Title
Evolved Error Management Biases in the Attribution of Anger

Author
Galperin, Andrew

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Evolved Error Management Biases in the Attribution of Anger

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in Psychology

by

Andrew Galperin

2012

ABSTRACT OF THE DISSERTATION

Evolved Error Management Biases in the Attribution of Anger

by

Andrew Galperin

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2012

Professor Martie Haselton, Chair

Judgments of others' emotional states and inclinations had recurrent fitness consequences for ancestral humans. Such judgments, made under uncertainty, can result in false positives (overestimating an emotional state or trait) or false negatives (underestimating the same). When the costs of these two errors consistently and historically differed, natural selection will have favored a bias toward making the less costly error. The perception of anger entails such asymmetry. Averaged across situations over evolutionary time, underestimation of anger was the more costly error, as the fitness decrements resulting from physical harm or death due to insufficient vigilance were greater than those resulting from lost social opportunities due to excessive caution. I therefore hypothesized that selection has favored an overestimation bias in the evaluations of others' state anger and general anger-proneness (trait anger) relative to evaluations of other traits to which this error asymmetry does not apply. Moreover, I

hypothesized that additional attributes which make the actor more dangerous or make the observer more vulnerable increase the error asymmetry with regard to inferring state anger and trait anger, and should therefore correspondingly increase this overestimation bias. In the study described in Chapter 2 (N = 292), models photographed with a neutral expression and holding a potentially dangerous household object were judged to be angrier than those holding a harmless object, even though the held object provided no normative information about their emotional state. In Chapter 3, Study 1 (N = 161), a fictitious individual depicted in a vignette was judged to have higher trait anger than trait disgust even after controlling for the raters' perception of how emotionally he behaved. Moreover, trait anger ratings were more responsive than trait disgust ratings to behavioral cues of emotionality. In Chapter 3, Study 2 (N = 335), participants viewed images of angry or fearful faces. The interaction of factors indicating target's formidability (male sex), target's intent to harm (direct gaze), and perceiver's vulnerability (being female or high belief in a dangerous world) increased ratings of the target's trait anger but not trait fear. In aggregate, these results suggest a domain-specific bias in evaluating anger.

The dissertation of Andrew Galperin is approved.

Thomas Bradbury

Gregory Bryant

Daniel Fessler

Kerri Johnson

Matthew Lieberman

Martie Haselton, Committee Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

ACKNOWLEDGMENTS

My years at UCLA have been extremely rewarding and intellectually stimulating, and I have many people to thank for that. First and foremost, I would like to thank my primary advisor, Martie Haselton, who was helpful, patient, and extremely supportive throughout my graduate career. She has played an immeasurable role in shaping how I think about both the intricacies and broader social impact of research. Martie also provided enormous help with my countless fellowship and job applications, and gave me a unique teaching opportunity that turned out to be one of my best experiences in the program. I am also grateful to the other members of my committee for their constructive feedback on my dissertation and for their overall guidance during the process: Dan Fessler, Kerri Johnson, Greg Bryant, Matt Lieberman, and Tom Bradbury.

I would like to thank Anne Peplau, who has been an unofficial mentor to me and so many other graduate students. She was instrumental in bringing the Interdisciplinary Relationships Science Program to UCLA, which funded me and played an important role in my development as a scholar. I would also like to thank my fellow (former) student Dave Frederick. Besides being a great friend, he was an invaluable source of advice on every aspect of graduate school. Among many other things, he was a model for how to work efficiently, write successful fellowship applications, and teach in an effective and engaging way. My acknowledgments would not be complete without thanking Jim Roney, who first introduced me to the wonderful and insightful ideas of evolutionary psychology when I was in college.

I am grateful for all the wonderful friends I have made here. I hope I will remain connected with them for the rest of my life. The UCLA psychology department, and especially

the social psych area, truly attracts a diversity of wonderfully eccentric individuals. They are the reason it is hard for me to move on from UCLA, and I will really miss all of them.

Last but not least, I am deeply thankful to my parents for being so loving and supportive. By contributing to and holding me to a high standard in every stage of my education, they are the ones who made it possible for me to receive a Ph.D.

Portions of this manuscript are in preparation for being submitted for publication. I would like to thank the co-authors whole-heartedly for their contributions to these articles.

Chapter 2: Galperin, A., Fessler, D. M. T., Haselton, M. G., Johnson, K. L., Holbrook, C., & Bryant, G. A. (in preparation). *Anger judgments are informed by affordances for doing harm.*

Chapter 3: Galperin, A., Fessler, D. M. T., Haselton, M. G., & Johnson, K. L. (in preparation). *Angry is as angry does (or looks): Biases in the attribution of anger.*

VITA / BIOGRAPHICAL SKETCH

EDUCATION

2007: M.A. in Social Psychology, University of California, Los Angeles
2003: B.A. in Psychology, University of Chicago

AFFILIATIONS

UCLA Interdisciplinary Relationships Science Program
UCLA Center for Behavior, Evolution, and Culture

PUBLICATIONS

Galperin, A., Fessler, D. M. T., Haselton, M. G., Johnson, K. L., Holbrook, C., & Bryant, G. A. (in preparation). *Anger judgments are informed by affordances for doing harm.*

Galperin, A., Fessler, D. M. T., Haselton, M. G., & Johnson, K. L. (in preparation). *Angry is as angry does (or looks): Biases in the attribution of anger.*

Galperin, A., Haselton, M. G., Frederick, D. A., von Hippel, W., Poore, J. C., Buss, D. M., et al. (under review). *Sexual regret*.

Haselton, M. G. & Galperin, A. (in press). Error management in relationships. To appear in J. A. Simpson & L. Campbell (Eds.), *Handbook of Close Relationships*. Oxford University Press.

Galperin, A., & Haselton, M. G. (in press). Error management and the evolution of cognitive bias. To appear in J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social Thinking and Interpersonal Behavior.* New York: Psychology Press.

Galperin, A., & Haselton, M. G. (2010). Predictors of how often and when people fall in love. *Evolutionary Psychology*, *8*, 5-28.

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., et al. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition, 27*, 733-763.

Frederick, D. A., Pillsworth, E. G., Galperin, A., Gildersleeve, K. A., Filossof, Y. R., Fales, M. R., et al. (2009). Analyzing evolutionary social science and its popularizations: Review of the book *The caveman mystique: Pop-Darwinism and the debates over sex, violence, and science*. *Evolution and Human Behavior*, *30*, 301-304.

Frederick, D. A., Galperin, A., Gildersleeve, K. A., Filossof, Y. R., Snider, J., Nguyen, D. et al. (2009). Sexuality from an evolutionary perspective: How did evolution shape the mating preferences of men and women? In T. Cooke (Ed.), *Sex and society*. New York: Marshall Cavendish.

ORAL PRESENTATIONS AT CONFERENCES

Galperin, A., Haselton, M. G., Poore, J., von Hippel, B., & Buss, D. M. (2011, April). *Gender differences in regretting sexual actions and inactions.* Talk presented at the 91st annual meeting of the Western Psychological Association, Los Angeles, CA.

Galperin, A., & Haselton, M. G. (2010, April). *Predictors of how often and when people fall in love.* Talk presented at the 90th annual meeting of the Western Psychological Association, Cancun, Mexico.

Galperin, A., Gonzaga, G. C., & Haselton, M. G. (2008, July). *The paradoxical effect of feeling sexual desire for one's partner on affiliation with attractive others.* Talk presented at the 2nd annual meeting of the International Association for Relationship Research, Providence, RI.

TEACHING AND MENTORING

College Academic Mentor (2011-2012)

Sex: From Biology to Gendered Society (Teaching Fellow; 2009-2011)

Evolution and Human Mating (Instructor; 2010, 2011)

Introductory Psychology (Instructor; 2011)

Psychological Statistics (Teaching Assistant; 2010)

Social Psychology (Teaching Assistant; 2006)

CHAPTER 1

Theoretical Framework

A myriad biases have been documented in human cognition. It appears that people often have inaccurate perceptions and make decisions that do not maximize their normative utility (Krueger & Funder, 2004). Rather than merely examining the proximate mechanisms via which these biases are instantiated, as most research has done, the emerging approach of "adaptive rationality" takes a more distal view: it argues that many biases can be best understood as adaptations that maximized survival and reproductive success ("fitness") in the ancestral environments in which they evolved (Haselton et al., 2009). As a byproduct of maximizing fitness, some of these adaptations also reduced the overall accuracy of judgments because fitness, not accuracy, is the ultimate currency on which natural selection operates.

An adaptive rationality approach to human cognition is useful not only for explaining known biases, but also for predicting previously undocumented ones. This dissertation hypothesizes a novel bias in the perception and attribution of anger, which is rooted in error management theory (Haselton & Buss, 2000; Haselton & Nettle, 2006), a major theory under the umbrella of adaptive rationality. This theory predicts that biases will evolve in human judgments and decisions that fit three criteria: 1) they were made under uncertainty; 2) they had consequences for fitness over the long expanse of human evolution; and 3) the possible perceptions or behavioral choices that could be made had consistently asymmetrical costs.

Judgments of other people's state anger (how angry an individual truly feels at the current moment) and trait anger (determining an individual's general propensity for anger) fit all three criteria. First, emotional states and traits alike are not directly observable, so their evaluations are inherently uncertain. Although emotional states can be signaled (e.g., via an anger expression),

such signals are not perfect indicators of an individual's true feelings or intentions. Second, anger reveals the intention to harm (e.g., Fessler, 2010), so judgments of anger had recurring consequences for self-protection and therefore fitness.

Third, anger inference can result in one of two errors: a *false-negative error* (underestimating the anger of an individual) or a *false-positive error* (overestimating anger). Because it was important to exercise caution around individuals who were either angry in the moment or anger-prone in general, the costs of committing these two errors were not equal. For state anger, underestimating how angry an individual is at the moment (false negative) poses the potential for harm to the perceiver, whereas overestimating his state anger (false positive) perhaps leads to avoiding the individual until the apparent danger has passed. Similarly for trait anger, assuming that an anger-prone individual is temperate (false negative) places the perceiver at risk of assault, whereas assuming that a temperate individual is anger-prone (false positive) merely leads to lost opportunities for mutualistic social interaction. Because the false negative was usually costlier than the false positive in the ancestral past, error management theory predicts that selection favored a false positive bias toward committing the less costly error in most situations. A tendency to commit this more benign error necessarily reduced the chances of committing the opposite and more costly error (Haselton & Nettle, 2006).

Chapter 2 and Chapter 3 are both self-contained research reports. The research in Chapter 2 tests the anger overestimation hypothesis for state anger, and the research in Chapter 3 tests it for trait anger. To circumvent the inherent lack of an objective baseline criterion against which to evaluate a bias, this research compared the perception and attribution of anger to other negative emotions (disgust and fear) as well as to other negative traits (e.g., dishonesty) for which error management logic does not predict as definitive a bias or any bias at all. It also examined

whether cues indicating an exaggerated the error asymmetry would increase the magnitude of the bias, a prediction which directly follows from the logic of error management (Galperin & Haselton, in press). Both Chapters 2 and 3 varied the momentary ability or long-term likelihood that targets would aggress against the observer, and Chapter 3 also measured the observers' self-perceived vulnerability – all factors that affected the asymmetry of underestimating versus overestimating state and trait anger. These manipulations were designed to provide no "real" information about the targets' state or trait anger, making it possible to conclude that the effects of the manipulations constituted biased judgments.

The discussion in Chapter 3 concludes the dissertation with a discussion of its contributions to both the social psychological and evolutionary psychological literatures on cognitive bias, as well as its practical implications.

References

Fessler, D. M. T. (2010). Madmen: An evolutionary perspective on anger and men's violent responses to transgression. In M. Potegal, G. Stemmler, & C.D. Spielberger (Eds.), *Handbook of anger: Constituent and concomitant biological, psychological, and social processes* (pp. 361-381). Springer.

Galperin, A., & Haselton, M. G. (in press). Error management and the evolution of cognitive bias. To appear in J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social Thinking and Interpersonal Behavior.* New York: Psychology Press.

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, 27, 733-763.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78*, 81-91.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47-66.

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27*, 313-327.

CHAPTER 2

Anger Judgments Are Informed by Affordances for Doing Harm

Judgments of others' emotional states and inclinations had recurrent fitness consequences for ancestral humans. Such judgments, made under uncertainty, can result in false positives (overestimating an emotional state or trait) or false negatives (underestimating the same). When the costs of these two errors consistently differ, natural selection favors an "error management" bias toward making the less costly error (Haselton & Buss, 2000; Haselton & Nettle, 2006; Galperin & Haselton, in press; Haselton & Galperin, in press). The perception of anger in a target possessed of the means to do harm entails such asymmetry. Because anger motivates harm infliction (Frank, 1988; Fessler, 2010; Sell, 2009), underestimating anger in someone possessing a potential weapon will generally be more costly than overestimating anger, as failing to anticipate an armed assault that occurs is more costly than exercising excessive caution. The same is not true, however, for judgments of other attributes (e.g., contamination disgust or fear) because, unlike anger, they are not linked to the likelihood of aggression. We tested the hypothesis that people's perceptions of state anger, but not other emotions, are upwardly biased by the target individual's transient capacity to do harm.

Perceptions of emotional states influence judgments of emotional traits via attributional processes (Jones & Davis, 1965; Kelley, 1972). Because behaviors (including emotional expressions) tend to be attributed to people's enduring dispositions (e.g., Gilbert & Malone, 1995), stronger perceived behaviors lead to judgments of stronger corresponding traits. Thus, we predicted that ratings of trait anger, but not other emotional or negative traits, would also be upwardly biased for targets holding dangerous objects.

Methods

American participants (N = 292; 81% Caucasian; 127 women, 159 men, 6 unspecified; mean age 30.4, SD = 10.8) completed an online survey via Mechanical Turk. Participants received $0.25 for survey completion.

Participants viewed a single photo of one of three male models. Each model faced the camera with a neutral expression. Four images of each model were created; in each image, the model held one of four household tools, two of which (a kitchen knife; garden shears) could be used as weapons, and two of which (a spatula; a watering can) were harmless tools used in the same activity. The tools and grasping hand/wrist were digitally inserted such that all other aspects of the photos of a given model were identical (see SM Figure 1 in Supplementary Materials).

Participants rated the model's emotional state (degree of anger, disgust, and fear), emotional traits (inclination to feel anger, disgust, and fear), and other negative traits (unpleasantness, dishonesty, incompetence) on 9 point scales anchored by "not at all" and "extremely" (states) or "much less than average" and "much more than average" (traits). The order of question items was randomized within the state and trait blocks. Participants also indicated which of four emotions (anger, disgust, sadness, or happiness) the model was purportedly induced to feel before assuming a neutral expression and being photographed. Demographic questions followed.

Results

A multivariate between-subjects ANOVA compared state ratings as a function of Tool Class (dangerous or harmless), Model, and Activity Type (cooking or gardening). As predicted,

participants in the dangerous condition rated models as more angry, relative to participants in the harmless condition, $F(1, 274) = 9.01$, $p = .003$, $d = .36$. Neither ratings of state disgust ($F(1, 274) = .464$, $p = .50$) nor fear ($F(1, 274) = .158$, $p = .69$) differed by condition (see Figure 1). When asked to identify the model's emotion prior to being photographed, 50.6% of participants in the dangerous condition categorized the model as angry rather than disgusted, sad, or happy, compared to only 38.6% in the harmless condition, a significant difference, $\chi^2(1, N = 290) = 4.18$, $p = .041$.

An identical multivariate ANOVA was conducted with the trait ratings as dependent measures. As predicted, relative to the harmless condition, participants in the dangerous condition rated models as more anger-prone ($F(1, 268) = 8.77$, $p = .003$). Participants in the dangerous condition also judged models as less disgust-prone ($F(1, 268) = 5.53$, $p = .019$), less fear-prone ($F(1, 268) = 4.60$, $p = .033$), less pleasant ($F(1, 268) = 13.4$, $p < .001$), and less honest ($F(1, 268) = 14.5$, $p < .001$), but equally competent ($F(1, 268) = .206$, $p = .65$) (see SM Figure 2 in Supplementary Materials).

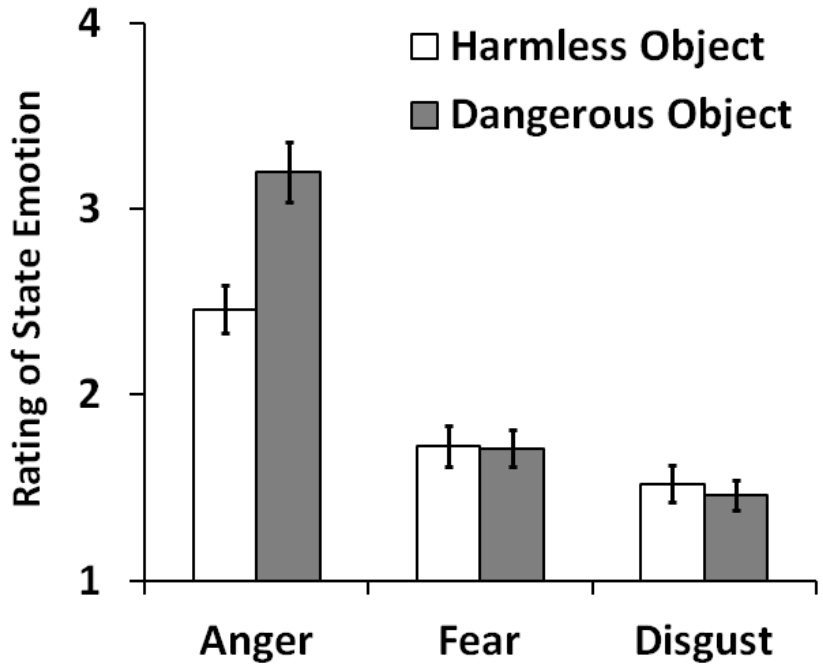
Discussion

The transient capacity to do harm entailed by possession of a tool having affordances as a weapon increased viewers' perceptions of state anger, but not disgust or fear, in models posed in a neutral manner. These cues of danger likewise increased the inference that the model was angry rather than disgusted, sad, or happy. Similarly, danger cues also increased trait anger ratings. The other two negative emotions – trait disgust and fear – showed the opposite effect, providing discriminant evidence that only anger attribution is increased by danger cues. These findings are consonant with an error management interpretation, as fearful and squeamish individuals are less likely to engage in aggression. Therefore, underestimating these traits in a

person who possesses a weapon is less costly than overestimating them. Although not predicted, but consistent with an error management perspective, ratings of unpleasantness and dishonesty were higher in the dangerous condition. Given norms prescribing harmonious relationships, aggressive individuals are, by definition, less pleasant; hence a bias in the attribution of aggression entails a bias in the attribution of unpleasantness. Likewise, because, like aggression, dishonesty entails violating norms and inflicting costs on others, assessments of dishonesty and aggressiveness are likely to be tethered. Importantly, these results are not explicable merely as an indiscriminate "reverse halo effect" (Nisbett & Wilson, 1977), as ratings of incompetence did not differ between conditions.

The shifts in attributions of both state and trait anger reveal a psychological bias, insofar as our tool manipulation was external to, and uninformative of, the target's true emotional state or disposition (see Supplementary Materials). More broadly, these findings highlight the value of interpreting psychological biases in their adaptive contexts (Haselton et al., 2009) and examining how core adaptive motivations, such as self-protection, influence cognition (Kenrick, Neuberg, Griskevicius, Becker, & Schaller, 2010).

Figure 1. State emotion ratings in the dangerous and harmless conditions.



Supplementary Materials

Object Selection. In manipulating the targets' affordance for inflicting physical injury, we were careful not to provide information about the targets' "real" emotional states or personality traits. Manipulating a characteristic inherent to the target (e.g., physical size) would likely shift state and trait anger ratings, but these shifts would not constitute a bias because physically imposing men are more aggressive in reality (Sell, Tooby, & Cosmides, 2009). In contrast, held objects are transient features of the environment that are external to the individual and thereby not informative of his or her emotional state or enduring disposition. This is true as long as these objects are not designed solely to injure others or thought to have been actively chosen by the targets for their dangerous potential. To provide context for the images, participants were told that we asked the target to model as if he were engaging in an everyday activity he enjoys (cooking or gardening, depending on the object). Under these conditions, our results arguably documented a true psychological bias.

Additional Analyses

State Ratings. In the ANOVA predicting state anger, there were no significant interactions of condition (dangerous or harmless) with object type (kitchen or gardening; $F(2, 274) = .001$, $p = .97$), condition with target ($F(2, 274) = 1.92$, $p = .15$), or of all three variables ($F(2, 274) = .627$, $p = .54$). This showed that the main result in the study was not driven by any one target or object type.

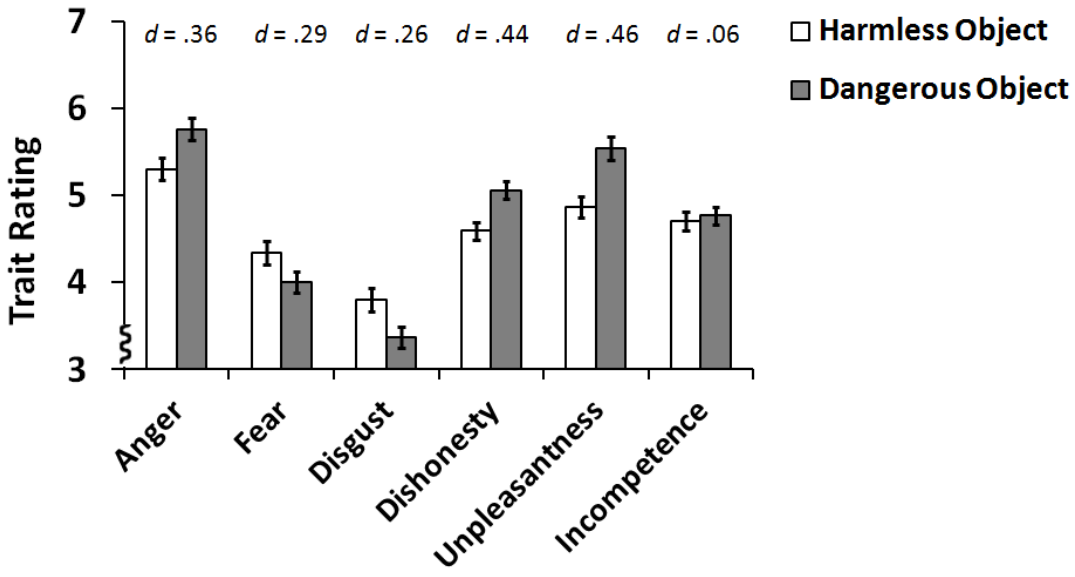
Trait Anger, Controlling for State Anger. An ANCOVA was run with trait anger as the dependent measure; condition, target, and object type as predictors; and state anger as a continuous covariate. The effect of condition on trait anger was no longer significant ($F(1, 273)$

= 3.33, p = .07), consistent with the idea that the higher trait anger ratings in the dangerous condition were a product of the higher state anger ratings.

Normative Responses and Bias. Given that the models posed with a neutral facial expression, truly normative responses in this study would have been 1 out of 9 on state emotion scales and a 25% (chance) likelihood of choosing each of the four emotions in the forced-choice question. Responses in the control (harmless object) condition were generally close to these normative values: state emotions were rated toward the bottom of the scale (as shown in Figure 1), and anger was selected only 34% of the time in the forced-choice question (although more frequently than chance, $\chi^2(1, N = 132) = 13.1, p < .001$). This increases confidence that the dangerous objects, not the harmless ones, produced biased perceptions. Systematic deviations between the normative answers and the harmless-condition answers can likely be attributed to the fact that even truly neutral male faces still appear somewhat angry due to the morphological overlap between angry and masculine facial characteristics (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Zebrowitz, Kikuchi, & Fellous, 2010).

SM Figure 1. All stimuli used in the study.

SM Figure 2. Trait ratings in the dangerous and harmless conditions.



References

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology, 92*, 179-190.

Fessler, D. M. T. (2010). Madmen: An evolutionary perspective on anger and men's violent responses to transgression. In M. Potegal, G. Stemmler, & C. D. Spielberger (Eds.), *Handbook of anger: Constituent and concomitant biological, psychological, and social processes* (pp. 361-381). Springer.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions.* New York: Norton.

Galperin, A., & Haselton, M. G. (in press). Error management and the evolution of cognitive bias. To appear in J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social Thinking and Interpersonal Behavior.* New York: Psychology Press.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117*, 21-38.

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition, 27*, 733-763.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78*, 81-91.

Haselton, M. G. & Galperin, A. (in press). Error management in relationships. To appear in J. A. Simpson & L. Campbell (Eds.), *Handbook of Close Relationships*. Oxford University Press.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47-66.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). San Diego, CA: Academic Press.

Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-26), Morristown, NJ: General Learning Press.

Kenrick, D. T., Neuberg, S. L., Griskevicius, V., Becker, D. V., Schaller, M. (2010). Goal-driven cognition and functional behavior: The fundamental motives framework. *Current Directions in Psychological Science, 19,* 63-67.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*, 250–256.

Sell, A. (2009). Applying adaptationism to human anger: The recalibrational theory. In P. R. Shaver & M. Mikulincer (Eds.), *Understanding and reducing aggression, violence, and their consequences.* Washington DC: American Psychological Association.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences, 106*, 15073–15078.

Zebrowitz, L. A., Kikuchi, M., & Fellous, J. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology, 98,* 175-189.

CHAPTER 3

Angry Is as Angry Does (or Looks): Biases in the Attribution of Anger

Abstract

Anger-prone individuals are volatile and frequently dangerous. Accordingly, inferring the presence of this personality trait in others was important throughout evolutionary history. This inference, made under uncertainty, can result in two types of errors: underestimation or overestimation of trait anger. Averaged across evolutionary time, underestimation will have been the more costly error, as the fitness decrements resulting from physical harm or death due to insufficient vigilance are greater than those resulting from lost social opportunities due to excessive caution. We therefore hypothesized that selection has favored an upwards bias in the estimation of others' trait anger relative to estimations of other traits not characterized by such an error asymmetry. Moreover, we hypothesized that additional attributes that i) make the actor more dangerous, or ii) make the observer more vulnerable increase the error asymmetry with regard to inferring anger-proneness, and should therefore correspondingly increase this overestimation bias. In Study 1 (N = 161), a fictitious individual depicted in a vignette was judged to have higher trait anger than trait disgust, and trait anger ratings were more responsive than trait disgust ratings to behavioral cues of emotionality. In Study 2 (N = 335), participants viewed images of angry or fearful faces. The interaction of factors signaling target's formidability (male sex), target's intent to harm (direct gaze), and perceiver's vulnerability (female sex or high belief in a dangerous world) increased ratings of the target's trait anger but not trait fear.

Angry Is as Angry Does (or Looks): Biases in the Attribution of Anger

Assessing others' personality traits is one of the most fundamental adaptive problems that the human social mind evolved to address. By allowing us to predict others' future behavior, understanding people's personalities facilitates navigating complex social interactions (Ross, 1977). However, a considerable obstacle stands in the way of accurately assessing traits: they are invisible. Past behavior is informative of underlying traits, but such inferences (especially from a single observation) are highly uncertain, for two reasons. First, behaviors are produced by a combination of enduring dispositions and fleeting situations, and the proper discounting of situational influences can only be achieved by observing the consistency of an individual's reactions across a variety of situations (Kelley, 1972), which is not always possible. Second, people manage their behaviors in a strategic manner, actively concealing negative traits.

Here, we explore the hypothesis that people assess a specific trait in others – the propensity to become angry – in a biased way that would have been adaptive in the ancestral past. Protecting oneself from harm is a fundamental human motive, directly linked to fitness (Kenrick, Neuberg, Griskevicius, Becker, & Schaller, 2010). Conspecifics were a primary source of danger for ancestral humans (Keeley, 1996), and the expression of anger reveals harmful intentions (Fessler, 2010; Frank, 1988; Sell, 2009). Equally important to detecting others' current anger is predicting when they might become angry. Doing so depends on estimating the inclination to become angry (trait anger) based on observations of angry behaviors in the past, a process we will refer to as "anger attribution." Importantly, perfect accuracy in this estimation is impossible.

Adaptive Rationality and Error Management

The "adaptive rationality" approach contends that the mind was shaped by selection to maximize fitness rather than accuracy in judgment (Haselton et al., 2009; see also Funder, 1995, and Krueger & Funder, 2004). Whenever fitness maximization did not coincide with accuracy maximization in a particular judgment domain in the ancestral past, human cognition will be characterized by seemingly irrational biases that are in fact "adaptively rational." Anger attribution is one such domain. Perceivers can commit one of two errors: underestimating an individual's trait anger (false negative) or overestimating it (false positive). The false negative was usually costlier than the false positive: assuming that an anger-prone individual is temperate places the perceiver at risk of assault, whereas assuming that a temperate individual is anger-prone merely leads to lost opportunities for mutualistic social interaction. Thus, being accurate on average (i.e., committing false negative and false positive errors with equal frequency) did not maximize fitness over evolutionary time. Rather, in line with Error Management Theory (Haselton & Buss, 2000; Haselton & Nettle, 2006), we hypothesize that selection will have favored a biased tendency to commit the less costly false positive – overestimating trait anger.

In the absence of objective baselines, investigating a hypothesized bias in judgment requires points of comparison. The above logic does not apply to the attribution of other negative emotional dispositions, as no equivalent cost asymmetry obtains across a variety of situations. For instance, across contexts, there is no generalizable difference in the costs of underestimating or overestimating another's propensity to experience disgust; hence we do not expect an evolved bias in this judgment. Judgments of traits such as this serve as the points of comparison in the current work. On the basis of this logic, we developed and tested the hypotheses that follow.

Hypothesis 1: Behaviors indicative of anger will be attributed to personality to a greater degree than behaviors indicative of other negative emotions.

Ancestral error cost asymmetries were not static. Rather, they were influenced by contextual factors (Galperin & Haselton, in press; Haselton & Galperin, in press). Psychological adaptations based on these asymmetries should therefore be influenced by relevant contextual factors. In the case of anger attribution, cues that the target is able or likely to aggress against the perceiver further increase the costs of underperceiving trait anger relative to overperceiving it. This exaggerated error asymmetry would in turn have made erring on the side of caution (i.e., overperceiving trait anger) all the more crucial, leading to an exaggerated dispositional bias.

Cues that another individual poses a threat include attributes of the target individual (e.g., where his anger is being directed), of the perceiver (e.g., self-perceived vulnerability), or a combination thereof. These factors should not affect assessments of other emotion traits because they do not affect the relevant error cost asymmetries.

Hypothesis 2: Increasing the danger that the target poses to the perceiver will increase the dispositional attribution of angry behaviors but not of other types of negative emotional behaviors.

The shift in trait anger ratings predicted by this hypothesis constitutes evidence for a true bias if the manipulated danger factors provide no normative information about the target's putative enduring traits. For this reason, some of the cues manipulated in the current research were transient (e.g., eye gaze direction), and thus not logically indicative of enduring personality.

We conducted two studies to test these hypotheses.

Study 1

Study 1 tested the idea that, all else equal, an unfamiliar individual would be viewed as more dispositionally prone to anger than to another negative emotion (disgust). We asked participants to read vignettes about a fictitious man who reacted with anger and disgust to situations commonly eliciting each emotion, then rate the protagonist's trait anger and disgust. Participants also rated the protagonist's state anger and disgust, and the degree to which they thought he overreacted or underreacted in expressing anger or disgust. These assessments played an integral role in testing the main predictions as covariates, predictors, and mediators. The first prediction, stemming from Hypothesis 1, was as follows:

Prediction 1: The target's trait anger will be rated higher than his trait disgust, and will remain so even after controlling for any systematic discrepancy between the perceived appropriateness of his anger and disgust reactions.

As outlined in Hypothesis 2, the cost asymmetry between underestimating and overestimating trait anger is hypothesized to vary as a function of the target's perceived dangerousness. If the target acts only mildly angry, he presents only a weak potential danger to the observer, and the cost asymmetry is weak. However, if the target seems enraged, he might be a present or future danger to the observer, and the cost asymmetry is strong. This suggests that the stronger the target's state anger reaction, the stronger the overestimation bias, and hence the higher his trait anger will be rated, regardless of how justified his reaction was. This biased pattern is not predicted for disgust: trait disgust is expected to scale with state disgust only to the extent that higher state disgust was judged to be an overreaction.

Prediction 2: Ratings of overreaction will fully mediate the positive association between state and trait ratings for disgust, but will not fully mediate this association for anger.

Methods

Participants and procedure. To prevent trait and state ratings from being artificially similar, participation occurred in two separate sessions held on different days. In exchange for course credit, 441 UCLA undergraduates from two Introductory Psychology classes completed the first session and were provided with a unique ID number. They were subsequently invited to participate in the second session online, and 161 of them did so over the next two months; these individuals constitute the sample. Participant sex and other demographics were not asssesssed.

Materials. In the first session, participants read two of four possible short vignettes about a fictitious college student, designed to portray an individual who was average in his anger and disgust reactions across two situations (see Appendix A). One ("weak") vignette of the pair presented to each participant consisted of a mildly anger- and disgust-provoking situation, followed by mild anger and disgust on the part of the protagonist. The other ("strong") vignette described more serious provocations of anger and disgust, followed by intense angry and disgusted reactions from the protagonist. None of the vignettes contained the words "angry" or "disgusted." Half of the participants read one pair of weak and strong vignettes, the other half read the other pair of weak and strong vignettes. The weak and strong vignettes were presented in randomized order.

In the first session, participants rated the target's trait anger and disgust relative to the average person in randomized order on 1-9 scales. Instructions explicitly asked for ratings of contamination disgust, in contrast to moral outrage (Rozin, Haidt, & McCauley, 2000; Tybur, Lieberman, & Griskevicius, 2009). In Session 2, which occurred between two weeks and two months after Session 1, participants read the same vignettes as before. They rated the absolute degree of the target's state anger and disgust on 1-9 scales, ranging from "none" to "extreme."

They also rated how justified his reaction was relative to the situation on a -3 to 3 scale, ranging from "extreme underreaction" to "extreme overreaction."

Results

Participants judged the target to have displayed somewhat above average anger (M = 6.08, SD = 1.35) and disgust (M = 5.95, SD = 1.32) on the absolute 1-9 scale; these means did not statistically differ, $t(159) = 1.85$, $p = .07$. Participants also rated the target as mildly overreacting in terms of both anger (M = .55, SD = 1.00; $t(160) = 7.01$, $p < .001$) and disgust (M = .25, SD = .90; $t(160) = 3.47$, $p = .001$), with the anger overreaction stronger than the disgust overreaction, $t(160) = 5.36$, $p < .001$.

Prediction 1: The target's trait anger will be rated higher than his trait disgust, and will remain so even after controlling for any systematic discrepancy between the perceived appropriateness of his anger and disgust reactions.

Before controlling for overreaction, the target's trait anger (M = 5.94, SD = 1.24) was rated higher than his trait disgust (M = 5.57, SD = 1.16; $t(160) = 3.88$, $p < .001$). To examine whether this difference remained significant after controlling for perceived overreaction, we used multilevel regression (HLM 7.0), as all of the measures were within participants. The dependent measure, trait emotion rating, was regressed on emotion type (anger or disgust) and the degree of the protagonist's behavioral overreaction. Not surprisingly, the more participants perceived the target as overreacting in terms of either anger or disgust, the more they rated him as dispositionally inclined to experience that emotion ($B = 0.48$, $p < .001$). Nevertheless, supporting Prediction 1, even with this variable controlled, the type of emotion was still significantly associated with the magnitude of the trait rating ($B = 0.21$, $p = .045$), such that marginal trait anger was still rated higher than marginal trait disgust.

Prediction 2: Ratings of overreaction will fully mediate the positive association between state and trait ratings for disgust, but will not fully mediate this association for anger.

The results of the two mediational models are shown in Figure 1. Participants attributed disgusted behavior to the target's personality in a normatively logical way: after controlling for perceived overreaction, there was no longer a significant association between state and trait disgust, i.e., overreaction fully mediated the state-trait association for disgust. In contrast, participants attributed angry behavior in a biased way: after controlling for perceived overreaction, there was still a significant (albeit weaker) association between state and trait anger, i.e., overreaction only partially mediated the state-trait association for anger. Thus, Prediction 2 was also supported.

Sobel tests indicated that overreaction mediated a significant amount of the total effect for both anger ($z = 3.20$, $p = .001$) and disgust ($z = 4.44$, $p < .001$). This means that controlling for overreaction significantly reduced the association between state and trait rating for both emotions, which was an expected result based on normatively logical discounting. The difference was that for anger, the residual direct association between state and trait was still strong after accounting for angry overreaction, whereas for disgust, there was no longer any direct association after controlling for disgust overreaction.

Discussion

Supporting Hypothesis 1, participants attributed more enduring anger than disgust to a fictitious man, even after we accounted for systematic differences between their perceptions of his state anger and disgust. Supporting Hypothesis 2, participants made increasingly dispositional attributions as the perceived level of anger the individual displayed increased, regardless of how justified his emotional reaction was; the same was not true of disgust. This

attributional pattern is consonant with an evolved error management bias sensitive to shifts in cost asymmetries, rather than with normative logic.

As noted earlier, in the absence of objective baselines, tests of error management hypotheses rely on relative points of comparison in confirming the existence of predicted biases. Disgust, a negative emotion that resembles anger in a number of respects (Smith and Ellsworth, 1985), performed this role in Study 1. However, in order to determine that the supportive evidence obtained in Study 1 was not an artifact of one particular comparison emotion, it is important to subject the hypotheses to additional tests using a different point of comparison. We therefore used fear, which differs greatly from both anger and disgust (Smith and Ellsworth, 1985), as the negative emotion control in Study 2. More generally, a main effect comparison of scale ratings of trait anger and any other negative emotion might be difficult to interpret. This underscores the importance of introducing additional manipulations that we hypothesized would affect the ratings of trait anger but not of other negative emotions, a key piece of our framework that was not explored in Study 1.

Study 2

Study 2 tested Hypothesis 1 using a new comparison emotion (fear) and Hypothesis 2 by manipulating the danger posed by the target to the perceiver. Participants viewed photographs of faces that varied in sex and eye gaze direction and expressed either anger or fear. Participants rated the images on trait and state anger or fear, respectively. This allowed us to test a number of predictions. Hypothesis 1 predicts that, collapsed across all other manipulations, dispositional anger ratings will be higher than dispositional fear ratings. Moreover, as in Study 1, this difference should remain significant even after accounting for the perceived strength of the anger and fear expressions. Controlling for this source of normatively logical inferences about the

targets' emotional traits ensures that any remaining difference between the ratings of trait anger and fear constitutes a bias.

Prediction 1A: Across conditions, dispositional anger ratings will be higher than dispositional fear ratings, and will remain so even after controlling for any systematic differences in the state intensity of the anger and fear images.

This prediction, directly paralleling Prediction 1 in Study 1, can be further qualified using the logic of Hypothesis 2: because, on average, men are more dangerous than women, underestimating a man's propensity to experience anger will have been especially costly over evolutionary history; the same would not have been true of fear. Accordingly, all else equal, the difference between dispositional anger and dispositional fear ratings should be especially pronounced for male targets.

Prediction 1B: The difference between dispositional anger and fear ratings will be higher for male than for female targets, and will remain so even after controlling for any systematic differences in the state intensity of the anger and fear images.

Eye gaze in the images was either direct (looking at the participant) or averted (looking to the side). An emotional expression coupled with direct gaze usually signals that the emotion is directed *toward* the perceiver. In the case of anger, direct gaze indicates that the target might have harmful intentions toward the perceiver – a possibility that is hazardous for the perceiver to ignore both in the moment and in future interactions. In such circumstances, it would be especially costly for the perceiver to underestimate the target's anger-proneness. Therefore, Hypothesis 2 predicts that perceivers will be especially biased toward assuming direct-gaze targets are anger-prone, thereby motivating the perceiver to avoid this threatening person in the future.

Target's sex and eye gaze should exhibit synergistic effects on judgments of dispositional anger, as a potentially dangerous male individual signaling that he is angry at the observer (via direct gaze) presents an especially potent combination of danger cues. Moreover, these two factors should interact with the perceiver's vulnerability to assault. Because women are less physically formidable than men, they should be especially sensitive to interpersonal cues of danger.

Prediction 2A: There will be a four-way interaction between emotion condition (anger or fear), the participant's sex, the target's sex, and the target's eye gaze, such that, to a greater extent than male participants, female participants will rate male targets expressing anger with direct gaze as more predisposed toward anger than male targets expressing anger with averted gaze. This contrast will not be significant in the fear condition.

More generally, because natural selection weighs the benefits of precaution against the costs thereof, psychological adaptations that serve to protect against the fitness costs of violence can be expected to calibrate to the individual's susceptibility to aggression (cf. Snyder et al., 2011). Two predictions that test Hypothesis 2 follow from this logic:

Prediction 2B: There will be a four-way interaction between emotion condition (anger or fear), the participant's self-perceived vulnerability, the target's sex, and the target's eye gaze, such that, to a greater extent than less vulnerable individuals, more vulnerable individuals will rate male targets expressing anger with direct gaze as more predisposed toward anger than male targets expressing anger with averted gaze. This contrast will not be significant in the fear condition.

Methods

Participants. Three hundred seventy-one U.S. participants responded to a posting advertising the online study of "perceptions of individuals" on Mechanical Turk, an increasingly popular website for questionnaire-based research (Buhrmester, Kwang, & Gosling, in press). Participants were paid $0.20 for 10 minutes of their time. Software settings prevented repeat participation from any given computer.

Of the 371 participants, 37 took either less than 5 or longer than 30 minutes to complete the survey. As outliers, these participants were excluded prior to analysis due to concerns about the validity of their responses, leaving 335 participants (189 women, 139 men, and 7 who did not specify their sex). The anger condition (N = 150) was run in its entirety approximately three months before the fear condition (N = 185), with identical sampling, payment, and advertising procedures for the two conditions. The average participant was 35.4 years old (*SD* = 12.9), and 76% of the participants were Caucasian.

Stimuli. Images of faces were selected from the standardized NimStim face set (Tottenham et al., 2009), which contains angry, fearful, and neutral faces posed by the same individuals. Four female and four male targets were selected on the basis of having the most readily identifiable anger expressions in pre-testing conducted by the creators of the face set. The same eight stimulus individuals were later used in the fear condition.

The original face set includes only images with direct gaze. To create averted gaze, angry, fearful, and neutral images were digitally altered (Photoshop Elements v. 8, 2009) by moving the irises and pupils to the eye's right side. These images and the unaltered images were then duplicated and flipped along the Y-axis for counterbalancing. Thus, participants saw one of four types of images: direct-gaze original, direct-gaze flipped, averted-gaze right, and averted-

gaze left (averted right flipped). In all analyses, the two direct-gaze conditions were collapsed into one condition, as were the two averted-gaze conditions.

In order to conduct a "frame-matching" perceptual task (described below), the expressions participants saw in the study were not the original angry or fearful images. Rather, each participant saw an image that was blended between the target's angry or fearful expression and the target's neutral expression. Blends were made using the website www.faceresearch.org in 10% intervals ranging from 50% to 110% of the angry or fearful expression.

Design and Measures. The design of the study was 2 (angry or fearful faces: between-subjects) x 2 (direct or averted gaze: between-subjects) x 2 (target sex: within subjects). In order to avoid arousing suspicion that these were the primary manipulations in the study, emotion and gaze were varied only between subjects. Each participant thus saw and rated each of the eight target individuals' images in randomized order, all of which were either angry or fearful, and all of which displayed either direct or averted gaze. All measures and tasks were completed for each target individual before the participant saw an image of the next target. Figure 2 presents a sample set of the eight faces that participants saw (see Appendix B for a sample trial).

Image ratings. For each of the eight targets, participants initially saw a single image. The degree of anger or fear in the image was randomly selected by the survey software to be either 70%, 80%, or 90% of the original angry or fearful expression. While the image was on the screen, participants answered three questions about using a 1-9 scale: 1) "How angry (scared) does the person look in this picture?"; 2) "Compared to the average person, how *often* do you think this person becomes angry (scared) in real life?"; and 3) "Compared to the average person, how *easily* do you think this person becomes angry (scared) in real life?" Question 1 was an

explicit measure of perceived current state level of anger or fear. The average of Questions 2 and 3 comprised this study's primary dependent measure of enduring trait level of anger or fear.

Frame-matching task. On the next screen, participants saw an array of image blends of the same target individual which varied in the neutrality of their expression. Participants were asked to select the image that matched the one they rated on the previous screen. Participants' error on this task was intended to be an implicit measure of their perception of state emotion in the images. In practice, this exploratory measure did not correlate with the state or trait measures, indicating that it was not a face-valid measure of perceived state emotion. It was therefore not used in further analyses.

Demographics. Following the image tasks, participants reported their sex, age, and ethnicity. To assess self-perceived vulnerability, participants then completed the Belief in a Dangerous World scale (BDW; Altemeyer, 1998), wherein a 5-point scale is used to indicate agreement with statements regarding the extent to which others are violent and antisocial, and life is full of hazards.

Results

Prediction 1A: Across conditions, dispositional anger ratings will be higher than dispositional fear ratings, and will remain so even after controlling for any systematic differences in the state intensity of the anger and fear images.

Collapsing across conditions, a one-way ANOVA was conducted predicting the trait rating (averaged across all eight targets) from the emotion condition (anger or fear). There was a significant effect of emotion condition ($F(1, 332) = 45.1$, $p < .001$), such that targets were judged to be more dispositionally inclined toward anger ($M = 5.32$) than toward fear ($M = 4.73$). Next, a similar ANCOVA analysis was conducted after adding in averaged state emotion ratings as a

continuous covariate. The state rating measure was higher for the anger (M = 5.09, SD = 2.17) than the fear images (M = 4.53, SD = 2.37; $t(333) = 4.74$, $p < .001$), and, as expected, it was positively associated with trait ratings, $F(1, 331) = 189$, $p < .001$. Controlling for the state rating moved the marginal means for trait anger (M = 5.18) and trait fear (M = 4.84) closer to each other than the observed means. However, this difference remained highly significant ($F(1, 331) = 296$, $p < .001$), supporting Prediction 1A.

Prediction 1B: The difference between dispositional anger and fear ratings will be higher for male than for female targets, and will remain so even after controlling for any systematic differences in the state intensity of the anger and fear images.

First, without controlling for the state ratings, we established that trait anger was rated higher than trait fear for both the male and female targets. For the analyses testing Prediction 1B, each participant's trait ratings were averaged for the four female targets and the four male targets. Without a covariate, trait anger was rated higher than trait fear for both female (*anger M* = 5.29, *fear M* = 5.06, $F(1, 332) = 4.58$, $p = .033$) and male images (*anger M* = 5.36, *fear M* = 4.39, $F(1, 332) = 7856$, $p < .001$).

Next, the state rating covariate was added to the models. With the covariate included, trait anger was no longer rated higher than trait fear for female targets, $F(1, 331) = .106$, $p = .745$. However, consonant with predictions, for male targets, trait anger remained higher than trait fear, $F(1, 331) = 64.7$, $p < .001$ (see Figure 3). Multilevel modeling confirmed that these results were significantly different for male and female targets. An analysis was run with trait rating as the dependent measure and emotion condition (L2; fear = 0, anger = 1), target's sex (L1), and state rating (L1) as predictors. The cross-level interaction of emotion condition X target's sex was highly significant ($B = .75$, $p < .001$). Hence, these analyses qualified the results

under Prediction 1A as not only being stronger for male targets, but as being true *only* for male targets. Therefore, Prediction 1B was supported.

Prediction 2A: There will be a four-way interaction between emotion condition (anger or fear), the participant's sex, the target's sex, and the target's eye gaze, such that, to a greater extent than male participants, female participants will rate male targets expressing anger with direct gaze as more predisposed toward anger than male targets expressing anger with averted gaze. This contrast will not be significant in the fear condition.

A 2x2x2x2 repeated-measures ANOVA was conducted to examine the effects of the manipulations. The dependent measure once again consisted of trait ratings averaged across the four same-sex targets. Emotion condition (anger or fear), gaze condition (direct or averted) and participant's sex were between-subjects variables, and target's sex was the repeated measure within participants.

There was again a main effect of emotion condition, such that targets were judged to be more predisposed toward anger than toward fear. There was also a significant main effect of the target's sex ($F(1, 319) = 25.9$, $p < .001$), which was qualified by a two-way interaction of target's sex and emotion condition, $F(1, 319) = 43.1$, $p < .001$. Post-hoc analyses showed that female targets were judged to be more dispositionally fearful than male targets ($p < .001$), whereas there was no difference between female and male targets' dispositional inclination toward anger ($p = .32$) (see Figure 3).

The 4-way interaction was not significant, $F(1, 319) = 1.15$, $p = .284$. However, there was a significant 3-way interaction between gaze, target's sex, and participant's sex ($F(1, 319) = 5.31$, $p = .022$); post-hoc analyses revealed no significant contrast pairings (all ps > .17). To clarify this result, a similar 2 (gaze: direct or averted) x 2 (participant's sex) x 2 (target's sex)

repeated-measures ANOVA was run separately for the anger and fear conditions. Importantly for Prediction 2A, within the anger condition, the same 3-way interaction between gaze, target's sex, and participant's sex was still significant, $F(1, 144) = 5.73$, $p = .018$. Post-hoc analyses revealed that female participants judged male targets to be more dispositionally angry with direct gaze than with averted gaze ($p = .047$). No other contrasts within this 3-way interaction were significant ($ps > .30$), and the entire 3-way interaction was not significant in the fear condition, $F(1, 175) = .783$, $p = .38$ (see Figure 4). Thus, Prediction 2A was mostly but not completely supported: as predicted, the critical 3-way interaction emerged for anger but not for fear, but the non-significance of the 4-way interaction suggests that this difference in patterns between anger and fear was not itself significant.

Prediction 2B: There will be a four-way interaction between emotion condition (anger or fear), the participant's self-perceived vulnerability, the target's sex, and the target's eye gaze, such that, to a greater extent than less vulnerable individuals, more vulnerable individuals will rate male targets expressing anger with direct gaze as more predisposed toward anger than male targets expressing anger with averted gaze. This contrast will not be significant in the fear condition.

To test this prediction, BDW was split at the median and substituted for participant sex into the earlier repeated-measures ANOVA. As before, the other three factors were emotion condition, gaze, and target sex. The 4-way interaction of all of the predictors was significant, $F(1, 318) = 5.33$, $p = .022$. Participants who were high in BDW and rated angry male faces provided higher ratings for trait anger with direct gaze than with averted gaze, $F(1, 318) = 4.29$, $p = .039$, $d = .53$. In contrast, this was not the case for participants who were low in BDW, $F(1, 318) = .011$, $p = .915$. This was also not the case for fear – indeed, there was a marginal opposite

trend for participants high in BDW to rate direct-gaze male fear faces as *less* dispositionally fearful than averted-gaze faces, $F(1, 318) = 2.97$, $p = .086$. Besides these, no other simple contrasts in the model were significant (ps > .19). Therefore, Prediction 2B was supported (see Figure 5).

Further analyses showed that the above result was driven largely by female participants. The 4-way interaction was significant for female participants, $F(1, 180) = 4.28$, $p = .04$), but not for male participants, $F(1, 130) = .93$, $p = .34$). Women who were high in BDW and rated angry male faces provided higher ratings for trait anger with direct gaze ($M = 5.61$, $SD = 1.05$) than with averted gaze ($M = 4.86$, $SD = .83$; $F(1, 180) = 6.21$, $p = .014$, $d = .79$); the equivalent contrast was not significant for men, $F(1, 130) = .11$, $p = .74$.

To verify these results using the original continuous BDW variable, two multilevel analyses similar to the above were conducted separately within anger and fear. These analyses used multilevel regression because repeated-measures ANOVA does not allow for the testing of interactions between continuous and dichotomous measures (e.g., BDW score X gaze condition). The dependent measure was trait emotion (either anger or fear) and the predictors were gaze (L2; dichotomous), BDW (L2; continuous), target's sex (L1; dichotomous), and all of their 2-way and 3-way interactions. For anger, the critical cross-level interaction of target's sex X gaze X BDW was significant, $B = .46$, $p = .049$; this was not the case for fear ($B = -.38$, p = .098). Once again splitting the sample by participant sex, this 3-way interaction for anger remained significant for women ($B = .72$, $p = .006$) but not for men ($B = -.32$, $p = .46$).

Discussion

Study 2 supported both primary hypotheses. Predictions 1A and 1B conceptually replicated and further elaborated on the results of Study 1, providing additional support for

Hypothesis 1. As in Study 1, across manipulation conditions, targets were judged to be more prone to becoming angry than to feeling a control negative emotion (in this case, fear). This remained true even after controlling for the images' perceived state emotional intensity. Moreover, this result was newly qualified: the effect was entirely driven by male targets. Female targets were considered more predisposed to anger than to fear only to the extent that participants perceived their anger expressions to be more intense than their fear expressions; once intensity was held constant, this effect disappeared. This result for female targets accords with normative attributional logic (Jones & Davis, 1965; Kelley, 1972). In contrast, male targets were judged to be more predisposed to anger than to fear above and beyond any rational indications that this was the case from the images. This result reveals an attribution process that is not rational in the classic sense, but rather is adaptively rational in its bias toward the error that has been consistently less costly over evolutionary history.

Figure 3 shows that the significant interaction that supported Prediction 1B was driven by lower ratings of women's marginal trait fear, relative to men. Ideal support for Hypothesis 2 would have emerged if the result was instead driven by higher ratings of men's marginal trait anger, relative to women. Nevertheless, the lack of such a difference is not problematic for the hypothesis because direct comparisons between the male and female targets are difficult to interpret in this design, given that their images could have systematically differed in a number of ways beyond sex. For instance, the target men's facial morphology or less drastic posed emotions, relative to the women's, might have led to low overall trait emotion ratings for the men. Direct comparisons of anger and fear *within* target sex are not problematic, in contrast, because the men's and women's fearful images are natural controls for their own angry images

in terms of morphology and skill in posing emotions. Prediction 1B was therefore based on such within-target-sex comparisons.

Men are stereotyped as being easily angered (Hess, Adams, & Kleck, 2005) and women as easily frightened (Hess, Blairy, & Kleck, 2000). Could these stereotypes account for the results that we interpret as supporting Prediction 1B? Male, but not female, targets were judged to have higher marginal trait anger than trait fear after controlling for the state emotional intensity in the images. One could argue that this result, rather than being elicited directly by viewing vivid emotional images, simply reflects participants' pre-existing notion that men are more predisposed to anger than to fear. However, if stereotyping were the best explanation, the pattern in Figure 3 would arguably show that men have higher perceived trait anger than women, which they do not. Moreover, it would be difficult to fit a stereotyping explanation to the more complex findings in this study that supported Hypothesis 2, so an error management bias remains the more parsimonious explanation for Prediction 1B.

Predictions 2A and 2B, testing Hypothesis 2, were largely supported. Results confirmed the idea that the dispositional attribution of angry expressions is increased by a combination of the target's strong danger cues (direct gaze, male target) and the participant's high vulnerability (if the participant is female, and especially so if she strongly believes that the world is dangerous). These highly specific results affirmed the notion that the estimation of trait anger involves a bias rooted in adaptive error management. In reality, an individual's enduring personality does not change with shifting gaze or when examined by a more vulnerable observer. However, participants' ratings of his personality *did* change based on these factors, pointing to a true bias. As predicted, this result was specific to the attribution of anger: fear attributions were not affected by gaze or target's sex.

General Discussion

Our findings constitute preliminary evidence that the estimation of trait anger is biased in an adaptively rational way. Study 1 showed that angry behaviors (especially when they are strong) are interpreted as reflecting on the actor's personality more than are disgusted behaviors, even when both are equally justified. Study 2 replicated this finding with a different comparison emotion, fear, and qualified the finding by showing that the trait anger overestimation bias is enhanced by combinations of factors associated with the target's ability and likelihood of aggressing against the observer. The complex interactions in Study 2 precisely supported the predictions, and are difficult to explain without appealing to error management logic. These findings are in line with recent research (Galperin et al., under review) which found that individuals in photos are judged to have higher trait anger when they are in possession of a potentially dangerous household object, relative to a harmless object.

Theoretical Implications

Cognitive Versus Behavioral Biases. The current research adds to the long list of documented cognitive biases rooted in error management (Galperin & Haselton, in press; Haselton & Galperin, in press). Some researchers have argued that such biases are unnecessary (and therefore unlikely to exist) because adaptive behavior, not cognition, is what ultimately affects fitness, and people can theoretically behave in adaptively biased ways without having corresponding cognitive biases (McKay & Dennett, 2009; McKay & Efferson, 2010). For instance, a woman could decide to avoid a man who has expressed anger toward her in the past without being biased in her estimation of his trait anger. We agree that the ultimate currency for fitness is adaptive behavior, but whether and when such behavior is motivated by biased cognition remains an empirical question. The corpus of evidence to which our results contribute

reveals cognitive biases in a variety of judgment domains (Haselton et al., 2009; Haselton & Buss, 2009).

Ingroups and Outgroups. The consequences of dealing with an anger-prone individual were not always negative for ancestral humans; rather, they depended on whether the individual was part of one's ingroup. In many contexts, a propensity for aggression would have been a highly valued quality in allies, as long as it was directed toward outgroups and facilitated successful intergroup competition or defense. The tests conducted in the current study were not intended to apply to allies in situations of intergroup conflict and, indeed, our findings suggest that participants implicitly treated unfamiliar individuals as non-allies by default. In ancestral populations, markers indicating sharing of cultural identity or strategic objectives likely facilitated cooperation and coordination between unfamiliar sets of individuals who were primarily acquainted only with their immediate neighbors (Boyd & Richerson, 2009; Henrich, 2004; Kurzban, Tooby, & Cosmides, 2001). In the absence of such salient cues, it will generally have been adaptive to evaluate strangers with caution, exactly as our participants did.

The Correspondence Bias and Negativity Bias. The Correspondence Bias (Gilbert & Malone, 1995; Ross & Nisbett, 1991) occurs whenever, to a logically unwarranted extent, people attribute others' behaviors to the target's enduring traits rather than to the situation. This bias has been documented across many judgment domains, including attitudes, moral character, competence, and emotionality. Researchers have typically focused on examining the mechanisms via which this bias operates consistently across domains, rather than examining its ultimate explanations (although see Andrews, 2001) or testing theoretically driven hypotheses about how it might differ between domains. While our results can conceivably be classed as an instance of the Correspondence Bias, our research speaks directly to the latter issues, as domain-general or

purely proximate explanations of the Correspondence Bias do not predict that angry behaviors will be attributed to enduring traits to a greater extent than disgusted or fearful behaviors.

An overarching pattern characterizing both our results and a majority of findings regarding the Correspondence Bias is that, when people evaluate others, bad looms larger than good (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Rozin & Royzman, 2001; Ybarra, 2002). This "negativity bias" facilitates adaptively attending to and addressing threats (Rozin & Royzman, 2001), and is manifested in people's tendency to attribute negative or socially undesirable behaviors especially strongly to enduring traits (e.g., Reeder & Spores, 1983; Ybarra, 2002). While the current results for anger (a socially undesirable trait) are consistent with this phenomenon, they also move beyond it by illustrating the adaptively rational ways in which context can affect the degree of the bias for anger but not for other negative emotions.

Future Directions

Future laboratory research could refine the methodology used in the current project. First, participants' vulnerability could be directly manipulated, perhaps via fear priming (e.g., Maner et al., 2005), rather than simply measured as in Study 2. Second, future research could examine the effects of purely situational variables external to the target – ones that provide zero information about the target's personality – on trait ratings. Although eye gaze is an approximation of such a variable, it might conceivably still be informative about the target's propensity to look directly into people's eyes, which might be informative of the target's other traits. The documentation of such purely situational effects would provide even stronger evidence for the existence an attributional bias. However, it is difficult to devise a situational manipulation that is completely unrelated to the target's disposition because people are likely aware that individuals assort into specific situations based on their dispositions (Gilbert & Malone, 1995; Swann, 1984).

Asking participants to make explicit dispositional judgments of individuals in vignettes or photos has debatable ecological validity. In the real world, people do not explicitly rate others' personalities immediately after first encountering them. Indeed, the most pressing adaptive problem is to infer the other's immediate state and intentions, rather than to judge how the other might act over prolonged periods. Nevertheless, the latter has great utility, and prior research shows that people do quickly and spontaneously infer traits in others even when not explicitly asked to do so (e.g., Winter & Uleman, 1984). To further alleviate concerns about ecological validity, future research could take place in more realistic settings. In the lab, observers could have face-to-face interactions with and rate confederates who vary in formidability and emotional behavior. Outside the lab, participants could be recruited in or placed into real-world settings that inherently vary in danger, such as subway stations at night versus during the day, and asked to rate vignettes or photos. Such a study design would eliminate the aforementioned confound between context and targets' personality traits, because the rater's real situational setting could not possibly be informative of fictitious targets' traits.

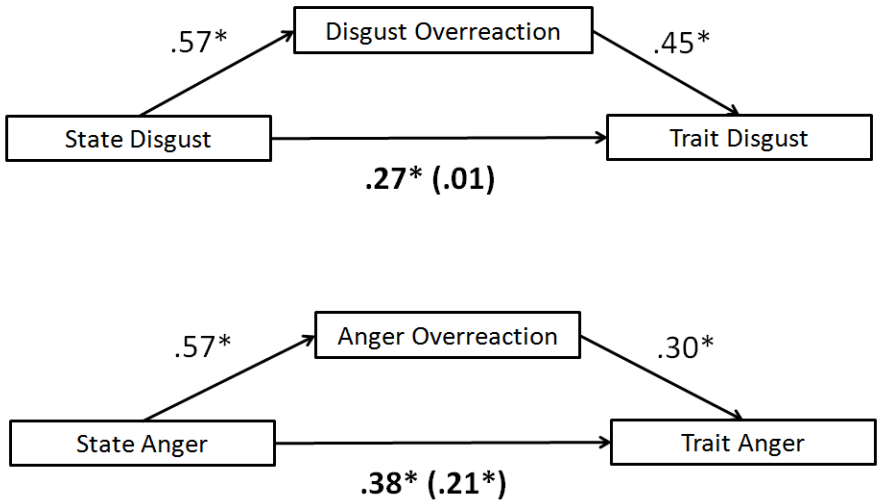
Theoretical and Practical Implications

This research illustrates the value of combining the fundamental motives approach (Kenrick et al., 2010) with the adaptive rationality approach (Haselton et al., 2009). Although we did not directly manipulate people's self-protection motive, we measured it (via the BDW scale) and arguably activated it by presenting vignettes and faces of angry individuals. When we did so, participants responded to these stimuli with biased trait anger ratings that would have minimized ancestral fitness costs in an adaptively rational manner.

This research has practical implications. Because people lean toward seeing the bad in others, they are likely to avoid interacting or forming relationships with individuals who made a

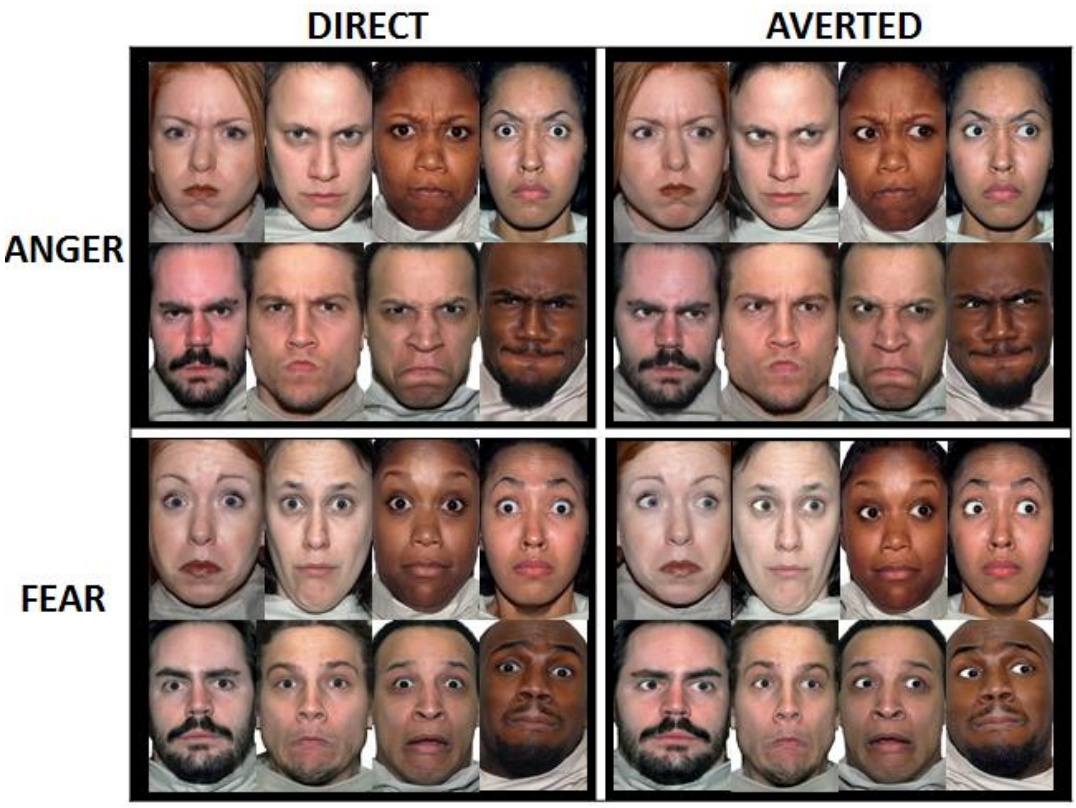
bad first impression even if they were situationally induced to behave this way. The specific case of the overestimation of trait anger suggests that people may avoid new acquaintances after a single instance of angry behavior, even if it was justified in the eyes of the perceiver. Moreover, this is probably especially likely to occur when the target is formidable (e.g., a muscular man) and when the observer is either chronically vulnerable or feels temporarily unsafe in the surrounding environment. Although these patterns were adaptive in the social environments of our ancestors, modern humans live in a much safer world than in the past (Pinker, 2011). Hence, the biased overestimation of trait anger might unfortunately lead people to mistakenly form negative impressions, eschewing relationships with others who might otherwise have become valued social partners.

Figure 1. Standardized regression coefficients for the relationship between ratings of state and trait emotion as mediated by perceived overreaction.



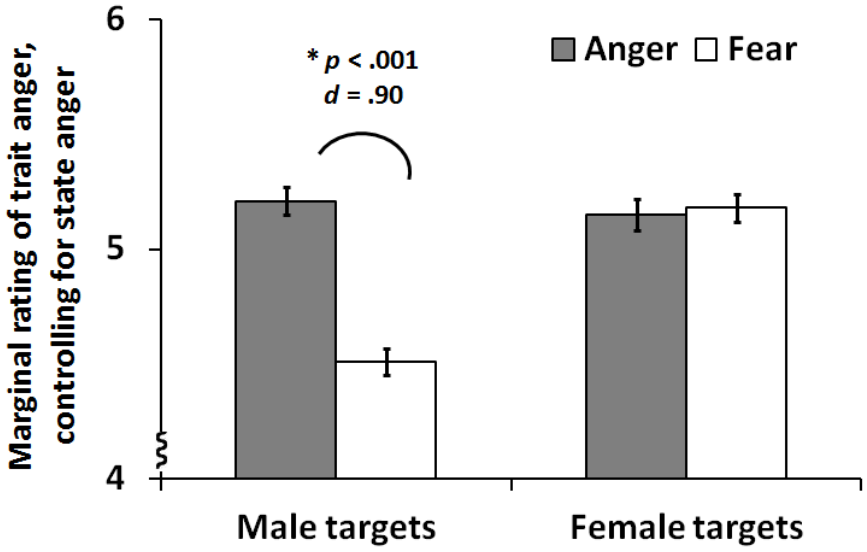
Note. The standardized regression coefficient between state and trait emotion controlling for overreaction is in parentheses.
* $p < .05$

Figure 2. Face stimuli in Study 2.



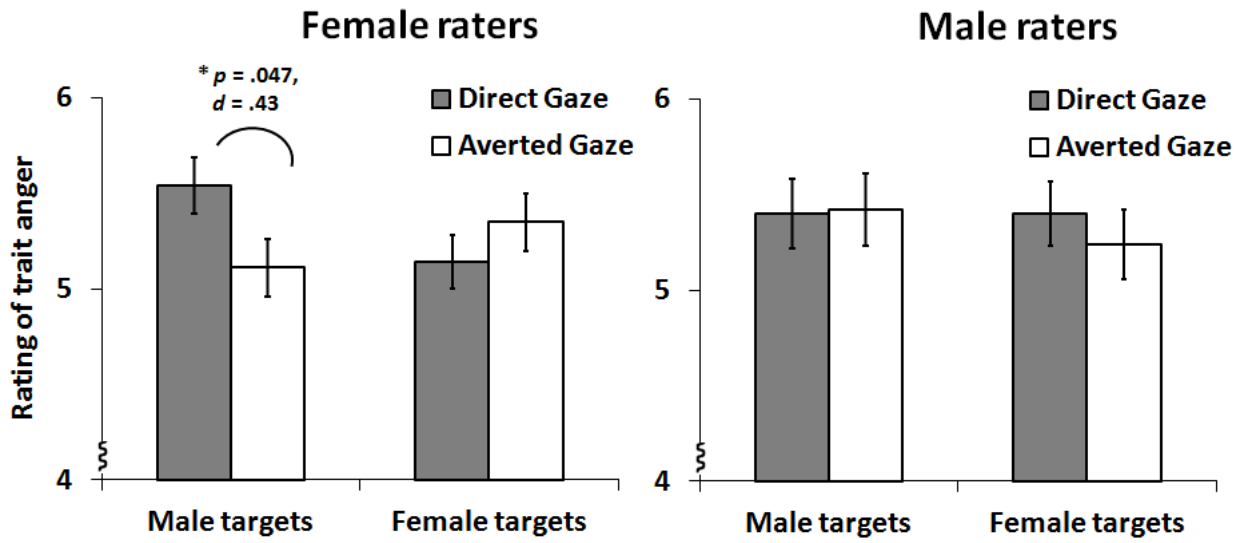
Note. Each participant saw and rated all eight faces (one at a time) that are depicted in one of the four quadrants. All faces depicted here are 90% blends of neutral and angry or neutral and fearful – that is, 10% less emotional than the original images in the Nimstim set. In the actual study, each target image was randomly selected to look 70%, 80%, or 90% angry or fearful.

Figure 3. The effects of targets' sex on participants' dispositional anger and fear ratings, controlling for participants' explicit ratings of state emotional intensity in the images in Study 2.



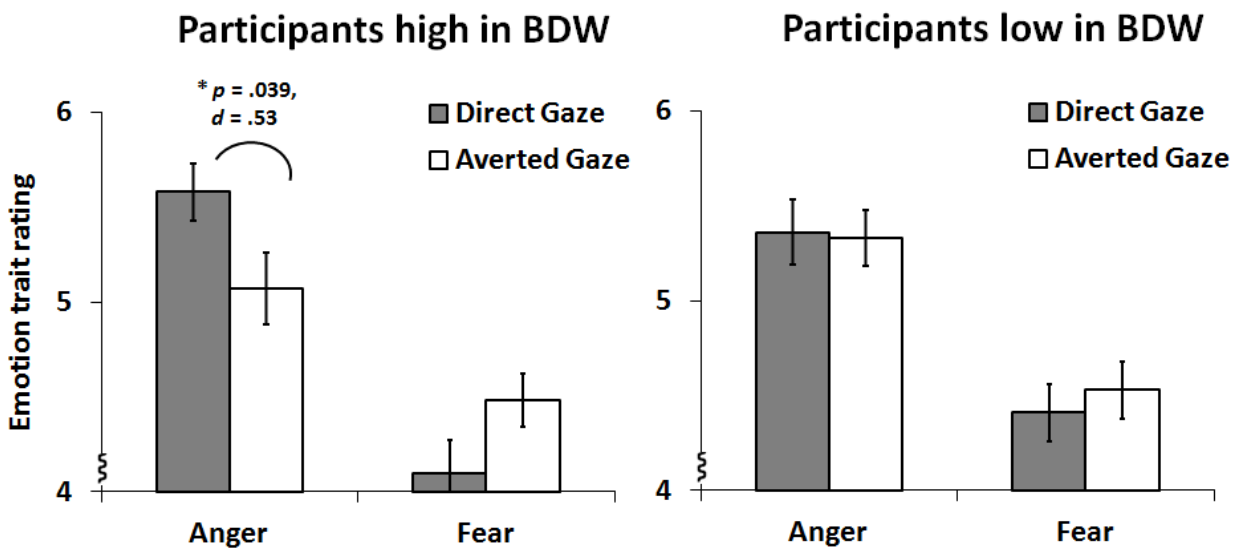
Note. The two-way cross-level interaction of emotion condition X target's sex was highly significant ($B = .75$, $p < .001$). The simple contrast between men's ratings of dispositional anger and dispositional fear was also significant ($p < .001$, $d = .90$).

Figure 4. The effects of gaze, target's sex, and participant's sex on participants' ratings of targets' predisposition toward becoming angry in Study 2.



Note. The 3-way interaction of gaze, target's sex, and participant's sex was significant for anger (p = .022); the same 3-way interaction for fear (not pictured here) was not significant (p = .38). The leftmost contrast between women rating male targets with direct vs. averted gaze was significant (p = .047, d = .43). No other contrast in this figure was significant (ps > .30).

Figure 5. The joint effects of participants' Belief in a Dangerous World and gaze direction on ratings of male targets' dispositional anger and fear in Study 2.



Note. The 4-way interaction between emotion condition, gaze, target sex, and BDW was significant (p = .022). The leftmost contrast between high-BDW participants rating angry male targets with direct vs. averted gaze was significant (p = .039). No other simple contrast in this figure was significant (ps > .086), and no simple contrasts were significant for participants rating female targets (ps > .19; not pictured here).

Appendix A:
Vignettes Used in Study 1

Note. Albert's angry reactions are bolded, and his disgusted reactions are underlined.

VIGNETTE 1 (WEAK)

Albert was out with several friends, having dessert at a restaurant. He briefly left the table to go to the bathroom, and when he came back, he saw that one of his friends had put ketchup on his ice cream, which Albert had not finished eating. **Seeing this, Albert did not look very happy.** To make up for his prank, the friend who did it proceeded to eat the ice cream with the ketchup on it. At this point, Albert became somewhat amused but made a face and said "That's nasty, man."

VIGNETTE 2 (STRONG)

Albert's roommate managed to clog their toilet and proceeded to flush it multiple times, hoping it would unclog itself. However, the toilet ended up overflowing, sending dirty water all over the bathroom floor. As this was happening, the roommate ran out of the bathroom, clearly panicked. Albert came over to see what all the commotion was about, and as he realized what happened, he covered his nose with his shirt and quickly closed the bathroom door. The roommate seemed reluctant to start cleaning up and suggested they just leave it alone for several hours until maintenance gets there. **Incredulous, Albert yelled at the roommate, "Dude, you're the one that made this mess, so it's your job to clean it up!"**

VIGNETTE 3 (WEAK)

Albert was at a party when a drunk, rowdy guy bumped into him and spilled beer all over his shirt. The guy was clearly not being careful or paying attention to his surroundings. **Albert yelled, "Hey, watch where you're going next time!" as the other guy started apologizing.** While cleaning himself up, Albert complained to one of his friends that his shirt smells awful, and eventually he went back to his dorm room to change.

VIGNETTE 4 (STRONG):

Albert ordered a chicken sandwich at a fast food restaurant. When he bit into it, he noticed that it was unusually chewy, and upon closer examination, it turned out that the chicken was almost raw on the inside. Recoiling, Albert spit out the sandwich, immediately grabbed his soda and gulped down half of it. He went back to the counter, explained what happened, and asked to speak to the manager. However, the manager seemed like he didn't care and even remarked that "a little bit of undercooked chicken won't kill you." **Albert got red in the face, raised his voice and told the manager that with that attitude, his restaurant will go out of business in no time. Albert then stormed out of the restaurant.**

Appendix B:
Sample Trial in Study 2

Study this image briefly, answer the questions below, and then go on to the next page.

How angry does the person look in this picture?

1	2	3	4	5	6	7	8	9
Not at all angry		Somewhat angry		Moderately angry		Very angry		Extremely angry

Compared to the average person, how <u>often</u> do you think this person becomes angry in real life?

1	2	3	4	5	6	7	8	9
Much less often than average		Somewhat less often than average		About average		Somewhat more often than average		Much more often than average

Compared to the average person, how <u>quickly</u> do you think this person becomes angry in real life?

1	2	3	4	5	6	7	8	9
Much less quickly than average		Somewhat less quickly than average		About average		Somewhat more quickly than average		Much more quickly than average

Which of these images is the same one you just saw?

○ ○ ○ ○ ○ ○ ○

References

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14*, 644-647.

Altemeyer, B. (1998). The other "authoritarian personality." *Advances in Experimental Social Psychology, 30*, 47–92.

Andrews, P. W. (2001). The psychology of social chess and the evolution of attribution mechanisms: Explaining the fundamental attribution error. *Evolution and Human Behavior, 22*, 11-29.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364,* 3281–3288.

Buhrmester, M. D., Kwang, T., & Gosling, S. D. (in press). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science.*

Fessler, D. M. T. (2010). Madmen: An evolutionary perspective on anger and men's violent responses to transgression. In M. Potegal, G. Stemmler, & C.D. Spielberger (Eds.), *Handbook of anger: Constituent and concomitant biological, psychological, and social processes* (pp. 361-381). Springer.

Frank, R. H. (1988.) *Passions within reason: The strategic role of the emotions.* New York: Norton.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.

Galperin, A., Fessler, D. M. T., Haselton, M. G., Johnson, K. L., Holbrook, C., & Bryant, G. A. (under review). *Anger judgments are informed by affordances for doing harm.*

Galperin, A., & Haselton, M. G. (in press). Error management and the evolution of cognitive bias. To appear in J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social Thinking and Interpersonal Behavior.* New York: Psychology Press.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117,* 21-38.

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences, 16,* 701-713.

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition,* 27, 733-763.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78,* 81-91.

Haselton, M. G., & Buss, D. M. (2009). Error management theory and the evolution of misbeliefs. *Behavioral and Brain Sciences, 32,* 522-523.

Haselton, M. G. & Galperin, A. (in press). Error management in relationships. To appear in J. A. Simpson & L. Campbell (Eds.), *Handbook of Close Relationships*. Oxford University Press.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10,* 47-66.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization, 53,* 3–35.

Hess, U., Adams, R. B., Jr., & Kleck, R. E. (2005). Who may frown and who should smile? Dominance, affiliation, and the display of happiness and anger. *Cognition and Emotion, 19,* 515–536.

Hess, U., Blairy, S., & Kleck, R. E. (2000). The influence of expression intensity, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal Behavior, 24,* 265–283.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). San Diego, CA: Academic Press.

Keeley, L. H. (1996). *War before civilization: The myth of the peaceful savage.* New York: Oxford University Press.

Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-26), Morristown, NJ: General Learning Press.

Kenrick, D. T., Neuberg, S. L., Griskevicius, V., Becker, D. V., Schaller, M. (2010). Goal-driven cognition and functional behavior: The fundamental motives framework. *Current Directions in Psychological Science, 19,* 63-67.

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27,* 313-327.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98,* 15387-15392.

Maner, J. K., Kenrick, D. T., Neuberg, S. L., Becker, D. V., Robertson, T., Hofer, B., et al. (2005). Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology, 88,* 63-78.

McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences, 32,* 493-561.

McKay, R. T., & Efferson, C. (2010). The subtleties of error management. *Evolution and Human Behavior, 31,* 309-319.

Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. New York: Viking.

Reeder, G. D., & Spores, J. M. (1983). The attribution of morality. *Journal of Personality and Social Psychology, 44,* 736 – 745.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). San Diego, CA: Academic Press.

Ross, L., & Nisbett, R. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.

Rozin, P., Haidt, J., McCauley, C., Dunlop, L., & Ashmore, M. (1999). Individual differences in disgust sensitivity: Comparisons and evaluations of paper-and-pencil versus behavioral measures. *Journal of Research in Personality, 33,* 330-351.

Rozin, P., Haidt, J., & McCauley, C. R. (2000). Disgust. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions, 2nd Edition* (pp. 637-653). New York: Guilford Press.

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

Sell, A. (2009). Applying adaptationism to human anger: The recalibrational theory. In P. R. Shaver & M. Mikulincer (Eds.), *Understanding and reducing aggression, violence, and their consequences.* Washington DC: American Psychological Association.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality & Social Psychology*, *48*, 813-838.

Snyder, J. K., Fessler, D. M. T., Tiokhin, L., Frederick, D. A., Lee, S. W., & Navarrete, C. D. (2011). Trade-offs in a dangerous world: Women's fear of crime predicts preferences for aggressive and formidable mates. *Evolution and Human Behavior*, *32*, 127-137.

Swann, W. B. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91*, 457-477.

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A. et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242-249.

Tybur, J. M., Lieberman, D. L., & Griskevicius, V. (2009). Microbes, mating, and morality: Individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology, 29*, 103-122.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*, 237-252.

Ybarra, O. (2002). Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin, 128,* 421-441.