Racial Bias and Its Relationship with Moral Blame

by

Reagan T. Heath

Honors Thesis

Appalachian State University

Submitted to the Department of Psychology
in partial fulfillment of the requirements for the degree of

Bachelor of Arts

May, 2017

Approved by:

_____
Andrew Monroe, Ph.D., Thesis Director

_____
Rose Mary Webb, Ph.D., Second Reader

_____
Andrew Smith, Ph.D., Departmental Honors Director

Racial Bias and Its Relationship with Moral Blame

Reagan T. Heath

Appalachian State University

Abstract

The proposed studies attempt to synthesize the two areas of research: moral judgment and racial bias in hopes to establish a link between a perpetrator's race and the amount of blame an individual gives them. In cognitive research, the process of the blame requires a step-by-step process of noticing an event, identifying whether an agent is involved, deciding whether the event was intentional, and then reviewing the agent's justifications, obligations to prevent the occurrence, and their ability to prevent the occurrence. Racial bias research has clear evidence of prejudice between own-race and other-race attitudes. This can be shown through the Implicit Association Test (IAT) and the Attitudes Towards Blacks (ATB) scale. The hypothesis in this research is that participants will blame other-race perpetrators more, and own-race perpetrators less than a no-race-given control for the same moral trespass. There are two competing theories on how this works. The first is that racial attitudes directly bias blame judgments without affecting the underlying information processing leading to blame. The second theory hypothesizes that racial attitudes intensify blame judgments by changing the way people evaluate the informational components on which blame depends (i.e., causality, intentionality, reasons, preventability). . The present research tests these two competing theories using a combination of previously validated measures, such as the IAT and ATB, as well as recently developed vignettes that will measure participants' amount of blame for race-specific perpetrators. Overall, the research hopes to conclude that race has a significant relationship with the process of blame.

Keywords: racial bias, prejudice, moral judgment, blame

Racial Bias and Its Relationship with Moral Blame

Recent, heavily publicized, police shootings of unarmed Black men have sparked increased concern over racial bias in police decisions to use lethal force. One proposal to address these shootings is to require police to wear body camera so that there would be first-hand evidence of the event. However, people's reactions to recent cases where shootings were videotaped show that people often disagree on the acceptability of what occurred. In essence, two people can view the same footage, and come to completely separate conclusions. This difference in perceptions also causes problems when choosing where to place blame.

While numerous studies have detailed the inputs to blame (Malle, Guglielmo, & Monroe, 2012; 2014), relatively little work to date has examined the process people go through from observing a moral violation to rendering and expressing a moral judgment of blame. The current work seeks to examine the process of blaming, and more specifically how personal racial biases might influence that process. Past research demonstrates broad consensus that racial beliefs bias moral and legal judgments; however, it remains unclear exactly how these beliefs operate in the process of blame.

The current work proposes two experiments to test how racial bias might influence moral judgments of blame. Below I review the existing research on the criteria and the process of blame; I review the evidence for racial bias influencing judgments in non-moral domains; and I outline my predictions and proposed experiments.

The Path Model of Blame

Malle et al. (2014) recently proposed the Path Model of Blame. This model is unique in that it not only defines the necessary informational components for blame, but it also specifies the information processing structure people move through from perceiving a moral violation

(e.g., a dead body on the ground) to rendering a moral judgment of blame. That is, the Path Model goes beyond previous moral judgments models that enumerate the relevant inputs to blame, by specifying how those inputs go together in a cognitive process of blaming.

According to the model, the process of blame begins with detecting a moral norm violation. This requires a system of broadly agreed upon norms of behavior against which perceivers can compare violations. Moral norms vary widely. Some researchers argue that moral judgments arise in response to violations of distinct moral domains, including care, sanctity, fairness, authority, and ingroup loyalty (Graham, Haidt, & Nosek, 2009). Other researchers suggest that moral norms relate to the motivation to regulate different types of social relationships such maintaining cooperation or social hierarchies (Rai & Fiske, 2011); whereas others distinguish between behaviors that *should* be performed (i.e., prescriptive norms) and behaviors that *should not* be performed (i.e., proscriptive norms; Janoff-Bulman & Carnes, 2013). Regardless of the specific norms that perceivers are sensitive to, detecting moral norm violations can be done without Theory of Mind (ToM) or emotionality, as individuals with autism can reliably detect norm-violating events (Zalla, Sav, Stopin, Ahade, & Leboyer, 2009); also psychopaths can do it (Maxwell & Le Sage, 2009).

After detecting a moral violation, perceivers evaluate whether the cause of the event was a morally responsible agent. The concept of agency emerges early in development based on features such as self-propelledness (Premack, 1990) and goal-directed action (Woodward, 1998). Bandura (2006) argued that agency requires four key mental capacities: intentionality, forethought, self-reactiveness, and self-reflectiveness (Bandura, 2006). Thus, agency requires that an agent can control one's behavior and act in light of relevant norms (see Monroe, Dillion, & Malle, 2014). If the cause of the event is not a morally responsible agent (e.g., if lightning

struck the victim dead or if an infant accidentally shot the victim), then little or no blame is assigned. If, however, the cause of the event was a morally responsible agent then perceivers consider whether the agent caused the outcome intentionally.

Judgments of intentionality are unique in the process of blame. Intentionality is a critical input to blame. People quickly perceive intentionality in everyday situations (Malle & Holbrook, 2012), often perceptually (Scholl & Tremoulet, 2000) or as part of scripts (Schank & Abelson, 1977), but they may also more carefully consider if an agent acted intentionally if actions are ambiguous (Guglielmo & Malle, 2010a, 2010b; Monroe & Reeder, 2011) or if the weight of the judgment demands it (Reeder, 2009). Moreover, previous research demonstrates that intentionality amplifies blame (Darley & Shultz, 1990; Gray & Wegner, 2008; Lagnado & Channon, 2008; Ohtsubo, 2007; Young & Saxe, 2009).

Intentionality also structures the process of blaming by bifurcating information processing onto one of two independent tracks (Monroe & Malle, 2017). If a behavior is judged to be intentional, then perceivers evaluate the agent's reasons for acting. Providing morally good reasons for harm (e.g., wanting to protect one's family from attack) will mitigate blame, while morally bad reasons (e.g., attacking someone in cold blood) will exacerbate it (Howe, 1991; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Tetlock, Self, Singh, 2010). Recently, Greene et al. (2009) found that inferences about motives shape people's moral judgments in sacrificial moral dilemmas, such as the trolley problem, where participants are asked to a decision to either allow five people to die, or to save them by killing one other person. When the decision to kill one to save five is described as malicious, participants are much less likely to accept his actions as morally permissible; however, when sacrificing one workman is described

as a side effect of a goal to save the lives of the others, acceptance increases (Greene et al., 2009).

By contrast, if the behavior is deemed unintentional, perceivers evaluate whether the agent had the *obligation* to prevent the outcome, and whether the agent had the *ability* to prevent the outcome. In this instance, the observer is determining alternate possibilities (what could or should have happened; Mandel & Lehman, 1996). The ability to prevent an outcome is based on the agent's cognitive capability to foresee harm and her physical capability to prevent such harm. As such, when an agent foresaw a negative event and did nothing to stop it, they were assigned significantly more blame (as opposed to when the agent did not foresee the negative event). This factor is referred to as *foreseeability*, and has been shown in both children and adults (Nelson-le Gall, 1985; Shaw & Sulzer, 1964). Foreseeability would be considered a measure of a person's mental capacity, and as such, people with mental disability are assigned less foreseeability and thus less blame when negative events occur. However, physical capacity is also a measure an agent's ability to prevent a negative event. For example, obesity not due to a medical condition intensifies blame rather than mitigating because the agent could have controlled his or her obesity (Weiner, 1995). Unfortunately, this also extends to victim blaming in rape cases, where perceivers begin to think of alternate behaviors that victim could have taken (Catellani, Alberici, & Milesi, 2004). This type of behavior during the blame process can then lead to self-blame, as many sexual assault victims end up blaming themselves (Davis et al., 1996; Janoff-Bulman, 1979; Janoff-Bulman & Wortman, 1977). This happens because victims focus on their actual capacity to take alternate actions, but do not take into account their obligation to take such actions.

Comparatively, not as much research has examined agent's obligations to prevent negative events. One of the main factors, though, that affects obligation to prevent is the agent's role. For example, doctors have the obligation to keep their patients healthy, so if a patient dies due to a doctor's negligence, blame is assigned because the doctor did not fulfill his or her duty. All this being said, if the person either had no obligation to or was unable to prevent the negative circumstance, blame is not generally assigned. However, if they were expected to prevent, and they were also fully able to do so but did not, then blame in varying degrees is assigned.

Information Processing, Preset Values, and Bias

In addition to making predictions about the structure of information processing en route to blame, the Path Model makes a second set of predictions about how people set the values of each individual concept in the model (e.g., intentionality, reasons, or preventability). The model predicts that once a concept is activated, perceivers gather information relating to that concept, and then use that information to determine its value (e.g., resolving that "Fred intentionally hit Jim."). Information acquisition includes the retrieval of knowledge (remembering the agent's role), informational searching (questioning an agent's intentionality or justifications), and simulation (what could the agent have done better?) among other things. This part of the process can either be automatic or deliberate, when information about a concept is not immediately obvious. Cases where this is automatic rely on event-implied information (e.g., a man standing over a dead body holding a smoking gun), which is information gained from observing the event or its results.

Importantly, one way people can fill in information required for a concept is with preset values. These are shortcuts that are activated by knowledge structures such as a target's race, occupation, age, religion, and many other constructs. For instance, preset values may be

associated with certain roles (as mentioned earlier with doctors), group membership (e.g., rivals always intentionally harm us) or erroneous beliefs. For example, people who subscribe to rape myths may have systematically different preset values for the preventability of sexual assault (i.e., believing that the victim could have prevented the assault had she been dressed differently or not been drinking) compared to people who reject such myths (Grubb & Turner, 2012; McCaul, Veltum, Boyechko, & Crawford, 1990).

These presets values are critically important, as they may be one way systematic bias may enter into the process of blame. If people have different implicit or explicit preset values when making moral judgments of White versus Black norm violators, then one would expect that based on these different presets people could observe the same behavior and yet draw different moral conclusions about the agent. While no study to date has examined such a question, evidence that racial bias influences other types of morally-relevant decisions is well documented.

Racial Bias in Decision-Making

There is a large field dedicated to investigating racial bias, and how it affects the ways people view racial out-group members, though as of yet there has been no research solely dedicated to how these attitudes enter the blame process. Racial prejudice is defined as negative attitudes, beliefs, or emotions motivated by a target's race (Blumer, 1958; Herek, 2004) However, whereas prejudice refers to a particular set of biased attitudes a person may hold, discrimination refers to specific behaviors motivated by prejudicial attitudes. Acts of discrimination may include excluding out-group members from certain activities, preferring to socialize solely with in-group members, giving privileges to in-group members, and violence against out-group members.

Implicit and Explicit Bias

Importantly, researchers differentiate between explicitly held prejudicial attitudes and implicit prejudice. While both correspond to an individual's biased attitudes, implicit bias refers to a suite of attitudes that are typically unconscious and automatic (see Gawronski, Hofmann, & Wilbur, 2006; Kahn & Martin, 2016); whereas explicit prejudice refers to a person's consciously accessible, controlled, and (perhaps even) endorsed biases (Rydell & McConnell, 2006; Thurstone, 1928; Thurstone & Chave, 1929).

Because implicit and explicit prejudice operate at different speeds and levels of conscious awareness, researchers use different tools to measure them. For example, common measures of explicit prejudice include the Attitudes Toward Blacks scale (ATB) or the Modern Racism scale developed by (McConahay,1986). These measures ask people to self-report their attitudes towards a target group, which can be any chosen group, such as Blacks or homosexuals. For example, one question on the ATB scale is: "I worry that in the next few years I may be denied my application for a job or a promotion because of preferential treatment given to minority group members." Most measures of explicit bias are designed to measure beliefs that individuals openly state they hold, and are more frequently used than measures of implicit bias (Greenwald & Banaji, 1995).

Whereas explicit prejudice is consciously accessible, measures of explicit prejudice rely on people's self-reported attitudes. Implicit prejudice, by contrast, is thought to be automatic and not available to introspection; therefore, measures of implicit prejudice rely on more subtle measures, such as the speed with which people are able to make positive versus negative associations with different racial targets. One of the most prominent measures of this kind is the Implicit Association Test (IAT) developed by Greenwald and Banaji (1995). The IAT measures implicit prejudice by presenting participants with a series of association tasks with stimuli of

cropped photos of Black or White target faces. Participants are also given positive or negative stimuli words such as “good” and “beautiful” or “bad” and “terrible.” In the first task, participants are asked to pair the White faces with positive words, and the Black faces with negative words. In the second task, this is reversed. Participants must pair Black faces with positive terms, and White faces with negative terms. The IAT then measures the response times for each task, as well as the number of mistakes each participant made, and then calculates a level for the participant’s implicit biases.

Not only are explicit and implicit prejudice measured with different tools, but they also are differentially predictive of behavior. On one hand, explicit biases are conscious and sometimes endorsed, while implicit biases are unconscious and cannot be controlled by those who hold them (Dovidio, Kawakami, & Gaertner, 2002). Dovidio, Kawakami, Johnson, Johnson, and Howard (1997) showed that explicit prejudice correlated with deliberate types of discriminatory behavior whereas implicit attitudes correlated with only unintentional behaviors (bodily movements, eye twitching) indicative of discomfort when in the presence of an other-race individual. This suggests that explicit attitudes map more closely onto conscious, intentional behaviors, whereas implicit attitudes better predict automatic or unintentional behaviors.

Therefore, research suggests that implicit bias can be found even when explicit bias is not. However, whether a person acts on their implicit biases or not, actions can be perceived as intentionally biased by some, and unintentionally biased by others. It has also been shown that these implicit biases can be detected through psychological testing, such as the IAT, and what this testing has revealed is that not all biases are directly discriminatory. This is because discrimination implies intentionality, and implicit biases are not intentional. Certain biases, in fact, favor members of the in-group based on shared physical traits based on how easily and

accurately a person is recognized, rather than persecuting out-groups members for traits not shared.

Racial Bias in the World

Much research examines how racial bias operates in the real world. One area where racial bias presents itself is employment. This effect is shown mostly in terms of ethnicity and race with significant disadvantages for Hispanic and African Americans (Braddock & McPartland, 1987; Culp & Dunson, 1986). There are multiple steps in the employment process, one of the first being a job interview. Job interview refers to inviting an applicant to a meeting, in order to gain first person experience of their character and qualifications. Here, both explicit and implicit biases can play major roles, as either direct hatred, or "classic racism" as Kirscheman and Neckerman (1991) refer to it, or cultural misunderstandings. This happens because inter-race interactions are still affected by factors like suspicion (Blauner, 1989). This could also have to do with Own-Race-Bias, as interviewers of any race will prefer same-race applicants. Kochman (1983) also points out that White and Black people also misidentify one another's behavioral cues, leading to misunderstandings. These factors taken together, as well as the overwhelming white majority of the population and work force, can stack the odds against minority applicants. Finally, employment tests are exams that measure aptitude for specific job skills. Meta-analyses for research looking into the correlation between test scores and job performance found that the tests favored non-minorities (Burstein & Pitchford, 1990; Hartigan & Wigdor, 1989). These findings suggest (strongly) that minorities are significantly disadvantaged in employment, due to the biases against them, which more often than not, are actually biases favoring the majority's in-group, and are not intentionally discriminatory, but sometimes produce that effect.

Another area that racial bias works against minorities is police shootings. Minorities often self-report police abuse (Weitzer & Tuch, 2004), and bias in law enforcement is difficult to measure. Correll, Park, Judd, and Wintenbrink (2002) developed a videogame in his article the Police Officer's Dilemma as a measure of people's implicit against Black men in shooting decisions. In this task subjects make quick decisions – in the order of seconds – of whether to shoot or not to shoot a person holding an ambiguous object in his hand. The test varies the race of the target person (White vs. Black) and the object the person is holding (a gun vs. a tool). The task measures two key variables: the speed with which people make shoot/don't shoot decisions and the number of errors people make (e.g., shooting an unarmed person or failing to shoot an armed person). Results show that people are more likely to mistakenly shoot unarmed Black men. Interestingly, however, Correll et al. (2002) found that this error pattern holds for both White and Black participants.

Correll et al. (2007) developed another experiment to test the effects of implicit bias on shootings made by non-police officers. When not given enough time to consciously process a situation, people will most likely mistakenly shoot the unarmed Black man. They showed that although laypeople again showed the shooter bias effect against black targets as opposed to white targets, police officers did not show such a bias. However, both police and laypeople showed a reaction time bias. Both sets of participants made faster shoot decisions when the target was an armed Black man compared to an armed White man. By contrast when the target person was unarmed it took both sets of participants longer to decide not to shoot the unarmed Black man compared to deciding not to shoot the unarmed White man. Thus, while police did not show the shooter bias that laypeople did, the pattern of response times demonstrate that even trained police

find it easier to associate Black targets with danger compared to White targets. However, as time was manipulated in order for faster decisions to be made, error rates in shooting decisions rose.

Finally, racial bias is also heavily present in the courtroom. Statistically, White jurors will treat Black defendants worse than White defendants (Kang, Bennett, Carbado, & Casey 2012). Finally, research by Levinson, Cai, and Young (2009) looked into mock jury cases. The researchers found that when the defendant was Black, they were much more likely to be deemed guilty. Furthermore, when asked to recall, most of the mock jurors could not consciously state the defendant's race, implying that it was their implicit biases that affected the overall guilty/not guilty decision. Finally, the researchers looked into the legal ramifications of the IAT, but instead of using legally ambiguous terms like "good" or "bad," they proposed using "guilty" and "not guilty." The results of their "Legal IAT" found that participants did show implicit biases with stronger associations being drawn between Blacks (as opposed to Whites) and "Guilty," as well as Whites and "Not Guilty."

Racial bias is a far-reaching issue that affects society in many ways. In some cases, bias is explicit, making it easy to detect and deal with. However, bias can also be implicit, or unconscious, making it very difficult for even the person with the biases to notice, let alone correct. Implicit bias reaches in many aspects of the world, from employment to law enforcement, and as of yet, there has not been a method produced that can significantly reduce implicit biases.

Synthesis

The Path Model (Malle et al., 2014) proposed a model of blame that outlines the necessary criteria for rendering moral judgments of blame as well as specifying the information processing requirements for blame. Critical in this model is the prediction that the values of

informational nodes of the model (e.g., intentionality) can be preset by people's individual beliefs and attitudes. If correct, these informational presets would allow us model how bias might get into the process of blame.

My project examines one possible relationship between a specific type of bias, racial prejudice, and moral blame. Whereas previous research has established ways of measuring explicit and implicit and demonstrated their relationships with morally relevant outcomes such as discriminatory behavior (Blumer, 1958; Herek, 2004), hypothetical shooting decisions (Correll et al. 2002), hiring decisions (Braddock & McPartland, 1987; Culp & Dunson, 1986), and jury decision making (Levinson et al., 2009), no work to date has specifically examined the relationships of these biases with moral decision-making more broadly. In doing so, I outline and test two alternative hypotheses for how racial attitudes might affect moral judgments of blame.

One hypothesis for how racial attitudes might affect blame judgments is that racial attitudes intensify blame judgments directly, leaving information processing (e.g., considering an agents reasons or intentions) unaffected. From this hypothesis, people who strongly endorse racially-biased attitudes would make harsher moral judgments of blame compared to people who do not endorse racially-biased attitudes; however, both people with high and low racial biases would appraise a target's intentions, reasons, and causal contributions identically.

An alternative hypothesis, derived from the Path Model of Blame (Malle et al., 2014) is that racial attitudes intensify blame judgments by changing the way people evaluate the informational components on which blame depends (i.e., causality, intentionality, reasons, preventability). From this view, people who strongly endorse racially-biased attitudes will be more likely to inflate a target's causal contributions to a negative outcome; to perceive targets as causing harm intentionally (rather than unintentionally); to perceive agents as acting for

unjustified or morally bad reasons (compared to morally good reasons); or to perceive agents as being able to prevent harm (rather than not) compared to people who do not endorse racially-biased attitudes. Thus, from this view, people who strongly endorse racially-biased attitudes arrive at harsher overall blame judgments, because of bias in their perceptions of the morally relevant evidence (causality, intentions, reasons, preventability).

In order to evaluate each of these hypotheses, experimental designs are being aimed at each theory. To test whether the outcome is being affected (rather than information processing), a moral updating paradigm is being proposed to test explicit biases in the context of the blame model. The participants will be given time to think about the blame judgments they are making, which creates the proper circumstances for explicit biases to show. Participants will also take the IAT and ATB in order to have a standardized measure of bias. Contrariwise, to test whether information processing is being affected, a reaction time paradigm is proposed, since the quick answer model will not allow for explicit biases to show, leaving only implicit biases to affect the participants' responses. These two studies together should both produce results that more clearly define which of the two hypotheses about the nature of the relationship between racial bias and moral blame is better supported.

Proposed Methods

To test these opposing hypotheses, I am proposing two experiments. Study 1 will use a reaction time paradigm to test whether people with higher racial (compared to people with low racial bias) bias more can more easily detect negative mental states for African American targets compared to Caucasian targets. In this study, I will assess participants' explicit and implicit racial attitudes using the ATB (Brigham, 1993) and the IAT (Greenwald & Banaji, 1995) respectively. Afterwards, participants will complete a reaction time paradigm where they will be

presented with brief stimuli sentences (e.g., "[agent] a successful filmmaker, only gives his ailing mother $20 a month.") and then asked to make a speeded mental state inference about the agent's bad/good reasons, bad/good intentions, or capacity/obligation to prevent the action (e.g., "Did the behavior reveal a certain INTENTION the actor has?"). I will manipulate the name of the agent, within-subjects, such that in half of the trials the agent's name will be a stereotypically African American name (e.g., Darius) or a stereotypically White name (e.g., Chuck) in the other half of trials. Thus, this study will test whether people with higher explicit or implicit bias (relative to people with low bias) more quickly and more frequently detect negative mental states for African American targets compared to White targets.

Study 2 will use a moral updating paradigm to test how racial attitudes affect the way people use causal mental state information (good vs. bad reasons, intentional vs. unintentional, preventable vs. unpreventable) to update previously made moral judgments. As in Study 1, I will assess explicit and implicit racial attitudes using the ATB (Brigham, 1993) and the IAT (Greenwald & Banaji, 1995) as part of a broader demographic questionnaire. After these measures, participants will complete the moral updating paradigm. In this paradigm, participants will be presented with a sparse description of a moral event (e.g., [agent] killed Frank) and asked to make an initial moral judgment (e.g., how much blame does [agent] deserve). Then participants will be presented with some additional information describing the agent's reasons, intentions or ability to prevent the event, and then allowed to update their blame judgment if they want to. As in Study 1 I will manipulate the name of the agents, within-subjects, such that in half of the trials the agent's name will be a stereotypically African American name (e.g., Darius) and a stereotypically White name (e.g., Chuck) in the other half of trials. In this way I can examine whether people with high racial bias (relative to low) asymmetrically attend to and use blame-

exacerbating information (e.g., bad reasons, intentional, preventable) over blame-mitigating information (e.g., good reasons, unintentional, unpreventable) in updating moral judgments of blame for African American targets.

References

Abelson, R., & Schank, R. C. (1977). Scripts, plans, goals and understanding. *An inquiry into human knowledge, 1*, 5-10. doi:10.2307/412850

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*, 164-180. doi:10.1111/j.1745-6916.2006.00011.x

Blauner, B. (1989). *Black lives, white lives: Three decades of race relations in America*. Berkeley, CA: University of California Press. doi:10.1007/BF02903827

Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific Sociological Review, 1*, 3-7. doi:10.1007/978-1-349-26403-2_3

Braddock, J. H., & McPartland, J. M. (1987). How minorities continue to be excluded from equal employment opportunities: Research on labor market and institutional barriers. *Journal of Social Issues, 43*, 5-39. doi:10.1111/j.1540-4560.1987.tb02329.x

Brigham, J. C. (1993). College students' racial attitudes. *Journal of Applied Social Psychology, 23,* 1933-1967. doi:10.1111/j.1559-1816.1993.tb01074.x

Burstein, P., & Pitchford, S. (1990). Social-scientific and legal challenges to education and test requirements in employment. *Social Problems, 37,* 243-257. doi:10.1525/sp.1990.37.2.03a00080

Catellani, P., Alberici, A. I., & Milesi, P. (2004). Counterfactual thinking and stereotypes: The nonconformity effect. *European Journal of Social Psychology, 34*, 421-436. doi:10.1002/ejsp.206

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83,* 1314-1329. doi:10.1037//0022-3514.83.6.1314

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology, 92,* 1006-1023. doi:10.1037/0022-3514.92.6.1006

Culp, J., & Dunson, B. H. (1986). Brothers of a different color: A preliminary look at employer treatment of white and black youth. *In The Black Youth Employment Crisis* (pp. 233-260). Chicago, IL: University of Chicago Press.

Darley, J. M., & Shultz, T. R. (1990). Moral rules: Their content and acquisition. *Annual Review of Psychology, 41*, 525-556. doi:10.1146/annurev.ps.41.020190.002521

Davis, C. G., Lehman, D. R., Silver, R. C., Wortman, C. B., & Ellard, J. H. (1996). Self-blame following a traumatic event: The role of perceived avoidability. *Personality and Social Psychology Bulletin, 22,* 557-567. doi:10.1177/0146167296226002

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82,* 62-68. doi:10.1037/0022-3514.82.1.62

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510-540. doi:10.1006/jesp.1997.1331

Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition, 15*, 485-499. doi:10.1016/j.concog.2005.11.007

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*, 1029-1046. doi:10.1037/a0015141

Gray, K., & Wegner, D. M. (2008). The sting of intentional pain. *Psychological Science*, *19*, 1260-1262. doi:10.1111/j.1467-9280.2008.02208.x

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, *111*, 364-371. doi:10.1016/j.cognition.2009.02.001

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27. doi:10.1037//0033-295X.102.1.4

Grubb, A., & Turner, E. (2012). Attribution of blame in rape cases: A review of the impact of rape myth acceptance, gender role conformity and substance use on victim blaming. *Aggression and Violent Behavior*, *17*, 443-452. doi:10.1016/j.avb.2012.06.002

Guglielmo, S., & Malle, B. F. (2010a). Enough skill to kill: Intentionality judgments and the moral valence of action. *Cognition*, *117*, 139-150. doi:10.1016/j.cognition.2010.08.002

Guglielmo, S., & Malle, B. F. (2010b). Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality and Social Psychology Bulletin*, *36*, 1635-1647. doi:10.1177/0146167210386733

Hartigan, J., & Wigdor, A. (1989). Fairness in employment testing. *Science*, *245*, 14. doi:10.1126/science.2740906

Herek, G. M. (2004). Beyond “homophobia”: Thinking about sexual prejudice and stigma in the twenty-first century. *Sexuality Research & Social Policy*, *1*, 6-24. doi:10.1525/srsp.2004.1.2.6

Howe, E. S., & Loftus, T. C. (1991). Integration of intention and outcome information by students and circuit court judges: Design economy and individual differences. *Journal of Applied Social Psychology*, *22*, 102-116. doi:10.1111/j.1559-1816.1992.tb01524.x

Kahn, K. B., & Martin, K. D. (2016). Policing and race: Disparate treatment, perceptions, and policy responses. *Social Issues and Policy Review, 10*, 82-121.

Kang, J., Bennett, M., Carbado, D., & Casey, P. (2011). Implicit bias in the courtroom. *University of California School of Law Review., 59*, 1124. doi: 10.1111/sipr.12019

Kirschenman, J., & Neckerman, K. M. (1991). “We’d love to hire them, but...”: The meaning of race for employers. *The Urban Underclass*, *203*, 203-224.

Kochman, T. (1983). The boundary between play and nonplay in black verbal dueling. *Language in Society*, *12*, 329-337. doi:10.1017/S0047404500009970

Janoff-Bulman, R. (1979). Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of Personality and Social Psychology*, *37*, 1798-1809. doi:10.1111/sipr.12019

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the moral landscape moral motives and group-based moralities. *Personality and Social Psychology Review*, *17*, 219-236. doi:10.1177/1088868313480274

Janoff-Bulman, R., & Wortman, C. B. (1977). Attributions of blame and coping in the" real world": severe accident victims react to their lot. *Journal of Personality and Social Psychology*, *35*, 351-363. doi:10.1037/0022-3514.35.5.351

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*, 754-770. doi:10.1016/j.cognition.2008.06.009

Levinson, J. D., Cai, H., & Young, D. (2009). Guilty by implicit racial bias: The guilty/not guilty Implicit Association Test. *Ohio State Journal of Criminal Law*, *8*, 187-208.

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology, 102,* 661-684. doi:10.1037/a0026790

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2012). Moral, cognitive, and social: The nature of blame. In J. Forgas, K. Fiedler, and C. Sedikides (Eds.), *Social thinking and interpersonal behaviour* (pp. 311-329). Philadelphia, PA: Psychology Press.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry, 25,* 147-186. doi:10.1080/1047840X.2014.877340

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology, 71,* 450-463. doi:10.1037/0022-3514.71.3.450

Maxwell, B., & Le Sage, L. (2009). Are psychopaths morally sensitive? *Journal of Moral Education, 38,* 75-91. doi:10.1080/03057240802601680

McCaul, K. D., Veltum, L. G., Boyechko, V., & Crawford, J. J. (1990). Understanding attributions of victim blame for rape: Sex, violence, and foreseeability. *Journal of Applied Social Psychology, 20,* 1-26. doi:10.1037/a0026790

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. *Personality and Social Psychology Bulletin, 9,* 551-558. doi:10.1177/0146167283094004

Monroe, A. E., Dillon, K. D., & Malle, B. F. (2014). Bringing free will down to Earth: People's psychological concept of free will and its role in moral judgment. *Consciousness and Cognition, 27,* 100-108. doi:10.1016/j.concog.2014.04.011

Monroe, A. E., & Malle, B. F. (2017). Two paths to blame: Intentionality directs moral information processing along two distinct tracks. *Journal of Experimental Psychology: General, 146*, 123-133. doi:10.1037/xge0000234

Monroe, A. E., & Reeder, G. D. (2011). Motive-matching: Perceptions of intentionality for coerced action. *Journal of Experimental Social Psychology, 47*, 1255-1261. doi:10.1016/j.jesp.2011.05.012

Nelson-le Gall, S. A. (1985). Motive–outcome matching and outcome foreseeability: Effects on attribution of intentionality and moral judgments. *Developmental Psychology, 21*, 332-337. doi:10.1037/0012-1649.21.2.332

Ohtsubo, Y. (2007). Perceived intentionality intensifies blameworthiness of negative behaviors: Blame-praise asymmetry in intensification effect. *Japanese Psychological Research, 49*, 100-110. doi:10.1111/j.1468-5884.2007.00337.x

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1-16. doi:10.1016/0010-0277(90)90051-K

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*, 57-75. doi:10.1037/a0021867

Reeder, G. D. (2009). Mindreading: Judgments about intentionality and motives in dispositional inference. *Psychological Inquiry, 20*, 1-18. doi:10.1080/10478400802615744

Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: the role of perceived motive. *Journal of Personality and Social Psychology, 83*, 789-803. doi:10.1037/0022-3514.83.4.789

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology, 91,* 995-1008. doi:10.1037/0022-3514.91.6.995

Shaw, M. E., & Sulzer, J. L. (1964). An empirical test of Heider's levels in attribution of responsibility. *The Journal of Abnormal and Social Psychology, 69,* 39-46. doi:10.1037/h0040051

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4,* 299-309. doi:10.1016/S1364-6613(00)01506-0

Tetlock, P. E., Self, W. T., & Singh, R. (2010). The punitiveness paradox: When is external pressure exculpatory–And when a signal just to spread blame? *Journal of Experimental Social Psychology, 46,* 388-395. doi:10.1016/j.jesp.2009.11.013

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33,* 529-554. doi:10.1086/214483

Thurstone, L. L., & Chave, E. J. (1929). The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the church. Chicago, IL: University of Chicago Press. doi:10.1037/11574-000

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York, NY: Guilford Press.

Weitzer, R., & Tuch, S. A. (2004). Race and perceptions of police misconduct. *Social Problems, 51,* 305-325. doi:10.1525/sp.2004.51.3.305

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69,* 1-34. doi:10.1016/S0010-0277(98)00058-4

Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47,* 2065-2072. doi:10.1016/j.neuropsychologia.2009.03.020

Zalla, T., Sav, A. M., Stopin, A., Ahade, S., & Leboyer, M. (2009). Faux pas detection and intentional action in Asperger Syndrome. A replication on a French sample. *Journal of Autism and Developmental Disorders, 39,* 373-382. doi:10.1007/s10803-008-0634-y