). Similarly, people who read about a killer who had been caught (vs. at large) were less like to detect moral (vs. non-moral) words—only when the words were presented ambiguously (Experiment 2). We presented stimuli at durations that have been previously identified as short enough to reduce visibility (Kouider & Dehaene, 2007; Gelskov & Kouider, 2010) and only found effects of the motivation manipulation when stimuli were presented close to the threshold of perceptual awareness (i.e., when visibility was sufficiently degraded). These experiments suggest that satiating justice needs can alter word detection in a top-down fashion—leading to less frequent detection of moral words.

The current research expands the scope of moral psychology by bridging the field with the study of visual word recognition, an aspect of visual perception more generally (Gantman & Bavel, 2015). We argue that often morality “wins out” in conscious awareness (Gantman & Bavel, 2014), especially when moral motives are activated (vs. satiated). To do this, we applied two domain-general principles of goal activation to the domain of morality: (a) active goals promote accessibility of goal relevant stimuli and (b) satiated goals lead to decrements in accessibility of goal related information. Representations of valuable objects are more accessible (Balceitis, Dunning, & Granot, 2012) and active goals shape value (Förster, Liberman, & Friedman, 2007). We suggest that the moral pop-out effect (and its context sensitivity) fit within this framework. Specifically, early lexical processing of moral content allows for a lower threshold to conscious awareness. Although we would like to suggest that the intersection of moral psychology and perception is an exciting new avenue for future research, we in no way mean to suggest a “moral module.” The motivational effects reported here are part of a domain-general process in which motives tune perceptual processing toward goal-relevant stimuli. Multiple processes have been implicated in moral judgments, decisions and actions, including mind perception, (Chakroff & Young, 2015), impression formation (Uhlmann, Pizarro, & Diermeier, 2015), face perception (Singer, Kiebel, Winston, Dolan, & Frith, 2004), reward processing (Delgado, Frank, & Phelps, 2005), and emotion and reasoning (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; for a review, Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015).

We recognize that our current method cannot fully distinguish between perceptual detection per se and cognitive accessibility. However, all of the reported effects remain unchanged in our model when we include RTs, which have historically been used to

![Figure 5](image-url)

**Figure 5.** Moral words are recognized more frequently than non-moral words—especially when words are perceptually ambiguous (at 40–60 ms stimulus durations). Frequency of correct categorization of letter strings as words (i.e., detection; Y-axis) increases as stimuli are presented for more time on screen (X-axis). Throughout, overall means are displayed for ease of interpretation despite interdependence.

![Figure 6](image-url)

**Figure 6.** People exposed to just (vs. unjust) information via a false New York Times article showed a diminished moral pop-out effect when letter strings were presented around the threshold for visual awareness (40–60 ms). Frequency of correct categorization of words (Y-axis) was greater for moral versus non-moral words in the unjust condition, but not in the just condition (X-axis). Overall means are displayed for ease of interpretation despite interdependence. Bars represent standard errors.
measured cognitive accessibility in lexical decision tasks (e.g., Neely, 1977). Moreover, a simple cognitive accessibility explanation should predict that justice needs increase detection across all stimulus durations. Given that we only observe effects of motivation on the detection of moral words when they are presented ambiguously, we suspect that cognitive accessibility alone cannot fully explain how exposure to justice moderates the moral pop-out effect. As such, we suspect that moral relevance may play a role in word detection. This is clearly an important direction for future research.

**Alternative Mediators of Moral Pop-Out**

**Lexical properties of moral versus non-moral words.** Lexical and symbolic stimuli are distinct from detecting other stimuli (e.g., objects, faces, or colors). First, word reading in English is left to right. Second, words do not resemble their referents the way that seeing a picture of a ball resembles an actual ball. Indeed, it is difficult to know what the visual component of a word like “just” or “should” might be. As such, it is essential that strictly lexical properties of the stimulus (that do not have to do with word meaning) do not explain our effects. When we constructed the moral and non-moral word lists, we ensured that word length and frequency did not differ between the two groups (Gantman & Van Bavel, 2014). These are, however, population values (Davies, 2008) rather than sample means. To make absolutely certain that small differences were not influencing our effects, we included word length and frequency into our model. In Experiment 1, when we include both word length and frequency into the model, we found that the significant interaction between word type and just world condition remains significant ($p = .059$). In Experiment 2, when we include both word length and frequency into the model, we find that our significant three way interaction between ambiguity, word type and justice condition remains ($p = .05$). As such, lexical properties like word frequency and length cannot explain differences in moral word detection.

To further ensure that lexical properties do not explain the motivational sensitivity of moral word detection, we have also examined the role of the number of syllables, as well as the number of orthographic and phonological neighbors of the words (Marian, Bartolotti, Chabal, & Shook, 2012). While there is a significant effect of number of syllables, as would be expected ($p = .04$), the number of syllables in the words accounts for neither the moral pop-out effect ($p < .001$), nor the interaction between word type and justice needs ($p = .058$). Likewise, while the number of orthographic neighbors to the words marginally predicts whether the word is correctly categorized ($p = .09$), this does not account for the moral pop-out effect ($p < .001$), or the interaction between word type and justice needs ($p = .066$). Finally, we also find that the number of phonological neighbors, which marginally predicts whether a word will be correctly detected ($p = .088$), does not account for the moral pop-out effect ($p < .001$), or the reported interaction between word type and justice needs ($p = .059$). Thus, the number of syllables, orthographic, or phonological neighbors each cannot account for the motivational sensitivity of moral word detection.

**Moral pop-out at Trial 1.** We also ruled out the possibility that participants merely show moral pop-out because the moral words may be more related to each other than the non-moral words, and so prime future moral words as the experiment goes on. Although semantic priming cannot fully explain the effects of the justice motive manipulations presented here, we tested whether we see the moral pop-out effect on the very first trial when no prior moral trials could have primed it. We examined Experiment 1, where all letter strings were presented at ambiguous durations, so that the full dataset, rather than a subset, could be used. There is no significant three-way interaction between justice motive condition, word type, and order, $p = .54$, so subsequent analyses were merely exploratory. We coded order such that the first trial was set to zero, allowing us to look at differences in moral pop-out on the first trial in the experiment. In the unjust condition, on the first trial, there is a simple main effect of word type, $p = .002$. In the just condition, on the first trial, there is no simple main effect of word type, $p = .53$. When justice motives are activated (but not satiated), moral pop-out occurs on the very first trial.

Taken together, we suggest that moral content affects word detection in a way that is sensitive to moral motives only when stimuli are perceptually ambiguous. That said, so far we have only found the moral pop-out effect—and its modulation by moral motives—with lexical stimuli. Others have shown effects of moral perception with nonword stimuli (see Gantman & Van Bavel, 2015). For instance, neutral faces associated with negative (vs. positive or neutral) gossip (e.g., “told a racist joke at a party”) dominate longer in binocular rivalry (Anderson, Siegel, Bliss-Moreau, & Barrett, 2011) and changes in deviations from pure whiteness are perceived differently by people who vary on trait and state disgust (Sherman, Haidt, & Clore, 2012). That said, this is a relatively new and unexplored aspect of moral cognition and it is premature to conclude that motivation will generalize beyond lexical content. However, future work should follow-up this basic finding using a variety of different experimental methods and stimuli.

**Future Directions**

We see many avenues for future research. First, future research would strongly benefit from using other types of visual stimuli such as faces and objects, which would allow for the possibility of generalized moral perception beyond moral word detection. Second, we suspect that the influence of moral concerns on perception is not limited to vision, but may extend to other sensory modalities. For example, moral content might be especially difficult to ignore when presented against competing auditory speech (i.e., in a dichotic listening task). In addition, future research could examine whether this effect generalizes to other moral motives, for example, unity, hierarchy (Rai & Fiske, 2011), social order, and communal solidarity (Janoff-Bulman & Carnes, 2013). Finally, further research would benefit from using multiple methods of analysis to investigate the process underlying the moral pop-out effect. For example, using neuroscience methods such as electroencephalography could help elucidate where in the processing stream the visual system shows attunement to moral content. In the case of the lexical-decision task presented here, we would expect differences to emerge at word categorization, once some preliminary lexical processing has taken place (e.g., P300) suggesting participants allocate extra attentional resources, boosting the motivationally relevant moral content up to conscious awareness. This might change for other, nonlexical stimuli where semantic processing is...
not involved. Further understanding the role of motivation and its effect on perception will help elucidate when, how, and in what context people detect moral content.

Implications

Motivated moral perception. The current research is also consistent with a broader body of work on motivated perception (see Bruner & Goodman, 1947; Balcetis & Dunning, 2010; Balcetis et al., 2012) and top-down effects on perception more generally (Lupyan & Ward, 2013; Adams, Ambady, Nakayama, & Shimojo, 2010; Balcetis & Lassiter, 2010; Gilbert & Li, 2013). Traditional wisdom asserts that prior states like beliefs and desires play no role in determining the content of early vision, and so visual systems are an autonomous module (Fodor, 1983), and thus cognitively impenetrable (Pylyshyn, 1999; Firestone & Scholl, 2015).

In that vein, it is has been suggested that the moral pop-out effect is due to semantic priming, such that individuals who complete the lexical-decision task in our experiments become primed with the concept of morality throughout the course of the experiment (Firestone & Scholl, 2015). We agree that semantic processing must be at work in this experimental procedure (how else would our participants know words like kill and die?) however, a semantic priming explanation cannot explain the current research. We directly manipulated justice motives using virtually identical semantic content and found diminished moral word detection. Not only does a semantic priming account fail to explain why we find motivational sensitivity of moral word detection, but we have also ruled out a number of alternative explanations empirically (see Gantman & Van Bavel, 2016). We did not find evidence that these effects are due to fear, and we found evidence for the moral pop-out effect (across participants) at the first trial. This offers convergent evidence that the activation of a moral motive, such as the need for justice, may alter our “moral perception” and make us more or less likely to detect moral lexical stimuli in our environment in a top-down manner.

Conclusion

We suggest that this work is part of an emerging trend in the study of moral psychology. There is emerging evidence that morally relevant content can influence not only word detection but perception more generally (for a review, see Gantman & Van Bavel, 2015). In the current article, we used core principles of motivation, finding evidence that satiating (vs. activating) justice motives can reduce the frequency with which moral (vs. non-moral) words reach perceptual awareness. These findings have implications for models of moral psychology, understanding factors that enhance word detection generally, as well as avenues for understanding when and what kind of information reaches the awareness of people in their daily lives as they encounter seemingly limitless text.

References


Received July 21, 2016
Revision received August 23, 2016
Accepted September 9, 2016