

Error Management and the Evolution of Cognitive Bias

Martie G. Haselton^{1,2,3,4} and Andrew Galperin^{1,3,4}

(these authors contributed equally to this work)

1. UCLA Department of Psychology
2. UCLA Department of Communication Studies
3. UCLA Center for Behavior, Evolution, and Culture
4. UCLA Interdisciplinary Relationships Science Program

2/18/2011

Correspondence to:

Martie G. Haselton
Department of Communication Studies
University of California, Los Angeles
Box 951538
Rolfe Hall, Room 2322
Los Angeles, California 90095
(310) 206-7445; (310) 206-2371 Fax
haselton@ucla.edu

Abstract

Error management theory (EMT) proposes that when the costs of different types of errors are asymmetrical in their fitness consequences, natural selection will create biased cognitive mechanisms that maximize the least costly error (Haselton & Buss, 2000; Haselton & Nettle, 2006). Since the time of its initial publication, EMT has produced dozens of new hypotheses and empirical results characterizing human cognition. With a focus on the last decade of research developments, we summarize evidence of error management biases across a variety of social psychological domains, ranging from perceptions of romantic attraction to social prejudices, cooperative behaviors, and the judgment of personality traits. We then cover emerging theoretical developments, such as the role of context in affecting the magnitude of biases predicted by EMT. We conclude by addressing a recent challenge to the theory – the notion that selection should preserve true beliefs, and therefore is expected only to bias *behavior*, and not *underlying beliefs* (cognition), in order to manage error costs (McKay & Dennett, 2009; McKay & Efferson, 2010). We discuss several hypotheses about why, in order to manage error costs, selection targets *beliefs*.

Error Management and the Evolution of Cognitive Bias

In the 1990s comedy “Dumb and Dumber”, Jim Carrey plays a dim-witted character named Lloyd Christmas. Lloyd is chasing after an attractive woman, Mary, and at one point asks her directly: “What are my chances?” When she replies that his chances are not good – namely, one out of a million – Lloyd pauses, seemingly disappointed. Then gradually, a huge smile comes over his face, and he exclaims, “So you are telling me there is a chance? YEAH!!”

This response is funny because it seems irrationally optimistic, but in fact it illustrates a well-documented phenomenon: men appear to systematically overestimate women’s sexual interest (e.g., Abbey, 1982; Haselton & Buss, 2000). Psychologists have offered a variety of explanations for this bias. Perhaps men are simply not very good at perceiving or remembering women’s nonverbal cues (Farris, Treat, Viken, & McFall, 2008a; Treat, Viken, Kruschke, & McFall, 2010), or, possibly, interacting with women imposes high cognitive load on men and confuses their thinking (Karremans, Verwijmeren, Pronk, & Reitsma, 2009). Another possibility is that social pressure, perhaps transmitted via the popular media, causes men to view women as sexual objects and therefore encourages them to view social interactions through a “sexualized lens” (Harnish, Abbey, & DeBono, 1990). These explanations share in common the implication that something has gone wrong in men’s thinking – some sort of limitation or confusion is to blame.

An alternative view, and the one that is the focus of this chapter, is that men’s bias and a variety of other social cognitive biases can be understood as adaptations produced by evolution. In short, from the perspective of error management theory (EMT), one

expects the evolution of a bias when it minimizes the net fitness cost of errors in judgment and decision making – even if that bias produces more errors than alternative psychological designs. In this chapter, we review EMT and the last decade of research testing hypotheses derived from the theory. The theory, we argue, has been influential in explaining cognitive biases, including men’s sexual overperception. We also discuss the future of EMT with a particular focus on recently emerging themes and challenges.

Error Management Theory

Error management theory (Haselton & Buss, 2000; Haselton & Nettle, 2006) applies to judgments under uncertainty. For instance, humans need to judge whether sticks are snakes and vice versa. We can make either a *false-positive error* (inferring that it is a snake when it is not) or a *false-negative error* (inferring it is not a snake when in fact it is). In making uncertain judgments like this, the costs of committing the two errors are often unequal; in this particular case, a false-negative might lead to being bitten by a snake, whereas a false-positive results only in added caution. This is why we judge sticks to be snakes when walking in the woods but rarely judge snakes to be harmless sticks. Because the costs of errors are asymmetrical, we err on the “safe side” by assuming the worst. Figure 1 depicts this decision making scenario.

Error management theory predicts that biases will evolve in human judgments and decisions whenever the following criteria are met: (a) the decision had recurrent impacts on fitness (reproductive success), (b) the decision is based on uncertain information, and (c) the costs of false-positive and false-negative errors associated with that decision were recurrently asymmetrical over evolutionary time. Detecting dangerous agents, such as snakes, fit all three criteria. This decision problem was present over evolutionary time

Figure 1. *The four possible combinations of actual presence of snakes and beliefs about snakes.*

		Belief	
		Snake Present	Snake Absent
True State of the World	Snake Present	Correct Detection	False Negative (more costly error)
	Snake Absent	False Positive (less costly error)	Correct Rejection

Note. The bias in this scenario errs toward false positives (seeing sticks as snakes) in order to minimize costly false negatives (failing to detect a real snake). In the general model of EMT biases, when the costs of the errors are reversed (e.g., women's inferences of men's commitment, as described in the text), the bias will be toward making false negatives.

and associated with fitness consequences, we had to make judgments about snakes when our eyes were not able to detect snakes with perfect accuracy, and the costs of a false negative error (failing to detect the snake and being bitten) were greater than the costs of false positive errors (unnecessary evasive precautions).

EMT is based on the same logic involved in signal detection theory (Green & Swets, 1966), which examined and formally modeled perceptual judgments under uncertainty. The significant contribution of EMT, beyond insights contained in signal detection theory, is the application of applying a similar logic to recurrent evolutionary costs and benefits that humans have faced throughout evolutionary history. According to EMT, because biases toward making the less costly type of error resulted in reproductive

fitness benefits, relative to other psychological designs, such biases became reliably-developing evolved features of human and animal minds.

Error management biases can be observed in both nonsocial and social domains. For example, in the nonsocial domain of snake detection, there is evidence that fear of snakes is more easily acquired and more difficult to extinguish than fears of other fear-relevant objects (Mineka, 1992). Another non-social error management bias is the tendency to judge the height of a vertical surface as greater when looking from the top rather than the bottom, which may reflect the costs associated with underestimating the danger of falling from a great height (Jackson & Cormack, 2007). We devote the following sections to a summary of key error management biases in social cognition. Over the past decade, EMT has integrated old findings and generated new ones across diverse social psychological domains including mating, avoiding dangerous persons, cooperation, and the attribution of behaviors to underlying attitudes and personality traits.

Biases in Mating

Sexual Overperception by Men. Like males in other species, human males are obligated to invest less in producing offspring and have higher reproductive potential than females (i.e., a man could potentially produce many more offspring over his lifetime than a woman could over hers; Clutton-Brock & Vincent, 1991). In the ancestral past, men could substantially increase their reproductive success by mating more often, whereas by virtue of the necessary time and energetic costs of pregnancy, women generally could not. Throughout evolutionary time, women, more so than men, gained fitness advantages by being selective in choosing partners (Trivers, 1972). Women

benefited from choosing men who either displayed cues of high-fitness genes that could be transmitted to offspring or provided resources that were helpful in raising offspring through their long juvenile period to reproductive maturity (Buss & Schmitt, 1993; Pillsworth & Haselton, 2006b). As a result of sex differences in reproductive opportunities and constraints, men are generally more sexually eager than women (Schmitt et al., 2003; Simpson & Gangestad, 1991) and more willing to engage in opportunistic sexual encounters (Clark & Hatfield, 1989; Li & Kenrick, 2006).

Because men benefited more than women from having a variety of sex partners, there has likely been selection on men for a keen ability to recognize cues of female sexual interest. This judgment, however, is made under considerable uncertainty and is prone to error. An error management perspective predicts that inaccurate judgments should be systematically biased toward overperception – perceiving more sexual interest than there really is. This is because missing a sexual opportunity due to underestimating sexual interest would have been more reproductively costly than overestimating sexual interest and wasting time pursuing a disinterested woman.

Although there have been challenges and competing explanations for this phenomenon (e.g., Farris et al., 2008a), the vast majority of published studies have produced results consistent with the error management account of sexual overperception. The key evidence for this bias, spanning a diversity of assessment methods, is summarized in Table 1 (also see La France, Henningsen, Oates, & Shaw, 2009, for a recent meta-analysis). These results are specific to men perceiving women and unlikely to be simply due to men overstating the sexual interest of all people. Generally, there is little evidence of a directional bias when men judge other men's sexual interest (for a

Table 1. *Key Empirical Evidence for Men's Sexual Overestimation Bias.*

Method	Result	Representative Citations
Face-to-face dyadic interactions (participant ratings)	Men rate women's interest in them as higher than women self-report	Abbey, 1982 Harnish et al., 1990 Henningesen & Henningesen, 2010
Face-to-face dyadic interactions (ratings of third-party observers)	Male observers rate women's interest as higher than do female observers	Abbey, 1982 Saal et al., 1989
Videos of dyadic interactions	Male video watchers rate female targets' interest as higher than do female video watchers	Shotland & Craig, 1988
Photos of dyadic interactions	Male photo viewers rate female participants' interest as higher than do female photo viewers	Abbey & Melby, 1986 Edmondson & Conger, 1995
Written scenarios	Male readers rate higher interest from women's hypothetical dating behaviors than do female readers	Haselton & Buss, 2000
Naturalistic experiences	Women (especially attractive women) report more incidences of men mistaking their friendliness for sexual interest than of men making the reverse mistake	Haselton, 2003
	Men overestimate their female friends' sexual interest	Koenig, Kirkpatrick, & Ketelaar,
Experiments	When primed with romantic thoughts, men perceive sexual interest in photos of neutral female faces (especially attractive faces)	Maner et al., 2005

discussion of this issue, see Abbey, 1982; Haselton & Buss, 2000; and LaFrance et al., 2009). The opposite-sex overperception bias is also not shared by women, who appear to either underperceive men's sexual interest (e.g., Abbey, 1982) or show no clear directional bias (e.g., Haselton & Buss, 2000), depending on the study.

Commitment Underperception by Women. During the energetically expensive period of pregnancy and lactation, women's reproduction requires substantial obligatory investment in offspring (Pillsworth & Haselton, 2006b; Trivers, 1972). These obligations have likely shaped women's preferences for mates who display convincing cues of long-term commitment and thus appear to be willing to provide resources during pregnancy and beyond (Buss & Schmitt, 1993). As with inferences of sexual interest, people must make inferences of commitment under conditions of incomplete information and sometimes outright attempts at deception from the target (Haselton, Buss, Oubaid, & Angleitner, 2005). Error management theory predicts a directional bias for these errors, but in this case, the bias is in women's perceptions of men: EMT predicts that women will underestimate men's commitment intent. This is because the consequences of overestimating commitment – having sex with a man who has little interest in continuing a relationship -- would have been greater than the consequences of underestimating commitment. The former could have resulted in a pregnancy without support from a partner, whereas the latter resulted in a temporary delay in a woman's reproduction while she assessed her partner's commitment.

There has been less work to examine this commitment underperception bias, but several studies have found support for it. Haselton and Buss (2000) asked participants to rate the likelihood that a variety of dating behaviors indicated an interest in a long-term

romantic relationship. The results showed that relative to male raters, female raters inferred less long-term interest (i.e. commitment intent) when men engaged in these behaviors. No such sex difference between raters emerged when rating women's long-term interest from identical behaviors (Haselton & Buss, 2000; also see Haselton et al., 2005, for a discussion of additional evidence). This result was recently replicated with participants in face-to-face interactions (Henningesen & Henningesen, 2010). Male-female stranger dyads participated in a five-minute conversation and afterwards filled out questionnaires about their own and their partner's perceived level of interest in a committed long-term relationship. As predicted by EMT, women underestimated the men's commitment, whereas men were not biased in their estimates of the women's commitment.

Other Mating-Related Biases. Error management theory has also stimulated research on judgment biases in other mating-related domains. For instance, people underestimate their romantic partners' forgiveness after a transgression (Friesen, Fletcher, & Overall, 2005), possibly prompting a more complete mending of the relationship. People overestimate the desirability of same-sex competitors (Hill, 2007), possibly to facilitate keener competition. Relative to women, men are more suspicious about partner infidelity, suggesting that they might overestimate how likely their partners are to be unfaithful (Andrews et al., 2008). This bias might help to protect against the high costs of cuckoldry. In sum, across the many judgments and decisions people make in the courtship context, there are a number of empirically supported biases that arise from the logic of error management theory.

Biases for Avoiding Dangerous People

Prejudice Against Outgroups. It is plausible that one of the greatest threats to life in ancestral environments was other people. Violent intergroup conflict was probably a constant feature throughout human evolution, and in modern environments, from traditional societies to industrialized nations, groups regularly wage deadly wars on one another (Keeley, 1996). Research has shown that people assume that out-group members are less generous and kind (Brewer, 1979) and more hostile and violent (e.g., Quillian & Pager, 2001) than members of one's own racial or ethnic group. Such effects are easily triggered and enhanced by increasing the salience of any kind of distinction between the ingroup and the outgroup (Brewer, 1979).

This bias can be understood from an error management perspective. Inferences about relatively unknown out-group members are uncertain. For ancestral humans, the costly false negative error was to miss aggressive intentions on the part of others. In contrast, the costs of the false positive error – overinferring aggressiveness in members of competing coalitions – were low. This cost asymmetry did not characterize assumptions about in-group members, in which unwarranted inferences of hostility or aggressiveness would have resulted in costly within-coalition conflict.

Avoiding sick people. If other people posed reliable threats of disease throughout evolutionary history (see Schaller & Duncan, 2007), humans could also possess adaptive biases that lead them to feel disgusted by and selectively avoid certain classes of others. And indeed, people require little evidence of illness or contamination to avoid someone, whereas they require much stronger evidence to warrant the inference that someone does not pose the threat of contagion (Kurzban & Leary, 2001; Park, Faulkner, & Schaller,

2003). This evidence can take the form of either propositional knowledge that someone has an unmarked disease, or of physical cues associated with disease. For instance, although people understand, intellectually, that mere contact is insufficient for the transmission of AIDS, they physically distance themselves from AIDS victims, express discomfort with even five minutes of contact, and even express discomfort with the thought that clothes they once wore would be worn by an AIDS victim in the future (Bishop, Alva, Cantu, & Rittiman, 1991; Rozin, Markwith, & Nemeroff, 1992).

Physical cues of disease that may precipitate avoidance include physical abnormalities, lesions, discoloration, impaired motor function, and atypical appearance of body parts. However, there are multiple factors that can cause such physical disfigurements: a swollen hand, for example, could be the result of a contagious infection or the result of an accident (e.g., falling), which is not contagious. Because it is difficult to know the source of a physical anomaly with certainty, error management logic predicts that humans will be biased to avoid phenotypically atypical others as if they are carriers of disease even when they are not. Thus, people may also treat other disabilities or phenotypic anomalies (e.g., obesity) as if they are produced by contagious disease.

Research has directly confirmed this hypothesis. For instance, a study using the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) showed that, after being exposed to a disease prime, participants associated disease with disabled but noncontagious individuals at an implicit level (Park et al., 2003). This result might explain why stigma associated with physical disabilities is so pervasive (Kurzban & Leary, 2001).

Biases for Navigating the Social Environment

The Social Exchange Heuristic. Standard economic principles predict that players in many single-interaction economic games should defect rather than cooperate, because defecting maximizes the player's monetary payoff. The interaction is not repeated and the players are usually anonymous strangers, so there is no incentive to signal cooperativeness for future interactions within the game or for the sake of reputation outside of the game. Yet, cooperation often occurs in these economic games (Camerer & Thaler, 1995; Caporael, Dawes, Orbell, & van der Kragt, 1989; Sally, 1995), and this cooperation is cross-culturally ubiquitous (e.g., Henrich et al., 2001).

From a view of the mind as a rational utility maximizer, this pervasive cooperation is a puzzling phenomenon. Players act as if they expect negative consequences of non-pro-social behavior even when they are aware that, objectively, such consequences are unlikely to follow. Yamagishi and colleagues have proposed that the costs of falsely believing that one can defect without negative consequences are often higher than cooperating when one could safely defect (Yamagishi, Terai, Kiyonari, Mifune, & Kanazawa, 2007). This bias – dubbed the “social exchange heuristic” – can be conceptualized as a combination of error management and an artifact of modern living. Although this is no longer the case in many modern settings, in ancestral environments the probability of repeated encounters would have been high and social reputation effects potent. Therefore, selection may have crafted the social exchange heuristic as an adaptation to this ancestral cost structure.

The social exchange heuristic is well illustrated by the ease with which people feel they are “being watched.” One study (Haley & Fessler, 2005) asked anonymous

strangers to play a series of dictator games (a type of economic game in which one “dictates” what portion of their endowment they will share with another player in the game) on the computer. For some of the participants, the researchers subtly manipulated visual cues by showing stylized eyespots as the computer’s desktop background. The effect of this manipulation was striking: when using a computer displaying eyespots, almost twice as many participants gave money to their partners, compared with the controls. Whether or not they were aware of it, in a sense these participants acted as if they were being “watched.”

Similar error management logic could explain the ubiquity of religious beliefs (Johnson, 2009): the belief that a higher power is observing and judging one’s behavior could be adaptive (and hence lead to the evolution of religious belief), because it promotes cooperation and is associated with the benefits of forgoing immediate self-interest in the service of long-term cooperative benefits.

The Negativity Bias in Attribution. Humans depend greatly on one another, but social partners can inflict costs on each other—for example, through aggression, cheating, or exploitation. Avoiding aggressive or selfish others has been a major selective pressure on human social cognition (e.g., Cosmides, 1989). Thus, it is plausible to expect that many of our initial social judgments are designed to help us avoid these poor social partners (Kurzban and Leary, 2001).

Social partners who have once demonstrated some negative social behavior might or might not be disposed to do the same again in the future. The false negative error is to assume that a person’s behavior is not representative of his or her long-term disposition and thus not take it into account in future interactions. The false positive error is to

assume someone is antisocially disposed because of a behavior that did not in fact represent his or her underlying dispositions, but was instead brought about by a more transient feature of the context. In making the false negative error, one risks becoming involved with a person who could later inflict harm. In turn, the cost of the false positive error might be the avoidance of someone who would in fact be a constructive social partner. This cost might be significant, but often not as high as the cost of being hurt or exploited. Thus, from an error management perspective, it is prudent to initially assume that social partners who have once behaved undesirably are likely to be “repeat offenders.”

Research has shown that there is indeed such a “negativity bias” in attribution (for reviews, see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; and Ybarra, 2002). That is, people view others’ negative behaviors as more diagnostic of their enduring dispositions than positive ones. People are initially more uncertain and skeptical about someone’s positive than their negative qualities (Gilbert, Tafarodi, & Malone, 1993). It takes fewer negative observations about a person to infer a corresponding negative trait than positive observations to infer a corresponding positive trait (Rothbart & Park, 1986). And once people have formed a negative opinion, they are more confident in it (Carlston, 1980) and process subsequent information about that person less attentively (Ybarra & Stephan, 1996) than if they had formed a positive opinion. This leads to negative impressions being difficult to disconfirm. In contrast, positive impressions can sometimes be disconfirmed with only one instance of negative behavior (Reeder & Coover, 1986).

The negativity bias is well illustrated in studies that ask people to judge a person’s moral character. One such study (Reeder & Spores, 1983) demonstrated that people make

attributions about morality in an asymmetric fashion. In the study, perceivers inferred that a person who behaved immorally (by stealing from a charitable fund) had an immoral disposition regardless of situational influences (whether his date encouraged him to steal money or donate money). In contrast, when presented with a moral behavior (in which the target donated money to the fund), perceivers did not always assume the target was a moral person. Instead, their inferences depended on situational cues: when the target was encouraged to donate money and did so, perceivers judged him as less moral than when the target was encouraged to steal and still donated the money. These results suggest that perceivers may be inclined to assume immorality regardless of mitigating circumstances; they make inferences of morality, on the other hand, in a more carefully qualified fashion. Along similar lines, studies have found that one dishonest behavior is enough to overcome a prior impression that a person is honest, but not vice versa (Reeder & Coovert, 1986), and that participants expect immoral acts exhibited by an individual in a particular situation to generalize across situations more easily than moral acts, and more easily than other negative or positive behaviors in non-moral domains (Trafimow, 2001).

Differential Evocation of Bias

Psychological adaptations are responsive to different environmental contexts (Gangestad, Haselton, & Buss, 2006). A critical aspect of error management logic is that ancestral asymmetries in costs were not static. Rather, they varied depending on context, such as the perceiver's individual characteristics and her social environment. If moderating contexts were recurrent, consistent in their effects, and indicated by the presence of reliable cues, we should expect judgment adaptations to respond to them with variable degrees of bias today. What follows are examples of systematic variations of

cost asymmetries and how they lead to variable biases in some of the domains we have already discussed. These moderating effects are also summarized in Table 2.

Men's Sexual Overperception is Affected by Women's Attractiveness. According to EMT, men's overperception of women's sexual intent evolved because it decreased the chances that men would miss sexual opportunities. If attractiveness is a cue of fertility (Symons, 1979), missing a sexual opportunity with an attractive woman would have been especially costly. In accordance with this prediction, women who are more attractive report more past episodes of sexual overperception by men (Haselton, 2003). In addition, a study by Maner et al. (2005) showed men and women clips of romantically arousing films and then asked them to interpret "micro-expressions" in photographs of faces that were actually neutral in expression. Men who watched the romantic film clip perceived sexual arousal in these neutral female faces, particularly when the faces were attractive.

In addition to documenting the effect of attractiveness, this study is also notable for its clear demonstration of a male bias. There has been controversy in the literature about what constitutes an appropriate criterion for men's judgments of women's interest (Farris, Treat, Viken, & McFall, 2008b). For instance, two commonly used criteria – women's reports of their own or other women's sexual interest – might also be systematically biased (Haselton & Buss, 2000). However, in the Maner et al. (2005) study, these concerns are not applicable because it used an objective criterion (i.e., the faces showed zero actual interest) and showed that men's judgments exceed this criterion.

Women's Commitment Underperception is Affected by Fertility Status. According to EMT, women's commitment underperception bias minimizes the chance of suffering

Table 2. *Notable Influences of Context on Error Management Biases.*

Bias	Moderating Variable is Attribute of...	Moderating Variable	Effect of High Level of Moderating Variable on Bias
Men's overestimation of women's sexual interest	Target	Female target's attractiveness	Bias ↑
Women's underestimation of men's relationship commitment	Observer * Target interaction	Fertility within ovulatory cycle * Male target's physical attractiveness	Bias ↓ (high fertility, high male attractiveness)
	Observer * Target interaction	Female observer's sexual interest in the male target	Bias ↓
	Observer	Fertility (pre- vs. post-menopause)	Bias ↑
Women's overestimation of men's sexual coerciveness	Observer	Fertility within ovulatory cycle	Bias ↑
Prejudice against outgroups	Environment	Ambient darkness	Bias ↑
	Observer	Fertility within ovulatory cycle	Bias ↑
	Observer	Pregnancy	Bias ↑
	Observer	Perceived vulnerability to disease	Bias ↑
Social exchange heuristic	Observer and Environment	Degree to which culture is collectivist	Bias ↑
	Environment	Presence of subtle cues of surveillance	Bias ↑

the costly reproductive consequences of having offspring with a man who will provide no paternal investment. Thus, the bias might be linked to the woman's fertility. In a recent study, Cyrus, Schwarz, and Hassebrauck (2011) found no evidence of the bias in an older sample that included postmenopausal women, but successfully replicated the results of Haselton and Buss in a younger sample using parallel methods. This result suggests that the bias is active only in women who face reproductive consequences as a result of commitment perceptions.

Perceptions of Coerciveness and Commitment Vary Across the Ovulatory Cycle.

An error management approach predicts psychological shifts within individuals across time based on changes in the relative fitness costs of false positive and false negative errors (Haselton & Nettle, 2006). These psychological shifts should occur in response to varying cues in the modern environment that were consistently associated with varying reproductive consequences throughout evolutionary history.

One type of shift occurs in women across the ovulatory cycle: the costs of having sex with an undesirable partner – a man who has low genetic fitness and will not invest in his children – rise as the chance of pregnancy increases near ovulation (Gangestad & Thornhill, 2008; Pillsworth & Haselton, 2006a). Noting this, Garver-Apgar and colleagues hypothesized that women will be particularly wary of sexual coercion when fertility is high, and they might avoid coercion by systematically increasing their perceptions of men's sexual coerciveness in men as ovulation approaches. This is precisely what their study showed: when women rated men appearing in videotaped interactions, women in their high-fertility phase rated the men as more sexually coercive than did women in other cycle phases (Garver-Apgar, Gangestad, & Simpson, 2007).

On the flip side, the benefits of having sex with a highly desirable partner – a man who has high genetic fitness – rise near ovulation when fertility is high (Pillsworth & Haselton, 2006a). In men, physical attractiveness is a hypothesized marker of fitness (Pillsworth & Haselton, 2006a). Even if such a man does not provide commitment, a woman can still receive genetic benefits for her offspring from a sexual encounter with him. In accordance with this logic, women elevate their standards for men's physical attractiveness in short-term sexual encounters relative to long-term relationships (Li & Kenrick, 2006). Thus, physically attractive men are more desired by women for short-term sexual encounters, which decreases these men's need to commit to a particular woman (Gangestad & Simpson, 2000). As a result, high levels of both attractiveness and commitment can be difficult to find in the same man.

Durante and colleagues (Durante, Griskevicius, Simpson, & Li, 2010) therefore hypothesized that encountering an attractive man who *might* be committed at high fertility involves an altered error asymmetry for the woman in perceiving his commitment. Unlike in most other circumstances, the false positive might actually be less costly than the false negative. The false positive error – assuming he is committed when he is not – could still lead to being abandoned but also gaining genetic benefits from the sexual encounter. The false negative error – assuming he is not committed when he actually is – could lead to missing out on the genetic benefits as well as on commitment.

In accordance with this new logic, Durante and colleagues found a striking shift in women's commitment perception when women judged attractive men at high fertility. In two separate sessions, one at low and one at high fertility, women were presented with the same photos of men who varied in attractiveness and commitment (described in a

vignette). Women then rated each man's interest in a committed long-term relationship and in investing in children with them. Results showed that at high fertility relative to low fertility, women thought the attractive men were more interested in commitment and in raising children. Possibly, this shift functions to increase women's motivation to consent to mating with attractive men when fertility is high.

Another recent study involving face-to-face interactions (Henningesen & Henningesen, 2010) showed that, the more sexually interested women were in their interaction partner, the more commitment they perceived from him. Although this result is correlational, it accords well with the findings of Durante et al. (2010) and might even suggest the mechanism (sexual interest) that allows women to adaptively shift their perceptions of men's commitment with shifting context.

Prejudice Against Outgroups is Increased by Vulnerability. A key error management hypothesis, as noted above, is that people overestimate the aggressiveness and hostility of outgroup members because, throughout evolutionary history, missing hostile intentions was more costly than overestimating it. This cost asymmetry probably became even more pronounced in circumstances when one was more vulnerable to attack. Based on this logic, Schaller and colleagues hypothesized and found that ambient darkness – a cue signaling increased vulnerability – increases racial and ethnic stereotypes connoting violence, but that it has little effect on other negative stereotypes (e.g., laziness or ignorance; Schaller, Park, & Faulkner, 2003; Schaller, Park, & Mueller, 2003). Following similar logic, women who were nearing ovulation and perceived they were vulnerable to sexual coercion showed increased hostility toward outgroup members (Navarrete, Fessler, Fleischman, & Geyer, 2009). Another striking demonstration is that

pregnant women, who have a heightened vulnerability to infection in the first trimester, show increased hostility toward outgroup members, who are often conceptualized as sources of disease (Navarette, Fessler, & Eng, 2007). Finally, research has shown that people who perceive that they are vulnerable to disease are less likely to have disabled friends (Park et al., 2003), possibly because disabled individuals trigger a disease avoidance bias.

The Social Exchange Heuristic Varies with Culture and Subtle Surveillance Cues.

Although the costs of selfish behavior were likely higher in most ancestral settings than the opportunity cost of cooperating, this asymmetry might have been even further enhanced by context. For instance, the costs of ostracism (resulting from selfish or dishonest behavior) may be particularly high in interdependent social contexts in which cooperation is either highly valued or especially necessary. As predicted by this logic, in Japanese collectivist samples, cooperation in one-shot experiments is higher than in the more individualist United States samples (Yamagishi, Jin, & Kiyonari, 1999). Moreover, the study described earlier (Haley & Fessler, 2005) showed that the social exchange heuristic, which is already pervasive, becomes even stronger when there are subtle cues of social surveillance, such as eyespots.

Recent Challenges to EMT

Recently, McKay and colleagues introduced a novel perspective of EMT, rooted in philosophy and economic modeling (McKay & Dennett, 2009; McKay & Efferson, 2010). They affirm the logic of EMT, but argue that to solve adaptive problems of the sort explained by EMT, humans do not need to possess *biased beliefs* if *biased actions* can accomplish the same ends while preserving true beliefs. McKay & Efferson (2010, p.

311) illustrate this point using the example of sexual overperception, by pointing out that having a strong belief that a woman is sexually interested is unnecessary to approach her. Instead, because for a man the payoff of a short-term sexual encounter, however improbable, is so large, even *accurately* knowing that there is only a small chance of success (and the corresponding large chance of incurring a small cost, such as a slap in the face) should not deter men from taking a chance.

We agree that this is a plausible theoretical possibility. Although the theory was advanced to explain cognitive biases, the core logic of EMT is neutral in predicting whether a bias will be cognitive or purely behavioral. The logic of EMT is satisfied as long as humans behave so that they minimize the more costly of the two errors in question. However, ultimately, the question of whether solutions to error management problems are sometimes rooted in biased *cognition* is an open issue that must be decided on a case-by-case basis with empirical research (see Haselton & Buss, 2009). The argument that, in theory, error management adaptations need not involve cognitive bias does not invalidate the possibility that the cognitive bias exists.

In the case of sexual overperception, we suggest that there is a compelling empirical reason to consider the possibility that this bias is cognitive rather than purely behavioral. As shown in Table 2, there is evidence across a range of studies that men actually do overestimate women's interest in face-to-face interactions when judging videotaped interactions and photos, in vignettes, real friendships (Koenig, Kirkpatrick, & Ketelaar, 2007), and in experiments in which women's faces objectively express zero interest (Maner et al., 2005). The notion that selection should not bias beliefs is difficult to reconcile with the fact that men appear to overestimate women's interest in all of these

varied ways, but especially in self-report measures. Self-reported estimations reflect biased *beliefs* rather than biased actions.

A number of other biases we have described in this chapter are also cognitive rather than purely behavioral. For instance, there may be behavioral adaptations to avoid falling off of cliffs (as demonstrated via the “visual cliff” paradigm; e.g., Campos, Langer, & Krowitz, 1970), without an accompanying bias that overestimates the height of the cliff from the top. However, humans do, in fact, appear to have an estimation bias (a biased *belief*) beyond any such behavioral biases. Similarly, one could easily imagine an adaptation to avoid an individual who committed a dishonest action, without necessarily holding the belief that this individual is generally dishonest. However, research has documented such a negativity bias in attribution (i.e., beliefs). Likewise, a cognitive bias is theoretically unnecessary for behaviorally avoiding potentially dangerous outgroup members, but again, research shows that when vulnerable, people are biased to see anger in neutral outgroup faces – a bias in beliefs.

A second major argument advanced by McKay & Efferson (2010) is that EMT’s novelty is contained entirely in the idea that *cognitive* bias can be adaptive. In contrast, they claim that the idea that *behavioral* bias can minimize costs can be derived entirely from expected utility theory (von Neumann & Morgenstern, 1944). Hence, EMT is useful only insofar as it predicts genuine cognitive biases (departures from Bayesian beliefs). Our response is two-fold. First, we believe we have presented a strong case that many of the biases explained and/or predicted by EMT are, in fact, cognitive. Second, we agree that the basic idea of biased responses to asymmetric costs existed before EMT, as reflected in expected utility theory and signal detection theory. However, we emphasize

that a key contribution of EMT lies in applying this logic to *recurring costs and benefits over human evolutionary history*, and not simply to present costs and benefits. In expected utility theory and signal detection theory, the utility maximizers are conscious animal or human agents or devices constructed by humans (e.g., smoke detectors or radars); in error management theory, the utility maximizer is natural selection, and the measure of utility is reproductive success. This idea of considering recurring ancestral costs and benefits is an important contribution to understanding biases, regardless of whether they are cognitive or behavioral.

Conclusion

Error management theory has been useful both for testing new hypotheses and integrating existing findings of bias which might otherwise seem singular or puzzling. In this chapter, we have described psychological biases in a variety of domains which fall under the theoretical umbrella of EMT. Some of these biases were discovered before the advent of EMT but can be at least partly explained by it. These include men's sexual overperception, certain types of prejudice against outgroups, disabled, and ill people, and negative attribution. In addition, EMT has motivated the recent discovery or empirical refinement of a number of phenomena: women's commitment underperception, the forgiveness bias, men's underestimation of their partners' faithfulness, both sexes' overestimation of the partner's attractiveness to competitors, women's overestimation of men's sexual coerciveness at high fertility, and the social exchange heuristic.

EMT incorporates hypotheses about variation in biases in response to contextual inputs. We have described a number of examples of these contingencies, one of the most striking of which was Durante et al.'s (2010) finding that women's commitment

underperception bias is sensitive to a specific set of contextual influences that simultaneously encompass both the observer's and the target's characteristics. These results demonstrate the utility of EMT for generating new findings.

In addition to the empirical findings being generated by EMT, there are also exciting new theoretical developments, such as McKay and colleagues' focus on the distinction between cognitive and behavioral biases. We believe this is an important theoretical refinement which provides an avenue for fruitful future research.

References

- Abbey, A. (1982). Sex differences in attributions for friendly behavior: Do males misperceive females' friendliness? *Journal of Personality and Social Psychology*, *42*, 830-838.
- Abbey, A., & Melby, C. (1986). The effects of nonverbal cues on gender differences in perceptions of sexual intent. *Sex Roles*, *15*, 283-298.
- Andrews, P. W., Gangestad, S. W., Miller, G. F., Haselton, M. G., Thornhill, R., & Neale, M. C. (2008). Sex differences in detecting sexual infidelity: Results of a maximum likelihood method for analyzing the sensitivity of sex differences to underreporting. *Human Nature*, *19*, 347-373.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323-370.
- Bishop, G. D., Alva, A. L., Cantu, L., & Rittiman, R. K. (1991). Responses to persons with AIDS: Fear of contagion or stigma? *Journal of Applied Social Psychology*, *21*, 1877-1888.
- Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*, 307-324.
- Buss, D. M., & Schmitt, D. P. (1993). Sexual Strategies Theory: An evolutionary perspective on human mating. *Psychological Review*, *100*, 204-232.
- Camerer, C., & Thaler, R. (1995). Ultimatums, dictators and manners. *Journal of Economic Perspectives*, *9*, 337-356.
- Campos, J. J., Langer, A., & Krowitz, A. (1970). Cardiac response on the visual cliff in prelocomotor human infants. *Science*, *170*, 196-197.

- Caporael, L., Dawes, R. M., Orbell, J. M., & van der Kragt, A. J. (1989). Selfishness examined. *Behavioral and Brain Sciences*, *12*, 683–739.
- Carlston, D. E. (1980). The recall and use of traits and events in social inference processes. *Journal of Experimental Social Psychology*, *16*, 303-328.
- Clark, R. D., & Hatfield, E. (1989). Gender differences in receptivity to sexual offers. *Journal of Psychology & Human Sexuality*, *2*, 39-55.
- Clutton-Brock, T. H., & Vincent, A. C. J. (1991). Sexual selection and the potential reproductive rates of males and females. *Nature*, *351*, 58-60.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187-276.
- Cyrus, K., Schwarz, S., & Hassebrauck, M. (2011). Systematic cognitive biases in courtship context: Women's commitment-skepticism as a life-history strategy? *Evolution and Human Behavior*, *32*, 13-20.
- Durante, K. M., Griskevicius, V., Simpson, J. A., & Li, N. P. (2010). *Ovulation leads women to overperceive commitment from sexy cads but not good dads*. Paper presented at the 22nd annual meeting of the Human Behavior and Evolution Society Conference, Eugene, OR.
- Edmondson, C. B., & Conger, J. C. (1995). The impact of mode of presentation on gender differences in social perception. *Sex Roles*, *32*, 169-183.
- Farris, C., Treat, T. A., Viken, R. J., & McFall, R. M. (2008a). Perceptual mechanisms that characterize gender differences in decoding women's sexual intent. *Psychological Science*, *19*, 348-354.

- Farris, C., Treat, T. A., Viken, R. J., & McFall, R. M. (2008b). Sexual coercion and the misperception of sexual intent. *Clinical Psychology Review, 28*, 48-66.
- Friesen, M. D., Fletcher, G. J. O., & Overall, N. C. (2005) A dyadic assessment of forgiveness in intimate relationships. *Personal Relationships, 12*, 61-77.
- Gangestad, S. W., Haselton, M. G., & Buss, D. M. (2006). Evolutionary foundations of cultural variation: Evoked culture and mate preferences. *Psychological Inquiry, 17*, 75-95.
- Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences, 23*, 675-687.
- Gangestad, S. W., & Thornhill, R. T. (2008). Human oestrus. *Proceedings of the Royal Society (B), 275*, 991-1000.
- Garver-Apgar, C. E., Gangestad, S. W., & Simpson, J. A. (2007). Women's perceptions of men's sexual coerciveness change across the menstrual cycle. *Acta Psychologica Sinica. Special Issue: Evolutionary Psychology, 39*, 536-540.
- Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology, 65*, 221-233.
- Green, D. M., & Swets, J. A. (1966). *Signal detection and psychophysics*. New York: Wiley.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior, 26*, 245-256.
- Harnish, R. J., Abbey, A., & DeBono, K. G. (1990). Toward an understanding of "The Sex Game": The effects of gender and self-monitoring on perceptions of sexuality and likeability in initial interactions. *Journal of Applied Social Psychology, 20*, 1333-1344.
- Haselton, M. G. (2003). The sexual overperception bias: Evidence of a systematic bias in men from a survey of naturally occurring events. *Journal of Research in Personality, 37*, 34-47.
- 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., Buss, D. M., Oubaid, V. & Angleitner, A. (2005). Sex, lies, and strategic interference: The psychology of deception between the sexes. *Personality and Social Psychology Bulletin, 31*, 3-23.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47-66.
- Henningsen, D. D., & Henningsen, M. L. M. (2010). Testing error management theory: Exploring the commitment skepticism bias and the sexual overperception bias. *Human Communication Research, 36*, 618-634.

- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Gintis, H., McElreath, R., et al. (2001). In search of Homo economicus: Experiments in 15 small-scale societies. *American Economic Review*, *91*, 73–79.
- Hill, S. E. (2007). Overestimation bias in mate competition. *Evolution and Human Behavior*, *28*, 118-123.
- Jackson, R. E., & Cormack, L. K. (2007). Evolved navigation theory and the descent illusion. *Perception & Psychophysics*, *69*, 353-362.
- Johnson, D. D. P. (2009). The error of God: Error management theory, religion, and the evolution of cooperation. In S. A. Levin (Ed.), *Games, Groups, and the Global Good* (pp. 169-180). Berlin; London: Springer.
- Karremans, J. C., Verwijmeren, T., Pronk, T. M., & Reitsma, M. (2009). Interacting with women can impair men's executive functioning. *Journal of Experimental Social Psychology*, *45*, 1041-1044.
- Keeley, L. H. (1996). *War before civilization: The myth of the peaceful savage*. New York: Oxford University Press.
- Koenig, B. L., Kirkpatrick, L. A., & Ketelaar, T. (2007). Misperception of sexual and romantic interests in opposite-sex friendships: Four hypotheses. *Personal Relationships*, *14*, 411-429.
- Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin*, *123*, 187–208.
- La France, B. H., Henningsen, D. D., Oates, A., & Shaw, C. M. (2009). Social-sexual interactions? Meta-analyses of sex differences in perceptions of flirtatiousness, seductiveness, and promiscuousness. *Communication Monographs*, *76*, 263-285.

- Li, N. P., & Kenrick, D. T. (2006). Sex similarities and differences in preferences for short-term mates: What, whether, and why. *Journal of Personality and Social Psychology, 90*, 468-489.
- 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.
- Mineka, S. (1992). Evolutionary memories, emotional processing, and the emotional disorders. *Psychology of Learning and Motivation, 28*, 161-206.
- Navarrete, C. D., Fessler, D. M. T., & Eng, S. J. (2007). Increased ethnocentrism in the first trimester of pregnancy. *Evolution and Human Behavior, 28*, 60-65.
- Navarrete, C. D., Fessler, D. M. T., Fleischman, D., & Geyer, J. (2009). Race bias tracks conception risk across the menstrual cycle. *Psychological Science, 20*, 661-665.
- Park, J. H., Faulkner, J., & Schaller, M. (2003). Evolved disease-avoidance processes and contemporary anti-social behavior: Prejudicial attitudes and avoidance of people with disabilities. *Journal of Nonverbal Behavior, 27*, 65-87.
- Pillsworth, E. G., & Haselton, M. G. (2006a). Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention. *Evolution and Human Behavior, 27*, 247-258.

- Pillsworth, E. G., & Haselton, M. G. (2006b). Women's sexual strategies: The evolution of long-term bonds and extra-pair sex. *Annual Review of Sex Research, 17*, 59-100.
- Quillian, L., & Pager, D. (2001). Black neighbors, higher crime? The role of racial stereotypes in evaluations of neighborhood crime. *American Journal of Sociology, 107*, 717-767.
- Reeder, G. D., & Coovert, M. D. (1986). Revising an impression of morality. *Social Cognition, 4*, 1-17.
- Reeder, G. D., & Spores, J. M. (1983). The attribution of morality. *Journal of Personality and Social Psychology, 44*, 736 - 745.
- Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*, 131-142.
- Rozin, P., Markwith, M., & Nemeroff, C. (1992). Magical contagion beliefs and fear of AIDS. *Journal of Applied Social Psychology, 22*, 1081-1092.
- Saal, F. E., Johnson, C. B., & Weber, N. (1989). Friendly or sexy? It may depend on whom you ask. *Psychology of Women Quarterly, 13*, 263-276.
- Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*, 58-92.
- Schaller, M., & Duncan, L. A. (2007). The behavioral immune system: Its evolution and social psychological implications. In J. P. Forgas, M. G. Haselton, & W. von Hippel (Eds.), *Evolution and the social mind: Evolutionary psychology and social cognition* (pp. 293-307). New York: Psychology Press.

- Schaller, M., Park, J. H., & Faulkner, J. (2003). Prehistoric dangers and contemporary prejudices. *European Review of Social Psychology, 14*, 105-137.
- Schaller, M., Park, J. H., & Mueller, A. (2003). Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes. *Personality and Social Psychology Bulletin, 29*, 637-649.
- Schmitt, D. P., Alcalay, L., Allik, J., Ault, L., Austers, I., Bennett, K. L., et al. (2003). Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. *Journal of Personality and Social Psychology, 85*, 85-104.
- Shotland, R. L., & Craig, J. M. (1988). Can men and women differentiate between friendly and sexually interested behavior? *Social Psychology Quarterly, 51*, 66-73.
- Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology, 60*, 870-883.
- Symons, D. (1979). *The evolution of human sexuality*. New York: Oxford University Press.
- Trafimow, D. (2001). The effects of trait type and situation type on the generalization of trait expectancies across situations. *Personality and Social Psychology Bulletin, 11*, 1463-1468.
- Treat, T. A., Viken, R. J., Kruschke, J. K., & McFall, R. M. (2010). Men's memory for women's sexual-interest and rejection cues. *Applied Cognitive Psychology*. doi: 10.1002/acp.1751.

- Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man: 1871-1971* (pp. 136-179). Chicago, IL: Aldine.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
- Yamagishi, T., Jin, N., & Kiyonari, T. (1999). Bounded generalized reciprocity: Ingroup favoritism and ingroup boasting. *Advances in Group Processes, 16*, 161–197.
- Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., & Kanazawa, S. (2007). The social exchange heuristic: Managing errors in social exchange. *Rationality and Society, 19*, 259-291.
- Ybarra, O. (2002). Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin, 128*, 421–441.
- Ybarra, O., & Stephan, W. G. (1996). Misanthropic person memory. *Journal of Personality and Social Psychology, 70*, 691–700.