Journal of Experimental Psychology: General
2016, Vol. 145, No. 2, 131–146


0096-3445/16/$12.00 http://dx.doi.org/10.1037/xge0000132

# Perceptual Dehumanization of Faces Is Activated by Norm Violations and Facilitates Norm Enforcement

Katrina M. Fincher and Philip E. Tetlock
University of Pennsylvania

This article uses methods drawn from perceptual psychology to answer a basic social psychological question: Do people process the faces of norm violators differently from those of others—and, if so, what is the functional significance? Seven studies suggest that people process these faces different and the differential processing makes it easier to punish norm violators. Studies 1 and 2 use a recognition-recall paradigm that manipulated facial-inversion and spatial frequency to show that people rely upon face-typical processing less when they perceive norm violators' faces. Study 3 uses a facial composite task to demonstrate that the effect is actor dependent, not action dependent, and to suggest that configural processing is the mechanism of perceptual change. Studies 4 and 5 use offset faces to show that configural processing is only attenuated when they belong to perpetrators who are culpable. Studies 6 and 7 show that people find it easier to punish inverted faces and harder to punish faces displayed in low spatial frequency. Taken together, these data suggest a bidirectional flow of causality between lower-order perceptual and higher-order cognitive processes in norm enforcement.

*Keywords:* punishment, morality, holistic processing, face processing, social functionalism

*Supplemental materials:* http://dx.doi.org/10.1037/xge0000132.supp



This article uses concepts, methods, and findings from perceptual psychology to define—conceptually and operationally—a new social psychological construct: perceptual dehumanization. Our analytical starting point is that humans are a deeply social species and that navigating the human social world requires response flexibility: a capacity to treat different types of people differently. This assumption implies there would be adaptive value to mechanisms that facilitate social response differentiation. Combining this functionalist perspective with empirical work on the centrality of face perception in social interaction (Haxby, Hoffman, & Gobbini, 2000), we hypothesize a causal chain linking social information, visual perception, and social behavior. The first link posits that the engagement of face-typical perceptual processing can be reduced by social cues signaling that the human being in question poses a threat to the social order and deserves harsh treatment. The second link suggests that the disruption of face-specific processing makes it easier for people to support harsh treatment. We propose that through this process perceptual dehumanization facilitates response differentiation. We test these hypothesizes in the domain of punishment.

Katrina M. Fincher and Philip E. Tetlock, Department of Psychology, University of Pennsylvania.

Support for this research comes from the Annenberg Endowed Chair at the University of Pennsylvania, the Wharton Risk Center Russell Ackoff Doctoral Student Fellowship and the Wharton Behavioral Lab. We thank Martha Farah, Rob Kurzban, Paul Rozin, Dan Simon, Nancy Kanwisher, and Jon Baron for their constructive comments on earlier drafts of this article. We would also like to acknowledge the extraordinarily helpful contributions of the Editor Max Weisbuch.

Correspondence concerning this article should be addressed to Katrina M. Fincher, University of Pennsylvania Psychology Department, 3720 Walnut Street, Philadelphia, PA 19104. E-mail: katrina.fincher@gmail.com

## Face-Specific Processing

A preference for face-like stimuli appears to be innate. Babies react to facial distress (Johnson, Dziurawiec, Ellis, & Morton, 1991) and newborns preferentially orient toward stimuli with face-like first-order relations (Johnson, Dziurawiec, Ellis, & Morton, 1991; Mondloch, Lewis, Budreau, Maurer, Dannemiller, Stephens, & Kleiner-Gathercoal, 1999). As adults, humans can recognize an inordinate number of individuals quickly, even when these individuals have different hairstyles, are presented from different viewpoints, in different lighting, or are even of different ages, that is, unless face typical processing is in some way disrupted (Ellis, Bruce, & De Schonen, 1992; Johnson, 2005). Put more reductively, face recognition stands out among humans' perceptual skills. This skill in recognizing faces is often attributed to specialized processing (Richler, Cheung, & Gauthier, 2011).

A large body of evidence suggests that face and object recognition involve qualitatively different processes—with faces eliciting more holistic and configural processing (i.e., as an integrated precept and based on the relationship between parts) and objects more analytical processing (i.e., attribute-based; Avidan, Tanzer, & Behrmann, 2011; Farah, Levinson, & Klein, 1995; Farah, Wilson, Drain, & Tanaka, 1995; Kanwisher, Tong, & Nakayama, 1998; Richler, Cheung, & Gauthier, 2011).

There are two competing modular accounts that explain the marked skills associated with face processing. The specific or encapsulated modularity account suggests that holistic processing developed to facilitate the processing of faces and that only face stimuli elicit holistic processing (Haxby et al., 2001; Haxby, Hoffman & Gobbini, 2000; Kanwisher, Downing, Epstein, & Kourtzi, 2001; Treisman & Kanwisher, 1998). This argument builds upon the premise that as a highly social species, facial recognition is particularly important to humans and, therefore, humans (and some primates) have acquired specialized mechanisms to recognize other's faces (Haxby, Hoffman, & Gobbini, 2000).

The domain general or unencapsulated modularity account (also known as expertise theory) suggests that holistic processing represents a more general strategy (Bukach, Gauthier, & Tarr, 2006). Studies supporting this domain-general or expertise argument suggest that face-typical processing could theoretically occur across an unbounded number of categories if the viewer has perceptual expertise and the targets require within category discrimination of complex stimuli (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Gauthier, Williams, Tarr, & Tanaka, 1998). Past work has shown that putatively face-typical processing mechanisms can occur for categories such as birds, cars, dogs, and even novel object classes such as Greebles (Gauthier et al., 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). While the domain general account suggests that face-typical processing may occur for more stimuli than just faces, both accounts maintain that faces should elicit face typical processing.

Although face perception research has thus far focused primarily on the stimulus features, social features may play an important role as well. If face perception emerged because of social needs, there is little reason to believe face specific processing occurs independently of selection pressures in the social world. An emerging body of work suggests that there are numerous instances where social information leads to attenuations in the typical processing of faces (Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007; Michel, Rossion, Han, Chung, & Caldara, 2006; Van Bavel, Packer, & Cunningham, 2011).

For example, group affiliation appears to moderate the perceptual processes a face engages. While many studies have shown that same-race faces are processed more "typically" than other-race faces (e.g., Michel et al., 2006), recent work suggests the difference is not only because of expertise as it occurs when the physical features of the face are held constant. For example, MacLin and Malpass (2001, 2003) documented differences in the processing of the same racially ambiguous face when it was believed to belong to a member of the same race. Moreover, group affiliation alone appears sufficient to elicit changes in processing (Bernstein, Young, & Hugenberg, 2007); minor indications of groups, such as university affiliation, have attenuated the facial inversion effect.[1]

However, this line of research has yet to explore the social functions of these changes. Face-typical processing marks the perceptual categorization of one human being by another as human (Hugenberg et al., 2015). Anything disrupting that processing can cause perceptual dehumanization.[2] And just as work on dehumanization in the classic sense has stressed the social functions of denials of humanity (e.g., Bandura, 1999; Haslam, 2006; Kelman, 1973; Opotow, 1990), we suggest that perceptual dehumanization can facilitate social goals. Figure 1 identifies the two process linkages that define perceptual dehumanization: social attitudes alter perceptual processing that in turn influences social behavior.

## A Social-Functionalist Framework for Attenuations in Face Perception

Although we have argued that face processing is a distinct perceptual phenomenon with deep implications for social interactions, we have not specified the functionalist preconditions for attenuations of face processing—or the consequences of such reductions. We predict that perceptual changes will facilitate response differentiation that supports the social order. And these changes are especially likely when we see people who threaten the social order and who, in revealing common parlance, "need to be punished."

We suggest that these targets of punitiveness provide a useful lens for viewing the workings of perceptual dehumanization. Punishment is a cross-cultural universal that invariably involves inflicting harm on a fellow human being and justifying that act as in defense of the social order (Aberle et al., 1950; Tetlock, 2002). Many scholars have argued that the ability to detect and punish cheaters is essential to maintaining all social systems and the act of punishment itself reifies the social order (Durkheim, 2008; Fehr & Fischbacher, 2004; Fehr & Gächter, 2000; Tetlock et al., 2007). In this view, maintaining social order requires a human capacity to punish. Although prior work suggests there are automatic mechanisms for detecting norm violators (Cosmides, Barrett, & Tooby, 2010; Cosmides & Tooby, 1992), that work has not examined the mechanisms that facilitate punishment itself. The theory of perceptual dehumanization fills this void by specifying such a mechanism.

Second, punitive behavior is inherently coercive and strikingly different from most human social interactions (Foucault, 1978; Pinker, 2011) which place a premium on voluntary cooperation (Loewenstein & Small, 2007). All of which raises the question: how do humans manage to engage so frequently in behavior that, were it not collectively sanctioned, would be regarded as profoundly antisocial? However, for the imprimatur of the state, jailing would be seen as kidnapping and execution as murder. This disconnect is easiest to observe in the disgust people express for the punitive practices of other cultures (e.g., American attitudes toward Sharia law or European attitudes toward the U.S. penal system) and perhaps the strongest disconnects emerge from our own disgust for historical punitive practices:

[1] There is a long history, beginning with the "new look," (Bruner, 1957), which suggests social processes can alter perceptual ones. Past work linking social phenomena and perception has focused on the perceptual mechanisms underlying these changes. This article uses concepts, methods, and findings drawn from perceptual psychology to help us conceptually and operationally define a new social psychological construct, perceptual dehumanization. It does not claim to advance a new theory of face-typical processing—or to be a contribution to our understanding of basic perceptual processes beyond the point that such processes can be as affected by social cues and the social functional context within which faces are embedded.

[2] On purely logical grounds, perpetual dehumanization is a form of dehumanization. Perceptual dehumanization is defined as any atypical processing of faces. While, the specific perceptual mechanisms that are being disturbed are an interesting and potentially important question, the construct itself is defined as denial of human like processing. Although, there important questions about how this basic form of dehumanization may relate to more explicit forms of dehumanization, and they are a very interesting topic for future research, they are not the focus of the present article.

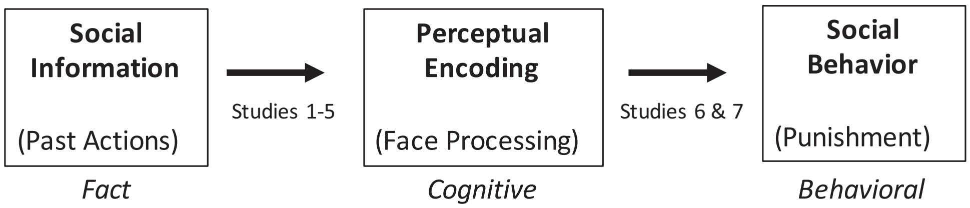


*Figure 1.* A visualization of the two causal links that support the theory of perceptual dehumanization.

> On 1 March 1757 Damiens the regicide was condemned "to make the amende honorable [. . .] the flesh will be torn from his breasts, arms, thighs and calves with red-hot pincers, his right hand, holding the knife with which he committed the said parricide, burnt with sulfur, and, on those places where the flesh will be torn away, poured molten lead, boiling oil, burning resin, wax and sulfur melted together and then his body drawn and quartered by four horses and his limbs and body consumed by fire, reduced to ashes and his ashes thrown to the winds" (Foucault, 1978, P.1).

Although modern readers of Discipline and Punishment recoil from its lurid descriptions, the account was taken from the popular paper Gazette d'Amsterdam—to say nothing of the vast crowd who gathered to see the spectacle. In highlighting the disconnect between historic and modern punitive attitudes, Foucault's account raises the question: what mental mechanisms allow a highly co-operative and social species to become indifferent to (or even enjoy) the suffering of others? This article cannot fully solve this social-functionalist puzzle but it does begin the process of assembling certain promising pieces.

Our inquiry centers on face perception: Do people process the faces of norm violators differently from those of others? And, if so, what is the functional significance of this differential processing? Does it make it easier to punish norm violators? To test the theory of perceptual dehumanization and establish this functionalist relationship, it is equally important to establish that changes in perception translate into changes in behavior and that changes in the social standing of human beings translate into changes in perception. We test the two parts of this theory separately, in two series of studies.

## Overview of Studies

Series 1 establishes that the desire to punish leads to an attenuation in face typical processing. In this series of studies, the dependent variables build on work on face processing and vary across studies to establish convergent validity and explore potential mechanisms. The independent variables build on work on punishment and aim to isolate punitive motives from other negative sentiments.

Studies 1 and 2 demonstrate an inhibition in the typical processing of perpetrators' faces. These studies use complementary methodological approaches. Study 1 decreases face-typical processing using facial inversion, whereas Study 2 increases face-typical processing using a low spatial frequency filter. By manipulating face-typical processing in opposite directions, we can rule out artifacts such as floor and ceiling effects.

Studies 3, 4, and 5 establish that: (a) punitive motives drive changes in face processing and; (b) configural processing is one of the mechanisms underlying perceptual changes. Study 3 rules out the valence of the action as a potential confound. It does so by holding the action constant and then identifying the face as that of either the victim or the perpetrator. If perceptual dehumanization effects are driven by punitive motives, one should observe less face-typical processing for the perpetrators than for the victims of any given action. Face-typical processing is measured using a facial-composite task, which more clearly isolates changes in configural processing from other possible perceptual mechanisms.

Studies 4 and 5 test the link between perceptual dehumanization and punishment by applying the culpable control model (Alicke, 2000; Alicke, Rose, & Bloom, 2011). It is well established that blame is a necessary condition for punishment. Therefore, perceptual dehumanization should occur when people see the perpetrator of the harm as blameworthy. As in other models of blame, the culpable control model posits blame to be rooted in volitional control. Therefore, Study 4 manipulates blame by varying whether the actions are intentional or unintentional and Study 5 manipulates blame by varying whether the perpetrator had the genetic-neurological capacity to act otherwise. Studies 4 and 5 both use facial offsetting to show as conclusively as possible that the changes are because of attenuation in configural processing.

Across these studies, the theory of perceptual dehumanization predicts a reduction in indicators of face-specific processing of perpetrators. This means we should consistently observe an interaction between the mechanisms of face typical processing and the social information on whether the target individual deserves to be harmed (in this case, punishment).[3] If we consistently observe attenuations in face-typical processing it would suggest social information can attenuate face perception, supporting our first link.

Series 2 establishes that reductions in face typical processing translate into increased punitiveness. The dependent variables of these studies build on social psychological work on punishment and measure changes in punitive drive. The independent variables build on perceptual work on face processing. Study 6 asks people

[3] Specifically, we predict that manipulations that decrease face-typical processing will have a smaller effect when the faces are associated with norm violators. Conversely, manipulations that increase face-typical processing will have a larger effect when the faces are associated with norm violators. However, we do not make a prediction about how this interaction will occur (e.g., greater accuracy at identifying the inverted faces of norm violators vs. reduced accuracy at identifying the upright faces of norm violators), we are merely predicting that an interaction will occur. Moreover, given that different manipulations of punitiveness may alter the salience of the norm violators relative to control conditions, it is possible that the interaction will work differently in different studies. However, what is relevant to our hypothesis is the interaction between face processing and manipulations of punitive sentiments.

to make punishment judgments for upright and inverted faces. Study 7 asks them to make punishment judgments for low and normal spatial frequency faces. As in Studies 1 and 2, this complementary-methodology approach was taken to rule out changes because of measurement artifacts (floor and ceiling effects)—but also to minimize demand effects (e.g., subjects might infer they were "supposed to" treat people whose faces were presented in unusual ways more harshly). In aggregate, these studies document the functional links among observer goals, target behavior, and face processing.

## Series 1

Studies 1 through 5 use a multimethod array of techniques to answer: Do people process the faces of norm violators differently from other faces? These studies test the hypothesis that the desire to punish an individual inhibits face typical processing, that is, the face is perceptually dehumanized. Study 1 and 2 test this idea by exploring whether markers of typical face perception interact with negative social cues that would typically induce observers to feel that a target actor deserves to be punished.

## Study 1

The theory of perceptual dehumanization predicts that the faces of norm violators are perceived differently from those of others. Study 1 uses the facial inversion effect to gauge the degree to which information about prior behavior inhibits face-typical processing. The facial-inversion effect illustrates face-typical processing: a 180 degree rotation of faces impairs recognition much more than a 180 degree rotation of comparably complex nonface objects (Farah, Tanaka, & Drain, 1995; Scapinello & Yarmey, 1970; Yarmey, 1971). An inhibition of the inversion effect suggests a disruption of face-typical processing. Our theory predicts a reduction in the face-inversion effect for faces linked to negative but not positive actions.

### Method.

***Participants.*** Participants were 48 (28 men and 20 women) students at a northeastern university who participated in partial fulfillment of a course requirement. Data was collected over a 2-week period. Sample size was determined by the number of students who volunteered to participate in the 12 days between the study posting and the end of the term. The average age of participation was 20.4 ($SD$ = .9) and average political ideology leaned liberal ($M$ = 5.2 on a 7-point scale).

***Stimuli.*** Seventy-two face stimuli were taken from Ballew and Todorov (2007). Our images were restricted to White male runner-ups in gubernatorial elections. Images were cropped to 150 × 215 pixels, placed on a standard background, and converted to grayscale.

***Procedure.*** We used a recognition memory task to measure perceptual dehumanization. For every participant, the experiment consisted of six blocks, each consisting of 12 face-action associations. As Figure 2 shows, each block consisted of a study phase and a test phase. In three blocks, faces appeared upright during both the study and test phases; in three blocks, the faces appeared as inverted during both the study and test phases.

In the study phase, participants learned 12 face-action pairings by simultaneously seeing a single face (either upright or inverted) and a single action on a screen for 8.5 s. Participants were told the individual in the photograph had performed the action described below the photograph. During the test phase, actions appeared serially below the face array, and the participant identified the face originally linked to the action in the learning phase from the array. Participants were asked to identify the correct face from the array by indicating the number associated with the face. Face arrays were created by randomly assigning each of the 12 faces a number.

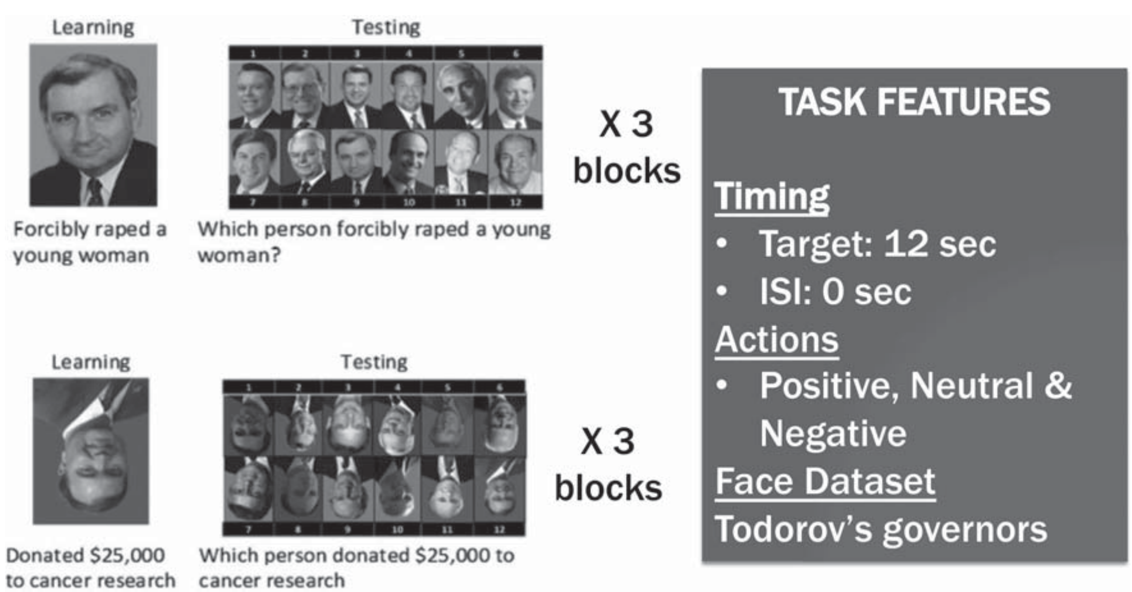


*Figure 2.* A description of the procedure used in Study 1. In the learning phase individuals saw 12 face-action pairings. In the testing phase, individuals saw the 12 actions appear below the faces. Participants saw three blocks of upright faces and three blocks of inverted faces.

Actions systematically varied in moral valence. Negative actions, norm violations, were restricted to actions punishable by five or more years in prison according to mandatory minimum sentencing guidelines. Whenever possible, the greater salience of negative things (Rozin & Royzman, 2001), we defined positive actions as intentional actions that resulted in 2 to 3 times the positive impact as the negative actions caused harm. For example, a negative action of stealing $10,000 dollars would be paired with a positive action of donating $30,000, or murdering one individual was paired with saving the lives of two individuals. Neutral actions were defined as actions that had no obvious impact on others. For example, "wore a blue shirt to dinner."

To control for difficulty or learning effects, the order of blocks was randomized across participants. Faces and actions were randomly assigned to one of six blocks. For each participant, the order of faces was randomly determined within each block. When any specific face appeared, the action associated with that face was randomly determined. Actions were shown in a random order below an array consisting of all 12 faces. This structure of randomization allowed us to control for three potential biases: (a) order effects; (b) bias because of a specific face; (c) bias because of a face-action pairing.

After completing the memory task, participants were asked their sex, age, and political beliefs. All measures and conditions are reported here.

**Results.** All participants were included in the data analysis. We hypothesized an interaction between the facial processing manipulations and the valence of the action; therefore, we conducted a 2 (orientation: upright, inverted) × 3(valence: positive, neutral, and negative) repeated measure analysis of variance (ANOVA), using the average response accuracy for each of the six act-face combinations. Results replicated the face inversion effect—participants were more accurate at identifying upright ($M$ = .61, $SD$ = .22) than inverted faces ($M$ = .48, $SD$ = .27) $F(1, 43) = 33.4$, $p < .001$, $\eta_p^2 = .41$.

The hypothesized interaction between the orientation of the face and the valence of the action also emerged, $F(2, 43) = 4.09$, $p < .05$, $\eta_p^2 = .15$. Participants were more accurate at identifying inverted faces paired with negative actions ($M$ = .54, $SD$ = .23) than with positive actions ($M$ = .45, $SD$ = .20), $t(46) = 4.4$, $p < .001$, $d' = .41$ or neutral actions ($M$ = .45, $SD$ = .21), $t(46) = 2.9$, $p < .002$, $d' = .40$. More important, they were *not* more accurate at identifying upright faces paired with negative actions ($M$ = .59, $SD$ = .22) than with positive actions ($M$ = .61, $SD$ = .23), $t(46) = .2$, $p = .42$ or neutral actions ($M$ = .59, $SD$ = .23), $t(46) = 1.58$, $p = .94$. This pattern of results shows a reduced inversion effect for faces linked to negative actions.

**Discussion.** The reduction in the facial-inversion effect, as shown by the interaction between valence and orientation, suggests that participants were no longer processing the faces of perpetrators of negative actions as they process other faces. Although these results indicate a change in processing, we are not claiming to have determined the specific perceptual mechanism underlying the changes in processing. The specific mechanisms underlying the facial inversion effect remain a topic of empirical debate (Sekuler, Gaspar, Gold, & Bennett, 2004; Young, Hellawell, & Hay, 1987). However, Study 1 suggests the faces of norm violators are perceived "differently" from those of others, and from a social psychological perspective, it is the very atypicality that matters most, not the exact character of the perceptual mechanism.

## Study 2

While Study 1 used an independent variable that impairs face-typical processing, Study 2 uses an independent variable that impairs analytic ("object-typical") processing. As a result, identifying the altered faces in Study 2 requires a greater reliance on face-typical processing. Therefore, Study 2 allows us to test the counterhypothesis that the improved accuracy scores in Study 1 were simply driven by increased effort or by measurement-scale artifacts.

Featural information about specific aspects of faces (e.g., shape of chin or nose) is finer grained than holistic or configural information. Therefore, removing high frequency spatial information disproportionally degrades featural processing (Goffaux, Hault, Michel, Vuong, & Rossion, 2005) relative to face-typical processing (e.g., configural). If individuals have difficulty processing the faces of norm violators in face-typical ways, this manipulation should reduce their ability to identify faces paired with norm violations (in contrast to the improvement in Study 1). We predict that removing high-spatial-frequency information should reduce the identifiability of the faces of severe norm violators relative to others.

**Method.**

***Participants.*** Participants were 208 (88 men and 120 women) students at a northeastern university who participated in exchange for payment. Data was collected over a 1-week period in the Wharton Behavioral lab. Sample size was determined by the number of participants who signed up to participate during the five day interval the study was scheduled to run. The average age of participants was 20.5 ($SD$ = .9) and participants reported a liberal ideological leaning ($M$ = 5.0 on a 7-point scale).

***Stimuli.*** Forty male faces with neutral expression were taken from the KDEF database (Lundqvist, Flykt, & Öhman, 1998). The faces were Fourier transformed and multiplied by high-pass Gaussian filters that preserved low (<8 cycles/face width) spatial frequencies (see Figure 2). Full spectrum (FS) faces were also used.

***Procedure.*** We used a recognition memory task to measure perceptual dehumanization. The experiment consisted of four blocks, each consisting of 10 face-action associations. In two blocks, the faces displayed were full spatial frequency, and in two blocks the faces were low spatial frequency (i.e., blurry). In each block participants completed a learning phase, followed by visual distraction and then a testing phase. In the learning phase, participants were told that they would learn of actions that the individuals depicted in the photographs had performed, and that their task was to learn to identify each individual by their past behavior. Participants learned 10 face-action pairings by simultaneously viewing a single face (either normal resolution or low spatial frequency) and a single action on screen for 12 s.

During the test phase, all 10 faces from the learning phase of the block were displayed in a single array that remained on screen. Actions appeared serially below the face array, and the participant identified the face originally linked to the action in the learning phase from the array. Face arrays were created by randomly assigning each of the 10 faces a number. Actions systematically varied in moral valence as in Study1. Face and action randomization used the same method as Study 1. After completing the memory task, participants were asked their sex, age, and political beliefs. All measures and conditions are reported here.

**Results.** All participants were included in analysis. We hypothesized an interaction between spatial frequency and the valence of action associated with the face; therefore, we conducted a 2 (spatial frequency: low frequency, full frequency) $\times$ 3(valence: positive, neutral, and negative) within-in subject ANOVA using the average response accuracy for each of the six act-face combinations. Not surprisingly, given the reduction in identifying information, participants were more accurate at identifying face-action pairings when faces were displayed at full spatial frequency ($M$ = .74, $SD$ = .25) than at low spatial frequency faces ($M$ = .66, $SD$ = .27), $F(1, 205) = 58.2$, $p < .001$, $\eta_p^2 = .48$. Most relevant, the hypothesized interaction between spatial frequency and valence emerged, $F(2, 204) = 38.8$, $p < .001$, $\eta_p^2 = .19$. As predicted, participants were significantly less accurate at identifying low spatial frequency faces paired with negative actions ($M$ = .60, $SD$ = .28) than low spatial frequency faces paired with positive actions ($M$ = .69, $SD$ = .25), $t(207) = 2.79$, $p = .002$, $d' = .34$ or neutral actions ($M$ = .68, $SD$ = .27) $t(207) = 2.05$, $p = .007$, $d' = .29$. Alternatively, participants were no more accurate at identifying full spatial frequency faces paired with negative actions ($M$ = .74, $SD$ = .27) than low spatial frequency faces paired with positive actions ($M$ = .70, $SD$ = .25), $t(207) = .07$, $p = .47$ or neutral actions ($M$ = .72, $SD$ = .26), $t(207) = .71$, $p = .76$.

**Discussion.** Consistent with the theory of perceptual dehumanization, participants processed the faces of norm violators differently from how they processed other faces. As predicted, participants were more impaired when low frequency faces were associated with negative actions than with neutral or positive actions. As shown in Figure 3, the results follow the opposite pattern as the results of Study 1 and thus rule out many artifactual explanations.

## Study 3

Studies 1 and 2 establish that the faces of perpetrators are processed differently. However, they do so by manipulating the valence of the action; this allows many features of the actions to vary, any of which could trigger the change in processing. Studies 3, 4, and 5 test the hypothesized function of perceptual dehumanization by linking changes in processing to punitive motives more specifically.

To establish that the desire to punish the perpetrator triggers perceptual dehumanization, the valence of actions needs to be controlled to rule out the possibility of a general negativity effect on face processing. Study 3 excludes this potential confound. To control for the negativity of the action but not the desire to punish, we use the same actions and vary whether the face is associated with the victim or the perpetrator of the harm. As a consequence, any change in the processing of faces across conditions should be because of differences in punitive motives, not the valence of the action.

Study 3 also explores the perceptual mechanism behind the changes observed in Studies 1 and 2. It uses the Face Composite Task (FCT) that more clearly implicates a specific perceptual impairment: configural processing. In the FCT, participants are asked to identify if the top halves of two faces are the same or different; however, the bottom halves associated with the sample (the face participants learn) and target (the face participants match) trials are always different. More important, when the top half is combined with a different bottom half, it creates the impression of a novel face. The composite image is a face that does not resemble the original person depicted in the top or bottom portion of the image; therefore, it is difficult to determine if the top-half is the same or different. This effect is produced by configural processing because it reduces participants' ability to selectively attend to the top half of the face when the composite face is presented as a single face. However, when the top and bottom half are misaligned, configural processing is disrupted and participants can

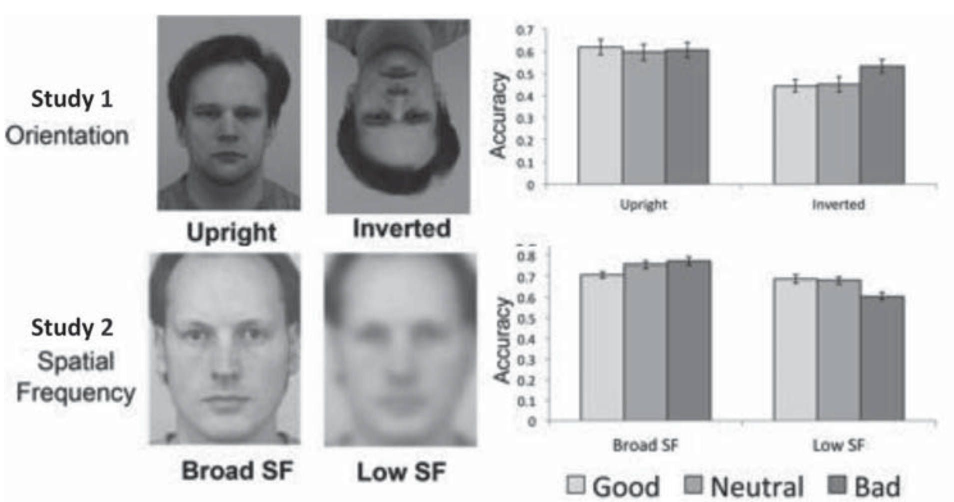


*Figure 3.* The results of Studies 1 and 2 suggest faces of norm violators are processed atypically. Study 1 suggests that pairing faces with negative social information attenuates the effects of inversion. Study 2 demonstrates that pairing faces with negative acts magnifies the effect of high frequency spatial filters.

more easily recognize the top half of the face. Therefore, it is easier to identify if the top halves are the same when the two halves are misaligned than when the two halves are fitted smoothly together (Young et al., 1987). The difference in performance across these two conditions indicates the effect of configural processing.

The theory of perceptual dehumanization predicts a reduction in the effect of offsetting faces when faces are linked to the perpetrators but not the victims of harm. A reduction in the effect of facial offsetting would suggest that valence does not drive perceptual dehumanization and that part of the disruption in processing is because of a failure to configurally processes the faces of perpetrators.

**Method.**

***Participants.*** Participants were 53 (24 men and 29 women) students at a northeastern university who participated in exchange for partial fulfillment of a course requirement. Data was collected over a 1-week period at the end of the semester. Sample size was determined by signup.

***Stimuli.*** Thirty-two White male face stimuli were taken from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) and used to form composites faces. All faces were fitted onto a $256 \times 256$ pixel white background and converted to greyscale. We created composite faces by pairing the top half of a face with the bottom half of another face. Luminance of the bottom half of each face was adjusted to match the top half. To ensure same and different trials were identical, only composite faces were used. We created offset faces by offsetting the bottom half of the face by 38 pixels. Each face appeared once as intact and once as offset.

***Procedure.*** To provide social information about each target face, we created a modified FCT. As in the standard facial composite task, subjects were instructed to attend only the top part of the face and performed a perceptual matching task. The experiment was comprised of 32 experimental trials. Each trial included reading a description of a harmful action and a matching task for both the victim and the perpetrator of the action.

As Figure 4 shows, in our task, each trial began with a verbal description of a crime on screen for 8500ms. In the description of the crime, both the victim and perpetrator of the action were identified by name. An example action is: “Dylan shot Jeff during a robbery.” After reading the description, participants performed a separate perceptual matching task for each of the actors as identified by name (e.g., “Dylan”) involved in the crime. Except for the displayed name, the perceptual matching components of each trial were standard. The matching task was performed separately for each of the targets.

In each matching trial, participants judged whether the top half of the target and sample face were the same or different. Specifically, after a 400 ms fixation, the target face and actor name (randomly determined to be the victim or the perpetrator) appeared and remained on screen for 600 ms. After a 1,000 ms interstimulus interval, the sample face appeared for 800 ms. Participants had 750 ms to decide, as quickly and accurately as possible, whether the top and bottom halves were the same or different. After a 400 ms

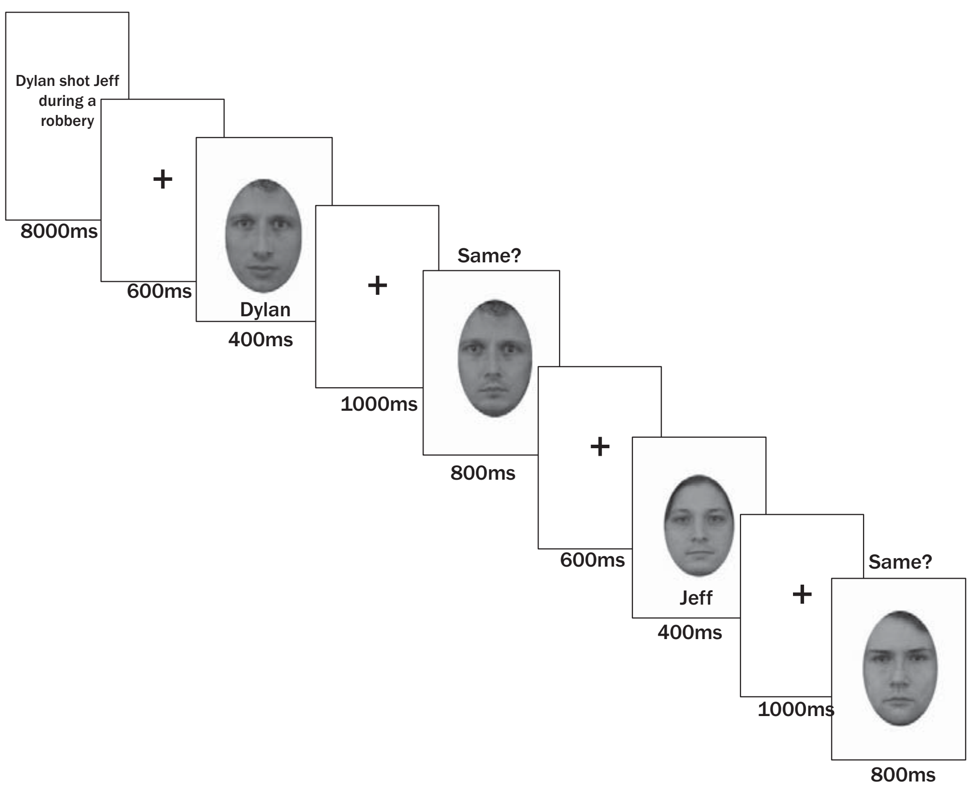


*Figure 4.* The procedure used in Study 3. The order of the victim and perpetrator of the action were randomized across participants. Participants indicated whether the top half of the faces matched by pressing a key.

interstimulus interval, a second target face and name (if the first name belonged to victim, the perpetrator's name was used, or vis-versa) appeared and remained on screen for 600 ms. After a 1,000 ms interstimulus interval, the sample face appeared for 800 ms. Again, subjects had 750 ms to decide whether these were the same or different. The order of the perpetrator and victim was counterbalanced.

The target and sample faces appeared at slightly different screen locations, to avoid subjects comparing a specific location of the display to perform the matching task. The participants expressed their choice by pressing a key on the left side ("f") or right side ("j") of the keyboard. Stickers with "S" for same or "D" for different were placed on the respective keys. In half of the trials, the top parts were identical for the sample and target face (demanding a "same" response). In the other half, the top differed for the sample and target face (demanding a "different" response). The bottom halves of the face always differed; however, this was irrelevant to the participants' task. Same-aligned and same-misaligned faces appeared twice as target faces: once in a same trial, once in a different trial; therefore, each of the 32 target top parts appeared approximately once in each condition.

**Results.** We hypothesized an interaction between alignment and the actor associated with the face; therefore we conducted a 2 (alignment: intact, offset) × 2 (actor: victim, perpetrator) repeated measure ANOVA using the sensitivity index for each of the four act-face combinations. A sensitivity index provides the separation between the means of the signal and the noise distributions, compared against the standard deviation of the noise distribution. We used this measure to control for threshold effects.

Results of Study 3 were largely consistent with those of Studies 1 and 2. Replicating past work on configural processing, participants showed greater sensitivity at identifying the top half of offset ($M$ = .61, $SD$ = 1.1) than intact faces ($M$ = .26, $SD$ = 1.3), $F(1, 53) = 32.5$, $p < .001$, $\eta_p^2 = .39$. Participants had insignificantly greater sensitivity at recognizing victims ($M$ = .51, $SD$ = 1.2) than perpetrators ($M$ = .37, $SD$ = 1.2), $F(1, 53) = 3.76$, $p = .058$, $\eta_p^2 = .07$.

Most important, there was a significant interaction between action type and alignment $F(2, 53) = 6.80$, $p = .01$, $\eta_p^2 = .12$. For victims, participants showed a typical pattern of face perception, with a greater sensitivity to the top half of offset faces ($M$ = .80, $SD$ = 1.1) than intact faces ($M$ = .43, $SD$ = 1.2), $t(52) = 2.78$, $p = .004$, $d' = .35$. However, for perpetrators, participants showed an abnormal pattern of face processing: they showed no more sensitivity to the top half of offset faces ($M$ = .22, $SD$ = 1.3) than intact faces ($M$ = .30, $SD$ = 1.3), $t(52) = 2.53$, $p = .99$, $d' = .07$.

**Discussion.** We observed a reduction in the effect of facial offsetting for the faces of perpetrators of harmful actions relative to the faces of victims. Study 3 helps to resolve three interpretive problems with the previous studies. First, the study controls for the effect of valence. Because the action is the same across conditions, the negativity of the action cannot drive the difference between conditions. Second, by presenting the information about the actor before displaying the face, the study controls for the time spent looking at faces versus actions. Third, the study uses a perceptual matching task rather than recall recognition; therefore, differences across conditions cannot be driven by differences in recall.

Moreover, Study 3 begins to identify the perceptual mechanisms underlying our effect. The facial composite task directly manipulates configural processing. Therefore, changes in processing in Study 3 can be more directly attributed to changes in configural processing. However, it should be noted that this study used a partial design. A partial design task is unbalanced because the top half of the face is always different. Therefore, there is still a potential problem with response bias (Gauthier & Bukach, 2007; Richler, Cheung, & Gauthier, 2011). Here it is worth noting an analytical difference between Studies 1 and 2 and Study 3. In Studies 1 and 2, participants selected the face they believed to be associated with a particular action. Participants cannot overidentify perpetrators because a response bias would require a threshold shift toward particular faces. By contrast, Study 3 is a perceptual matching task. In this case, if participants were biased to over-identify perpetrators they could simply respond that the faces match more frequently and, thus, response bias is a potential concern. To avoid this confound, we use the sensitivity index ($d'$) rather than accuracy.

Of interest to the authors, Study 3 showed that perceptual dehumanization effects were mostly driven by greater sensitivity to offset faces (or faces where processing was unimpaired), rather than to faces where processing was impaired as in Studies 1 and 2. Although our theory requires smaller effects of facial offsetting for perpetrators than for victims of harm, it does not make predictions about overall accuracy across conditions. Based on prior research, which suggests that negative information disproportionally captures attention, one would predict that Studies 1 and 2 (that contrast positive and negative information) may show the effect via opposite pathways from Studies 3, 4, and 5 (that contrast negative and negative information). This implies an overall effect of negativity on attention (and performance) in Studies 1 and 2 that is not present in Studies 3, 4 and 5 (because we controlled for it). However, this phenomenon is not relevant to the theory of perceptual dehumanization, which predicts differential attenuations of perceptual effects for which the theoretically relevant effect is the interaction.

## Study 4

Our core thesis is that attenuations in face-typical processing facilitate punishment. But the attenuation of face typical processing should only occur for actors whom people desire to punish. Although Study 3 demonstrates that perceptual dehumanization is not driven by valence, there are other possible confounds. Studies 4 and 5 further test the link between perceptual dehumanization and punishment using the culpable control model.

The phrase "culpable control" refers to the fact that the desire to blame or find someone culpable requires assessments of mental states and agentic control over outcomes ("mens rea" in the law). The culpable control model specifies the conditions under which individuals ascribe blame, such as whether situational pressures (e.g., coercion, provocation) or personal incapacities (e.g., ignorance, mental illness) are sufficient to excuse wrongdoing (Alicke, 2000; Alicke, Rose, & Bloom, 2011). The model postulates that one pathway through which culpability can be reduced is the link between behavior and its consequences—or volitional control.

Study 4 directly manipulates volitional control by varying the intention of the action. It measures changes in processing using the FCT. The theory of perceptual dehumanization, therefore, predicts a reduction in face-typical processing only when faces are linked to intentionally harmful actions. Therefore, we predict a reduction in the effect of facial offsetting for faces associated with intentional rather than unintentional harm.

**Method.**

***Participants.*** Participants were 204 (98 men and 106 women) students at a northeastern university who participated in exchange for payment. Data was collected over a 1-week period in the Wharton Behavioral lab. Sample size was determined by signup. The average age of participants was 20.5 (*SD* .9).

***Stimuli.*** Sixteen White male face stimuli were taken from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) and were used to form face composites. All faces were fitted onto a 256 × 256 pixel white background and converted to greyscale. Luminance of the bottom half of each face was adjusted to match the top half. Composite faces were used for both the sample and target face. We created offset faces by offsetting the bottom half of the face by 38 pixels. Each face appeared once as intact and once as offset.

***Procedure.*** Subjects performed a perceptual matching task in which they indicated if the top halves of two faces matched. Before each trial of the perceptual matching task, participants learned about an intentional or unintentional harm the target perpetrated.

All that varied was the use of the word intentional versus unintentional in descriptions of the act. For example, in the intentional condition the action read "Dylan intentionally shot his brother while hunting" while in the unintentional harm condition the action read, "Dylan unintentionally shot his brother while hunting." Which actions were associated with intentional and unintentional harms were counterbalanced across participants.

After reading the action, participants performed a perceptual recognition task. To ensure participants associated the face with the action, the name of the actor appeared below the face (e.g., "Dylan"). Each trial began with the display of an action on screen for 8,000 ms. After a 500 ms interstimulus interval, a face appeared on screen for 400 ms. After a 1,000 ms interstimulus interval, the sample face appeared for 800 ms. Participants had 750 ms to decide, as fast and accurately as possible, whether the top and bottom halves were the same or different.

To avoid subjects comparing a specific location of the display to perform the matching task, the target and sample faces appeared at slightly different screen locations. In half of the trials both the top and bottom half of the target and sample face were aligned. In half of the trials both the top and bottom half of the target and sample face were misaligned. As in Experiment 3, the participants expressed their choice by pressing a left versus right key on a keyboard placed in front of them. Same-aligned and same-misaligned faces appeared twice as target faces: once in a same trial, and once in a different trial. Target and sample faces always differed with regard to their bottom half. In half of the trials, the top parts were identical (same). In the other half, both top and bottom parts differed (different). Intentional and unintentional actions, as well as aligned and misaligned trials were randomly interleaved. The experiment comprised 32 experimental trials randomly mixed up across subjects.

**Results.** We hypothesized an interaction between alignment and the actor's intention; therefore, we conducted a 2 (alignment: intact, offset) × 2 (action: intentional, accidental) repeated measure ANOVA using the sensitivity index for each of the four act-face combinations. Results of Study 4 were largely consistent with those of Study 3.

Results suggest participants had greater sensitivity to recognizing the top half of offset faces ($M$ = .91, $SD$ = 1.3) than intact faces ($M$ = .46, $SD$ = 1.1) $F(1, 200) = 135.2$, $p < .001$, $\eta_p^2 = 1.14$. Participants displayed insignificantly greater sensitivity at identifying faces paired with unintentional ($M$ = .72, $SD$ = .9) than intentional ($M$ = .66, $SD$ = 1.0) harms $F(1, 200) = 2.076$, $p = .15$. As predicted, there was a significant interaction between action type and alignment, $F(2, 199) = 28.2$, $p < .001$, $\eta_p^2 = .28$. When the action was unintentional, participants showed a typical pattern of face perception: they showed greater sensitivity to offset faces ($M$ = 1.42, $SD$ = 1.3) than intact faces ($M$ = .34, $SD$ = 1.1), $t(52) = 5.43$, $p < .001$, $d' = .86$. When the action was intentional, participants showed an atypical pattern of face perception: they did not show a different sensitivity to offset ($M$ = .66, $SD$ = 1.3) faces than intact ($M$ = .44, $SD$ = 1.3) faces, $t(52) = .48$, $p = .83$, $d' =$ .17.

**Discussion.** In Study 4, we observe a reduction in the effect of facial offset for the faces actors associated with intentional, but not unintentional harm. This suggests that intentional harm leads to changes in visual processing. Study 4 strengthens the claim that perceptual dehumanization facilitates punitive action by linking perceptual dehumanaztion to the desire to punish. Specifically Study 4 suggests that culpability, a necessary precondition for punishment, is a necessary condition for perceptual dehumanization as well.

As in Study 3, the effect was primarily driven by different sensitivity to offset faces (where processing was unimpaired), rather than faces where processing was impaired. Again, as previously discussed, this change is likely because of negativity dominance and difference in attention across conditions across studies. What remains important is that all studies have shown greater attenuation of face-typical processing for harm-doers—and Study 4 underscores that it needs to be intentional harmdoing.

## Study 5

Study 5 provides convergent validity for the link between blame and perceptual dehumanization. While Study 4 manipulated culpable control by manipulating whether the action was performed accidentally or purposefully, Study 5 manipulates capacity constraints.

"Capacity" refers to the actor's mental state or ability to have acted otherwise. Past work on capacity constraints suggests that the same action can be interpreted differently depending on beliefs about the actor's ability to have acted differently (Alicke, 2000). In Study 5 we manipulate culpability by altering the degree to which the perpetrators' actions were genetically predetermined. In a between-subjects fully counterbalanced design, we manipulated participants' beliefs about the perpetrator's psychological capacity to have acted otherwise. We predicted a reduction in face-typical processing when faces are linked to actors who could have acted otherwise, but not when their actions were believed to be genetically predetermined.

**Method.**

***Participants.*** Participants were 126 (47 men and 79 women) students at a northeastern university who participated in exchange for course credit. Data was collected over a 2-week period. The average age of participation was 19.7 (*SD* 1.3) and the sample was liberal (mean 4.7 on a 7-point scale).

***Stimuli.*** Thirty-two face stimuli were taken from the KDEF database (Lundqvist, Flykt, & Öhman, 1998). Images were cropped to 250 × 350 pixels and converted to grayscale.

***Procedure.*** Before completing a recognition recall task, participants read an article documenting the contribution of genetic and environmental factors to pedophiliac behavior relative to impulsive violent behavior. In one condition, pedophiliac behavior was described as highly predetermined, and impulsive violent behavior was described as highly controllable. In the other condition impulsive violent behavior was described as highly predetermined, and pedophiliac behavior was described as highly controllable. Following the passage, participants completed a reading comprehension check. If they failed they were reshown the relevant information from the passage until they passed.

After completing the reading comprehension, participants completed the recognition recall task to measure perceptual dehumanization. In two blocks, faces were intact, and in two blocks faces were offset. In each block, participants completed a learning phase and then a testing phase. In the learning phase, participants learned eight face-name pairings by simultaneously seeing a single face (either intact or offset) and a single name on a screen for 8.5 s. Preceding the display of the face, participants learned whether the target had committed a pedophiliac or violent action through a screen that displayed "This target has committed a pedophiliac action" or "this target has committed an extremely violent action." Half of the individuals were described as pedophiliac whereas the other half were described as violent. Action type and names were randomly paired with faces.

During the test phase, all 8 faces from the learning phase of the block were displayed in a single arrangement that remained on screen. Names appeared serially below the face array, and the participant identified the face originally linked to the name in the learning phase. Display order in both the test and recall phase and face-name pairings were fully randomized.

**Results.** We conducted a 2 (behavior: pedophiliac, violent) × 2 (Offset: intact, Offset) × 2 (condition: pedophilia controllable, violence controllable) mixed-model ANOVA using the average response accuracy for each of the four behavior-face combinations. Results indicate an effect of facial offsetting; participants were more accurate at identifying intact ($M$ = .64, $SD$ = .22) than offset ($M$ = .52, $SD$ = .25) faces $F(1, 124) = 129.0$, $p < .001$, $\eta_p^2 = .39$. There was also a main effect of behavior type; participants were more accurate at identifying pedophiles ($M$ = .64, $SD$ = .21) than violent individuals ($M$ = .53, $SD$ = .26), $F(1, 124) = 33.598$, $p <$ .001, $\eta_p^2 = .22$.

There was an unpredicted interaction between behavior type and condition, $F(1, 124) = 4.71$, $p = .03$, $\eta_p^2 = .06$. Participants were more accurate at identifying pedophiles when they believed the pedophiles were responsible for their actions ($M$ = .68, $SD$ = .21) than when they were told pedophiles were not responsible for their actions, ($M$ = .60 $SD$ = .21), $t(128) = 1.94$, $p = .028$, $d' = .25$.

The hypothesized higher-order interaction also emerged among culpability prime, alignment and action $F(3, 205) = 7.3$, $p = .007$, $\eta_p^2 = .12$. Reductions in facial offsetting were attenuated by perceived psychological capacity constraints. When participants believed pedophilia was because of a psychological capacity constraint, participants showed greater accuracy at identifying intact ($M$ = 74, $SD$ = .02) than offset faces of pedophiles ($M$ = .55, $SD$ = .02), $t(52) = 3.76$, $p < .001$, $d' = .59$; however, when pedophilia was viewed as controllable, participants showed no greater accuracy at identifying intact ($M$ = .69, $SD$ = .02) than offset faces of pedophiles ($M$ = .63, $SD$ = .02) faces $t(52) = .525$, $p = .30$. Conversely, when participants believed violent behavior was because of a capacity constraint, they showed greater accuracy at identifying intact ($M$ = 55 =, $SD$ = .02) than offset faces of violent actors ($M$ = .39, $SD$ = .02), $t(52) = 2.8$, $p = .002$, $d' =$ .29; however, when the violent behavior was viewed as volitional, participants showed no greater accuracy at identifying intact ($M$ = .56, $SD$ = .02) than offset faces of violent actors ($M$ = .52, $SD$ = .02), $t(52) = 5.01$, $p < .001$, $d' = .67$.

**Discussion.** Study 5 found a reduction in the effect of facial offsetting only for actors who could control their actions. This suggests that capacity constraints moderate changes in visual processing, that is, perceptual dehumanization. Study 5 strengthens the claim that perceptual dehumanization facilitates punitive action by further establishing the link between perceptual dehumanization and the desire to punish.

Note that Studies 3 and 4 used a paradigm, which captured the effects of selective attention on configural processing. Therefore, participants perform best when configural processing is inhibited. By contrast, in Study 5 participants are responding the entire face. Therefore, they perform base when engaging in face typical processing. Thus, disruptions in face typical process should reduce accuracy. This suggests that the effects in these studies should run in opposite directions. And they did: the main effects of offsetting are consistently in opposite directions. The simple effects in the studies reflect these patterns.

Moreover, because the actions are held constant across conditions, the effects cannot be because of either differences in valence, the action or the interpretation of the actions. These results strengthen the claim that culpability, a necessary precondition for punishment, is a necessary condition for perceptual dehumanization as well.

## Series 2

Section 1 established a link between social cues and perceptual processes. The results from Studies 1 to 5 suggest that changes in punitive desire lead to changes in visual processing of faces. Section 2 establishes a functional relationship between changes in perceptual processes and changes in behavior by testing the hypothesis that reductions in face-typical processing facilitate punitive actions.

Studies 6 and 7 directly link perceptual dehumanization to an increased willingness to punish by testing the impact of face processing on punishment decisions. The independent variables, are the face processing manipulations used in Study 1 (Inversion) and Study 2 (Low Spatial Frequency). The dependent variable is a logarithmic sentencing scale (Darley, Carlsmith, & Robinson, 2001) used to measure punitive drive in social psychological research.

Our theory predicts that inversion, which decreases face-typical processing, should increase punishment. Conversely, low spatial frequency, which forces participants to engage in face-typical processing, should decrease punishment. Because these studies change face processing in opposite directions, when we take them together, we can establish that changes in punitive drives are because of changes in visual processing, not alternative explanations.

## Study 6

**Method.**

***Participants.*** Participants were 225 (125 men and 100 women) American workers on Mechanical Turk who participated for $1.00 payment. Based upon funding constraints the sample size was predetermined to be 225.

***Stimuli.*** Thirty face stimuli were taken from Ballew and Todorov (2007). Our images were restricted to while male gubernatorial runner-ups. Images were cropped to 150 × 215 pixels, placed on a standard background, and converted to grayscale.

***Procedure.*** Participants rated the appropriate punishment for 15 perpetrators with upright faces and 15 perpetrators with inverted faces on a unique 13-point scale adopted from the justice-research literature (Carlsmith, Darley, & Robinson, 2002; Darley et al., 2001). The scale ranged from "No sentence" to "Death sentence" with nonlinear intervals given in weeks, months, and years of incarceration. The actions were taken from Studies 1 and 2. Therefore, the actions were preselected for sentences longer than 5 years and used language from the federal sentencing guidelines. Examples include "Trafficking of $100,000 of drugs" and "Sexual Exploitation of Minors for $50,000 profit." The orientation of the faces was counterbalanced across participants.

At the beginning of the task, the participant received a message from the experimenter apologizing for a programing glitch and explaining that some of the faces may be displayed as upside down. To avoid spurious face-action effects, the face-action pairings were randomly determined for each participant. Additionally, the orientation of faces was randomly determined for each participant.

**Results.** We conducted a 2 (orientation: upright, inverted) × 2 (Block Order: Upright first, Inverted First) mixed model ANOVA on punitiveness. There was a significant effect of inversion on punitiveness, with inverted faces ($M$ = 8.1, $SD$ = .35) receiving significantly harsher punishments than upright faces ($M$ = 7.8, $SD$ = .37), $F(1, 222) = 64.8$, $p < .001$, $\eta_p^2 = .06$. This suggests inhibition of face typical processing facilitates punishment.

## Study 7

**Method.**

***Participants.*** Participants were 225 (117 men and 108 women) American Mechanical Turk workers who participated for $1.00 payment. Based upon funding constraints the sample size was predetermined to be 225.

***Stimuli.*** Thirty male faces with neutral expressions were taken from the KDEF database (Lundqvist, Flykt, & Öhman, 1998). The faces were Fourier transformed and multiplied by high-pass Gaussian filters that preserved low (<8 cycles/face width) spatial frequencies (see Figure 3). Full spectrum (FS) faces were also used.

***Procedure.*** Participants rated the appropriate punishment for 15 perpetrators with full spectrum faces and 15 perpetrators with low frequency faces on the same 13-point scale adopted from the justice-research literature as Study 6 (Carlsmith et al., 2002; Darley et al., 2001). At the beginning of the task, the participant received a message from the experimenter apologizing for the programing glitch, explaining that some of the faces may appear as low resolution depending upon the participants browser and monitor.

Thirty norm violations were randomly selected from Studies 1 and 2. Therefore, the actions were preselected for sentences longer than 5 years and used language from the federal sentencing guidelines. Examples actions are "Blackmailed a Congressman for $100,000" and "Committed first degree murder." The spatial frequency of the faces was counterbalanced across participants. To avoid spurious face-action effects, the action-face pairings were randomly determined for each participant.

**Results.** A 2 (Spatial frequency: Low only, Full) × 2 (Block Order: Upright first, Inverted First) repeated measure ANOVA revealed a significant effect of blurring on punitiveness, with low frequency faces ($M$ = 8.6, $SD$ = .32) receiving significantly more lenient punishments than normal faces ($M$ = 9.2, $SD$ = .29), $F(1, 222) = 53.2$, $p < .001$, $\eta_p^2 = .12$.

**Discussion.** Studies 6 and 7 directly link face-typical processing to willingness to punish. Study 6 impairs typical face processing and demonstrates that impairments in face processing increase punitiveness. Study 7 impairs featural processing and demonstrates that increased reliance on face-typical processing decreases punitiveness. As Figure 5 shows, inverted faces and low spatial frequency faces have opposite effects on punitiveness. More important, although both Studies 6 and 7 distort faces, in line with our predictions they affect punitiveness in opposing directions. Taken together, Studies 6 and 7 suggest an inhibition of face-typical processing functions to increase the punitive drive.

# General Discussion

Seven studies yielded substantial support for the key predictions of the theory of perceptual dehumanization. The results suggest that participants visually processed the faces of norm violators atypically and these changes in visual perception facilitated punitive behavior. Studies 1–5 use a multimethod array of techniques to produce convergent evidence that the visual mechanisms that enable face-typical processing are inhibited upon learning someone is a norm violator.

Studies 3 through 5 underscore that that these attenuations of face-typical processing are specific to norm violators. Study 3 controls for valence as a potential confound by using the same actions across conditions and only varying the role faces are assigned (victim or perpetrator). Study 4 and 5 more stringently test the functionalist logic of perceptual dehumanization by applying Alicke's model (2000) of culpable control and manipulating the degree to which observers see perpetrators as blameworthy. The results of these studies are consistent with the functionalist hypothesis that perceptual dehumanization effects are driven by the desire to punish.

Studies 6 and 7 establish a functionalist link between face processing and punishment. They directly manipulate face processing and show that face-typical processing reduces punitive drive. This establishes the link between visual processing and social behavior.

Returning to our core argument, to establish a functionalist relationship, face processing needs to interact with punitive be-

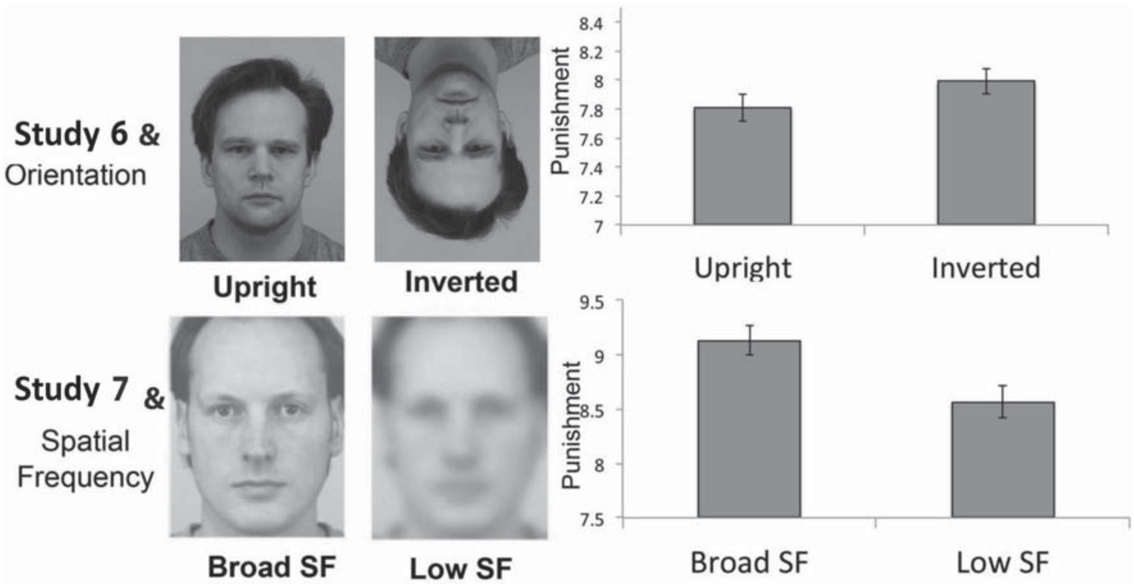


*Figure 5.* The results of Studies 6 and 7 directly link perceptual dehumanization to an increased willingness to punish. Study 6 demonstrates inversion increases punitiveness. Study 7 demonstrates that low frequency spatial reduce punitiveness.

havior in two ways: (a) social cues must induce changes in face processing; (b) changes in face processing must facilitate socially significant behavior. Series 1 established the link between social cues and face processing. Series 2 established the link between face processing and socially relevant behavior. The most parsimonious explanation is that changes in face-typical processing serve the social function of facilitating the punishment of norm violators. We posit this mechanism enables individuals to satisfy conflicting demands imposed by the complexities of social interdependence: balancing the need to get along and cooperate with the need to deter norm violations.

## Methodological Conclusions

A clear advantage of the perception literature is its emphasis on mechanism, which makes it easier than in much of social psychology to pinpoint the exact strengths as well as confounds in any particular paradigm. In Table 1, we outline the offsetting strengths and weaknesses of each paradigm used. Although each paradigm is potentially problematic, alternative artifactual explanations can be rendered implausible across these methodologically diverse studies. Weighing the entirety of the evidence, our results demonstrate impairment in the face typical processing of norm violators.

Readers familiar with the history of experimental psychology will recognize a resemblance between our findings and controversies dating back to the 1950s on the conditions under which social processes can influence perceptual processes (Bruner & Goodman, 1947). The controversies centered on two issues: (a) replicability and (b) the possibility that effects are because of response threshold shifts (or response biases), rather than perceptual shifts. We have shown replicability across different manipulations of perceptual processing (see Table 1), and across different types of facial and contextual parameters by manipulating display timing and

Table 1
*Convergent Operationalizations of Perceptual Dehumanization*

| Task (*Mechanism Impaired*) | Description | Negative task features | Confounds controlled for |
|---|---|---|---|
| Facial Inversion Effect (*face-typical processing*) | In a recognition memory, task facial inversion is used to impair face-typical processing | The mechanism underlying the impairment is unclear and the task uses recall. | Task structure controls for response bias |
| Low Frequency Resolution (*featural processing*) | In a recognition memory task removal of high spatial frequency information impairs featural processing | A recognition recall paradigm. | Effects are in the opposite direction, controlling for any attentional effects (controlling for attentional effects) |
| Facial Composite Task (*holistic processing*) | In a task facial offsetting is used to impair holistic processing | The "partial" design may be associated with a response bias. | Clearest manipulation of holistic processing; uses a perceptual matching task |

*Note.* Table 1 documents the strengths and weaknesses of each of the methods used to impair holistic processing.

display size. Additionally, we have taken a key step to eliminate response bias by designing a task in such a way that response bias cannot influence key variables. Response bias would be problematic if the independent variables (facial orientation and action valence) had been the response variable (and participants had been selecting an action to match a particular face). In Studies 1, 2, and 5, faces were randomized across actions and perceptual manipulations, and responses were uncorrelated with the main and interactive effects of the independent variables. A response bias would require a threshold shift toward particular types of faces. In these studies, faces were randomized across all other dimensions, rendering a systematic threshold shift a very unlikely counter-explanation.

Notably, we tested a potentially general theory in a specific domain. Many features of the punitive response make punishment well suited to test the theory of perceptual dehumanization. By studying punishment, researchers can isolate the effect of social information on face processing. Unlike researchers who study racial discrimination or ethnic genocide, those studying punitiveness can randomly pair facial stimuli to social categories. As a result, they are better positioned to show that perceptual effects are driven by the manipulation of social information, not by properties of the stimulus.

Second, in previous work on social information and face perception, attenuations in facial processing may be a generic response to negative associations (i.e., out-group derogation or enhanced positive associations with the in-group) and not relate to the *function* of (actively or passively) facilitating harm. More important, in the domain of punishment, the conditions under which individuals desire to inflict harm are well established. Prior work has decoupled punitive motives from generic negativity associated with particular actions. It is only when the perpetrator is seen as culpable (Alicke, 2000) or blameworthy (Malle, Monroe, Guglielmo, & 2014; Weiner, Graham, & Reyna, 1997) that people express retributive tendencies, (Tetlock et al., 2007).

Finally, it was possible to measure the effect of face processing on social behavior (punitive drives) directly. A large body of research has developed tools for measuring punitive attitudes (e.g., Carlsmith et al., 2002; McKee & Feather, 2008). Moreover, unlike work on racial attitudes, which finds considerable concealment or self-censorship of attitudes (Norton, Vandello, & Darley, 2004), harming norm violators is relatively socially acceptable (Tyler, 2006). Of course, perceptual dehumanization may occur more broadly; however, this remains a topic for future research.

## Theoretical Conclusions

It is also important to preempt confusion about the term "perceptual dehumanization," which is tempting to conflate with the much more extensively studied phenomenon of dehumanization. Several prominent theorists have identified the social construct of dehumanization as a mechanism that facilitates harming others. Opotow (1990) posited that dehumanization allows people to be placed outside the boundary beyond which moral rules apply. Bandura proposed that dehumanization is one way in which moral self-sanctions are selectively disengaged. And according to Kelman (1973), when people are divested of the agentic and communal aspects of humanity, they lose the capacity to evoke compassion and moral emotions.

Our research suggests perceptual dehumanization, like the historic construct, acts to facilitate harm. However, unlike the historic construct, perceptual dehumanization does not require endorsing negative stereotypes of the target group and likely functions with less awareness and on a more rapid time scale. Indeed, we propose that one mechanism for inhibiting cooperative impulses and facilitating punishment is rapid-fire shifts in the perceptual categorization of norm violators, in which the faces of norm violators cease to be seen as fully human.

Although we have shown that perceptual dehumanization facilitates punishment, it may function more broadly than just applying to norm violators. For example, prior research has demonstrated an attenuation in face-specific processing of individuals of other races (e.g., Michel et al., 2006) and out-group members (e.g., MacLin & Malpass, 2003). Table 2 lays out these potential parallels between the perceptual literature on attenuation of face typical processing and the social psychological literature on dehumanization and discrimination.

The fact that the attenuation of face-typical processing applies broadly, even to minimal (experimentally created) groups, suggests there may be an asymmetric relationship between perceptual dehumanization and more explicit forms of dehumanization. It may be possible to perceptually dehumanize without engaging in the explicit hatred or derogation of social-psychological dehumanization, but it is probably impossible to engage in nasty overt dehumanization without the subtle assistance of perceptual dehumanization. In short, our larger and more speculative claim is that perceptual dehumanization is a necessary but far from sufficient condition for dehumanization (as the term is used in the social psychological literature).

Table 2
*Convergent Evidence for Perceptual Dehumanization*

| Category | Perceptual effects | Dehumanization | Discrimination |
|---|---|---|---|
| Race | Michel, Rossion, Han, Chung, & Caldara, 2006; Rhodes, Hayward, & Winkler, 2006; Tanaka, Kiefer, & Bukach, 2004 | Jahoda, 1999; Cuddy, Rock, & Norton, 2007 | Ayres, 1991; Milkman, Akinola, & Chugh, 2012; Norton, Vandello, & Darley, 2004 |
| Group | Bernstein, Young, & Hugenberg, 2007; Hugenberg & Corneille, 2009 | Bar-Tal, 2000; Castano & Giner-Sorolla, 2006; Vaes, Paladino, Castelli, Leyens, & Giovanazzi, 2003 | Tajfel, 1970 |
| Gender | Bernard, Gervais, Allen, Campomizzi, & Klein 2012 | Nussbaum, 1999; Fredrickson & Roberts, 1997; | MacKinnon, 1987; Norton, Vandello, & Darley, 2004 |
| Dominate group | Shriver, Young, Hugenberg, Bernstein, & Lanter, 2008 | Hodson & Costello, 2007 | Sidanius & Pratto, 2001 |

*Note.* Table 2 documents the co-occurrence between the attenuation of face typical processing and the mistreatment of social groups.

Perceptual dehumanization may be so pervasive because it allows us to seamlessly navigate the emotional complexities of the social world. Human societies—unlike those of most other species—are based on large-scale cooperation among genetically unrelated individuals. As in other species, the social world of our ancestors contained individuals who were poised to exploit others if such acts were self-beneficial (Daly & Wilson, 1988; Duntley, 2005; Buss & Duntley, 2004). A sustainable social order requires beings that can switch readily between cooperation and punitiveness. Indeed, laboratory experiments and formal models alike suggest that the cooperative equilibria critical to human social groupings would be impossible without tit-for-tat rules and third-party punishment of norm violators (Axelrod & Hamilton, 1981; Fehr & Gächter, 2000; Nowak, 2006).

A larger issue lurking in the background are prevalent stereotypes about human nature, although alternative views in the social psychological literature do exist (Tetlock, 2002). It is generally believed that people (at least in affluent Western societies) have a default reaction of empathy in the face of others' pain. However, the ease with which our test subjects were able to "turn off" their empathy to do their social duty and punish norm violators and harm others more generally certainly calls the stereotype of empathy as a default part of human nature into question. This assumption is consistent with a good deal of work in developmental and social psychology (Hamlin, Wynn & Bloom, 2007; Tetlock et al., 2007).

Alternatively, one could make the opposite argument that empathy is required to overcome the natural tendency toward punitiveness and other forms of callousness. For example, one could posit that it makes greater survival sense for organisms to assume harmful actions or intentions and then to correct when this turns out to be untrue, than to begin with empathy and then need to overcome it. In this view, the humans' default is not to engage in face specific processing, and it is only when empathic processes are engaged that people do so. Past research on face processing has indeed suggested the possibility that "face-specific" processing is not the default processing of human faces (Van Bavel et al., 2011).

In summary, our results reveal previously experimentally unexplored connections across levels of analyses, between the more macro social phenomena of punishment to the more micro mechanisms of face perception. Our data also potentially bridge neuro-cognitive work on facial and object perception with social psychological work on norm enforcement and punitiveness. Perceptual dehumanization suggests that crucial adaptations to group life, such as our capacity to enforce norms essential for reaching and maintaining cooperative equilibria, are internalized at a surprisingly basic perceptual-cognitive level of mental functioning.

## References


Aberle, D. F., Cohen, A. K., Davis, A. K., Levy, M. J., Jr., & Sutton, F. X. (1950). The functional prerequisites of a society. *Ethics, 60,* 100–111. http://dx.doi.org/10.1086/290705

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126,* 556–574.

Alicke, M. D., Rose, D., & Bloom, D. (2011). Causation, norm violation, and culpable control. *The Journal of Philosophy, 108,* 670–696. http://dx.doi.org/10.5840/jphil20111081238

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia, 49,* 2541–2552. http://dx.doi.org/10.1016/j.neuropsychologia.2011.05.002

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211,* 1390–1396. http://dx.doi.org/10.1126/science.7466396

Ayres, I. (1991). Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review, 104,* 817–872.

Ballew, C. C., II, & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 17948–17953. http://dx.doi.org/10.1073/pnas.0705435104

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3,* 193–209. http://dx.doi.org/10.1207/s15327957pspr0303_3

Bar-Tal, D. (2000). *Shared beliefs in a society: Social psychological analysis*. Thousand Oaks, CA: Sage Publications.

Bernard, P., Gervais, S. J., Allen, J., Campomizzi, S., & Klein, O. (2012). Integrating sexual objectification with object versus person recognition: The sexualized-body-inversion hypothesis. *Psychological Science, 23,* 469–471.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18,* 706–712. http://dx.doi.org/10.1111/j.1467-9280.2007.01964.x

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123.

Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *The Journal of Abnormal and Social Psychology, 42,* 33–44. http://dx.doi.org/10.1037/h0058484

Buss, D. M., & Duntley, J. D. (2004). The evolution of gender differences in aggression. *Gender and Aggression*. New York, NY: Guilford Press.

Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology, 83,* 284–299. http://dx.doi.org/10.1037/0022-3514.83.2.284

Castano, E., & Giner-Sorolla, R. (2006). Not quite human: Infrahumanization in response to collective responsibility for intergroup killing. *Journal of Personality and Social Psychology, 90,* 804.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Colloquium paper: Adaptive specializations, social exchange, and the evolution of human intelligence. *Proceedings of the National Academy of Sciences of the United States of America, 107*(Suppl. 2), 9007–9014. http://dx.doi.org/10.1073/pnas.0914623107

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York, NY: Oxford University Press.

Cuddy, A. J., Rock, M. S., & Norton, M. I. (2007). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup helping. *Group Processes & Intergroup Relations, 10,* 107–118.

Daly, M., & Wilson, M. (1988). *Homicide*. New York, NY: Transaction Publishers.

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2001). The ex-ante function of the criminal law. *Law & Society Review, 35,* 165–190. http://dx.doi.org/10.2307/3185389

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107.

Duntley, J. D. (2005). Adaptations to dangers from humans. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 224–249). Hoboken, NJ: Wiley.

Durkheim, E. (2008). *The elementary forms of the religious life [1912]*. New York, NY: Oxford University Press.

Ellis, H. D., Bruce, V., & De Schonen, S. (1992). The development of face processing skills. [and Discussion]. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 335,* 105–111. http://dx.doi.org/10.1098/rstb.1992.0013

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia, 33,* 661–674. http://dx.doi.org/10.1016/0028-3932(95)00002-K

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628–634. http://dx.doi.org/10.1037/0096-1523.21.3.628

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35,* 2089–2093. http://dx.doi.org/10.1016/0042-6989(94)00273-O

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25,* 63–87. http://dx.doi.org/10.1016/S1090-5138(04)00005-4

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review, 90,* 980–994. http://dx.doi.org/10.1257/aer.90.4.980

Foucault, M. (1978). *Discipline and punish: The birth of the prison*. New York, NY: Pantheon Books.

Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory. *Psychology of Women Quarterly, 21,* 173–206.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103,* 322–330. http://dx.doi.org/10.1016/j.cognition.2006.05.003

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37,* 1673–1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2,* 568–573.

Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). Does visual subordinate-level categorisation engage the functionally defined fusiform face area? *Cognitive Neuropsychology, 17,* 143–164.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble' experts: A framework for studying expert object recognition processes. *Vision Research, 38,* 2401–2428.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception, 34,* 77–86. http://dx.doi.org/10.1068/p5370

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450,* 557–559. http://dx.doi.org/10.1038/nature06288

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10,* 252–264. http://dx.doi.org/10.1207/s15327957pspr1003_4

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293,* 2425–2430.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233. http://dx.doi.org/10.1016/S1364-6613(00)01482-0

Hodson, G., & Costello, K. (2007). Interpersonal disgust, ideological orientations, and dehumanization as predictors of intergroup attitudes. *Psychological Science, 18,* 691–698.

Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science, 33,* 1173–1181.

Hugenberg, K., Young, S., Rydell, R. J., Almaraz, S., Stanko, K. A., See, P. E., & Wilson, J. P. (2015). The face of humanity: Configural face processing influences ascriptions of humanness. *Social Psychological and Personality Science*. Advance online publication. http://dx.doi.org/10.1177/1948550615609734

Jahoda, G. (1999). *Images of savages: Ancients [sic] roots of modern prejudice in Western culture*. New York: Psychology Press.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience, 6,* 766–774. http://dx.doi.org/10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40,* 1–19. http://dx.doi.org/10.1016/0010-0277(91)90045-6

Kanwisher, N., Downing, P., Epstein, R., & Kourtzi, Z. (2001). Functional neuroimaging of human visual recognition. In R. Cabeza & A. Kingstone (Eds.), *The handbook on functional neuroimaging* (pp. 109–152). Cambridge, MA: MIT Press.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68,* B1–B11. http://dx.doi.org/10.1016/S0010-0277(98)00035-3

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18,* 1084–1089. http://dx.doi.org/10.1111/j.1467-9280.2007.02029.x

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues, 29,* 25–61. http://dx.doi.org/10.1111/j.1540-4560.1973.tb00102.x

Loewenstein, G., & Small, D. A. (2007). The Scarecrow and the Tin Man: The vicissitudes of human sympathy and caring. *Review of General Psychology, 11,* 112–126. http://dx.doi.org/10.1037/1089-2680.11.2.112

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet, 9,* 1–630.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47,* 1122–1135. http://dx.doi.org/10.3758/s13428-014-0532-5

MacKinnon, C. A. (1987). *Feminism unmodified: Discourses on life and law*. Cambridge, MA: Harvard University Press.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law, 7,* 98–118. http://dx.doi.org/10.1037/1076-8971.7.1.98

MacLin, O. H., & Malpass, R. S. (2003). The ambiguous-race face illusion. *Perception, 32,* 249–252. http://dx.doi.org/10.1068/p5046

Malle, B. F., Monroe, A. E., & Guglielmo, S. (2014). Paths to blame and paths to convergence. *Psychological Inquiry, 25,* 251–260. http://dx.doi.org/10.1080/1047840X.2014.913379

McKee, I. R., & Feather, N. T. (2008). Revenge, retribution, and values: Social attitudes and punitive sentencing. *Social Justice Research, 21,* 138–163. http://dx.doi.org/10.1007/s11211-008-0066-z

Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science, 17,* 608–615. http://dx.doi.org/10.1111/j.1467-9280.2006.01752.x

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science, 23,* 710–717.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., & Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science, 10,* 419–422. http://dx.doi.org/10.1111/1467-9280.00179

Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology, 87,* 817–831. http://dx.doi.org/10.1037/0022-3514.87.6.817

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314,* 1560–1563. http://dx.doi.org/10.1126/science.1133755

Nussbaum, M. C. (1999). *Sex and social justice*. New York: Oxford University Press.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues, 46,* 1–20. http://dx.doi.org/10.1111/j.1540-4560.1990.tb00268.x

Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. New York, NY: Penguin.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13,* 499–505.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22,* 464–471. http://dx.doi.org/10.1177/0956797611401753

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 296–320.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21,* 329–330. http://dx.doi.org/10.3758/BF03335807

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396. http://dx.doi.org/10.1016/j.cub.2004.02.028

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin, 34,* 260–274.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223,* 96–102.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition, 93,* B1–B9.

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review, 109,* 451–471. http://dx.doi.org/10.1037/0033-295X.109.3.451

Tetlock, P. E., Visser, P. S., Singh, R., Polifroni, M., Scott, A., Elson, S. B., & Rescober, P. (2007). People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility. *Journal of Experimental Social Psychology, 43,* 195–209. http://dx.doi.org/10.1016/j.jesp.2006.02.009

Treisman, A. M., & Kanwisher, N. G. (1998). Perceiving visually presented objects: Recognition, awareness, and modularity. *Current Opinion in Neurobiology, 8,* 218–226.

Tyler, T. R. (2006). *Why people obey the law*. Princeton, NJ: Princeton University Press.

Vaes, J., Paladino, M. P., Castelli, L., Leyens, J. P., & Giovanazzi, A. (2003). On the behavioral consequences of infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *Journal of Personality and Social Psychology, 85,* 1016.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23,* 3343–3354. http://dx.doi.org/10.1162/jocn_a_00016

Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research, 10,* 431–452. http://dx.doi.org/10.1007/BF02683293

Yarmey, A. D. (1971). Recognition memory for familiar" public" faces: Effects of orientation and delay. *Psychonomic Science, 24,* 286–288. http://dx.doi.org/10.3758/BF03329007

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759. http://dx.doi.org/10.1068/p160747

Received February 3, 2015
Revision received October 18, 2015
Accepted October 21, 2015 ■