E

Error Management Theory and the Evolution of Cognitive Bias

Bernardo U. Seixas and Laith Al-Shawaf
Department of Psychology, University of Colorado – Colorado Springs, Colorado Springs, CO, USA

Synonyms

Signal detection theory; Smoke detector principle; Statistical decision theory

Definition

Error management theory (EMT) is a theory of how humans make decisions under uncertainty. It proposes that when information is incomplete and the fitness costs of making false positive and false negative errors are consistently asymmetric over evolutionary time, natural selection favors decision mechanisms that are predictably oriented toward the less costly error. Thus, humans are equipped with numerous "better safe than sorry" cognitive biases that are adaptively biased toward the less costly of the two types of error.

The Logic of Error Management Theory

Error management theory is an application and extension of signal detection theory (Green & Swets, 1966) developed by Martie Haselton and David Buss (2000). While EMT initially emerged in the context of predicting and explaining sex differences in the perception of romantic interest and commitment, the logic of the theory applies to a wide variety of psychological phenomena that involve decisions under uncertainty, including anxiety disorders (Nesse, 2005) and a variety of cognitive biases (Haselton & Nettle, 2006; Haselton et al., 2009). As such, error management theory can be thought of as an overarching framework for how humans make inferences and decisions under conditions of uncertainty.

Error management theory and its forebear, signal detection theory (Green & Swets, 1966), are closely associated with the principles of statistical hypothesis testing, in which an observer faces four possible inferential outcomes: correctly concluding there is an effect, correctly concluding that an effect is absent, erroneously concluding that there is an effect when there is none (false positive, also known as type I error or α), or erroneously concluding that there is no effect when in fact there is one (false negative, also known as type II error or β). Often, there are different costs and payoffs for each of these outcomes; the roots of this realization go at least as far back as Blaise Pascal and his famed wager (Pascal, 1669/2008). Since all animals must

T. K. Shackelford (ed.), *Encyclopedia of Sexual Psychology and Behavior*,
https://doi.org/10.1007/978-3-031-08956-5_83-1

confront various decisions in which false positives and false negatives are possible outcomes, and since the costs of the two types of error for an organism's fitness are frequently different, natural selection is expected to have favored decision rules that reliably commit the less costly error given the relevant risks and payoffs of the local ecology. Table 1 shows the possible decisions an agent can make under uncertainty.

EMT as a General Theory of Cognitive Bias

Error management theory predicts that cognitive biases will evolve if the following three conditions are met: (1) there is noise or uncertainty in a decision-making domain; (2) the fitness costs of the two types of error (false positive and false negative) differ, such that one is more costly than the other; and (3) these conditions recur over evolutionary time (Haselton et al., 2015). Under these conditions, a cognitive system that is biased toward the less costly error will outperform even an unbiased system. This is because the "adaptively biased" system might make more mistakes on average, but these will be of the less costly type rather than the more costly type. Because natural selection does not minimize the *number* of errors the human mind makes (the crude error rate), it minimizes *net fitness costs* (Haselton et al., 2015); the above conditions lead to the evolution of decision-making systems that are adaptively biased toward the less costly error.

Because the logic of error management theory is not restricted to any specific content area, the theory is broadly applicable to many domains of psychology. The theory predicts that as long as the three preconditions are met, biased neurocognitive mechanisms will evolve, regardless of the content domain. Below we offer multiple examples across diverse domains.

The Auditory Looming Bias: On evolutionary grounds, Neuhoff (2001) proposed that one function of hearing is to provide the listener with a model of the environment that helps them avoid approaching threats. He predicted and found that people would perceive change in loudness to be greater for a sound increasing in intensity than for a sound decreasing in intensity by an equivalent amount. He further showed that listeners overestimate how close an approaching sound source is. These results show that people tend to overestimate the loudness of a sound, the closeness of its source, and how fast the sound is approaching, all which are helpful in anticipating and avoiding collisions and attacks. More recent investigation into auditory looming showed that, as predicted, individuals in poor physical condition have a larger looming bias than those in better condition, and that people exposed to an infant's cry exhibit larger auditory looming biases than those not exposed to it (Haselton et al., 2015).

Overestimation of Heights and the Descent Illusion: If you are standing at the top of a cliff, it is riskier to underestimate the height of the cliff than to overestimate it – overestimating it might make you more cautious, whereas underestimating it might lead to a reckless, fatal navigation decision. In this context, evolved navigation theory uses error management logic to predict that people have a general tendency to overestimate heights. Furthermore, the theory predicts that people will consider a surface to be higher when looking at it from the top than from the bottom, given that danger levels are much higher from the top than the bottom. These a priori predictions have been verified by empirical studies (Jackson & Cormack, 2007; Stefanucci & Proffitt, 2009), showing that error management principles can be used to predict new biases in visual perception and the psychology of navigation.

Anxiety: Psychiatrist Randolph Nesse has used the "smoke detector principle," which follows the same underlying logic as error management theory, to explain anxiety disorders as the outcome of an overactive behavioral defense system (Nesse, 2005). Defense systems (such as cough, fever, and panic) carry activation costs. However, the cost of failing to respond to a threat (a false negative) is usually larger than the cost of activating a defense when it is not needed (a false positive). Nesse used quantitative models to demonstrate that even under expectations of

Error Management Theory and the Evolution of Cognitive Bias, Table 1 Possible inferences an agent can make under uncertainty

	Effect is present	Effect is absent
Conclude there is an effect	Hit	False alarm (false positive or type I error)
Conclude there is no effect	Miss (false negative or type II error)	Correct rejection

optimality, we would expect defense mechanisms to be overexpressed (i.e., expressed sometimes even when they are not needed; Nesse, 2005). This refined understanding of how defense mechanisms work has important practical implications, as it can help inform our decisions of when it is (and is not) safe to block defenses such as fever, cough, and anxiety (Nesse, 2005, 2019).

Superstition, paranoia, and conspiracy theories: Recent research has suggested that the human tendency to believe in conspiracy theories about threatening outgroups may be the outcome of an error-management bias: If there is even a remote chance that the conspiracy is true, and if its truth would carry dangerous or fatal consequences for the individual, it may be more useful to err in the direction of conspiracy gullibility than to risk committing the false negative error and missing a dangerous threat (van Prooijen & van Vugt, 2018; Pinker, 2021).

The sexual overperception bias: Among ancestral men, the cost of missing a reproductive opportunity would have been much larger than the cost of erroneously perceiving sexual interest in a potential mate (Symons, 1979). This evolutionary logic explains the phenomenon of male overperception of sexual intent found in previous studies (e.g., Abbey, 1982, 1987) and also leads to an ancillary prediction which was subsequently verified by studies: that men will not overperceive sexual intent when their sister is the target, as the overperception of sexual intent is specifically geared toward unrelated individuals (Haselton & Buss, 2000).

The commitment skepticism bias: For ancestral women, it would have been more costly to overestimate a potential mate's commitment intent than to underestimate it, because the first error could have caused a woman to be abandoned by her mate during a critical time, whereas the second error would only impose the relatively minor cost of prolonging the courtship and evaluation period (Hurtado & Hill, 1992). This led researchers to predict a new bias in person perception: that women will display a systematic bias toward underperceiving men's commitment intent (Haselton & Buss, 2000). Confirming this prediction, Haselton and Buss (2000) found that women consistently underperceive men's commitment intentions. This was later confirmed by independent researchers (e.g., Cyrus et al., 2011; see also Al-Shawaf, 2016, for discussion of the factors that may erase or reverse these biases). These findings illustrate how error management considerations lead to the discovery of new cognitive biases in person perception.

How Does EMT Differ from the Traditional Heuristics and Biases Approach?

Readers will have noticed that Error Management Theory differs from the traditional heuristics and biases approach in psychology (Tversky & Kahneman, 1974) because it does not include the default assumption that mistakes are a product of bad biological or psychological design. Instead, it proposes that some cognitive biases may be adaptations to ancestral problems in which there were recurrent asymmetries in the two types of error (Haselton et al., 2015). In other words, some cognitive biases may be design *features* rather than design *flaws* (Ketelaar, in press). In this way, Error Management Theory suggests that while cognitive biases may appear irrational at first, some of them follow an adaptive logic – they conform to the net fitness costs and benefits of the two types of error present during the time that they evolved. This is why they are sometimes described as "adaptively rational" or "adaptively biased."

EMT Leads to New Predictions

Error Management Theory has helped to explain previous but puzzling findings, such as the male overperception of female sexual interest. It has also produced novel predictions and new findings, such as the discovery of women's commitment skepticism bias, the descent illusion, and the auditory looming bias.

Recently, researchers used Error Management Theory to predict and discover two new cognitive biases: the Male Overperception of Attractiveness Bias (MOAB) and the Female Underperception of Attractiveness Bias (FUAB). The researchers used error management logic to predict that in initial person perception, men would be more likely to overperceive a woman's attractiveness under conditions of uncertainty, whereas women would underperceive a man's attractiveness.

They found that when men and women were asked to rate the attractiveness of opposite-sex targets whose pictures were blurred (and thus conditions were uncertain), men tended to overestimate the attractiveness of unattractive images. On the other hand, when images were blurred, women tended to underestimate the attractiveness of attractive men. The researchers then suggested and tested competing decision rules to see which one best accounts for the results, emphasizing that research will benefit from paying more attention to the specific algorithm the mind uses to generate these inferences (Lewis et al., 2022; see also Lewis et al., 2023).

The above examples, which represent just a sample of the total, suggest that EMT is a rich conceptual framework that generates many new hypotheses about human cognition. Additionally, its logic can be applied to fields other than psychology, including biology, medicine, and engineering. In ethology, for instance, it has long been known that animals possess risk-reduction strategies when foraging, such that they often skip a meal in order to avoid possible predation (Johnson et al., 2013). In medicine, defense mechanisms such as cough and allergy have also been hypothesized to follow the smoke alarm principle (Johnson et al., 2013). We encourage the reader to apply the logic of error management to new areas of psychology, inference generation, and decision-making.

More broadly, Error Management Theory brings a fresh perspective to the debate on human rationality (Haselton et al., 2009). This debate has been dominated by a perception that humans are incorrigibly flawed, falling well below the benchmarks of formal logic and Bayes' Theorem in the arena of decision-making (Haselton et al., 2015; Pinker, 2021). EMT suggests that far from being poor decision-makers, as the heuristics and biases approach suggests, humans are in fact "better than rational": When perfect accuracy is impossible, we are adapted to make frequent but inexpensive mistakes that ancestrally increased our fitness (on average) and that outperform more accurate decision-making systems that distribute their mistakes more evenly between the two types of error. Error management theory shows human cognition and decision-making in a new light and helps us understand what it means to be "adaptively biased."

Cross-References

- Auditory Looming Bias
- Cognitive Bias
- Commitment Skepticism Bias
- Fever
- Mate Choice
- Panic
- Sexual Overperception Bias
- Smoke-detector Principle

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. (1987). Misperceptions of friendly behavior as sexual interest: A survey of naturally occurring instances. *Psychology of Women Quarterly, 11*, 173–194.

Al-Shawaf, L. (2016). Could there be a male commitment skepticism bias and a female sexual overperception bias? Novel hypotheses based on error management

theory. *Evolutionary Psychological Science, 2*(3), 237–240.

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*(1), 13–20.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969–1912). Wiley.

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*(1), 81.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*(1), 47–66.

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*(5), 733–763.

Haselton, M. G., Nettle, D., & Andrews, P. W. (2015). The evolution of cognitive bias. In *The handbook of evolutionary psychology* (pp. 724–746). Wiley.

Hurtado, A. M., & Hill, K. R. (1992). Paternal effect on offspring survivorship among ache and Hiwi hunter-gatherers. In B. S. Hewlett et al. (Eds.), *Father-child relations: Cultural and biosocial contexts* (pp. 31–55). Aldine de Gruyter.

Jackson, R. E., & Cormack, L. K. (2007). Evolved navigation theory and the descent illusion. *Perception and Psychophysics, 69*, 353–362.

Johnson, D. D., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends in Ecology and Evolution, 28*(8), 474–481.

Ketelaar, T. (in press). Defect or design feature?: Toward an evolutionary psychology of the role of emotion in motivated reasoning. In L. Al-Shawaf & T. Shackelford (Eds.), *The Oxford handbook of evolution and the emotions*. Oxford University Press.

Lewis, D. M., Al-Shawaf, L., Semchenko, A. Y., & Evans, K. C. (2022). Error management theory and biased first impressions: How do people perceive potential mates under conditions of uncertainty? *Evolution and Human Behavior, 43*(2), 87–96.

Lewis, D. M., Al-Shawaf, L., Semchenko, A. Y., & Evans, K. C. (2023). Creative ideas, but elementary mistakes: A reply to Brandner and colleagues to promote best practices in error management theory research. Evolution and Human Behavior. https://www.sciencedirect.com/science/article/pii/S1090513823000090?casa_token=z1ovH3TqcqkAAAAA:or_OyJar6U8dAvdif4zpFEpeBmZnUg89r7M_GO0FcQ8sgiD5JsYVV81jVq_BnmpN3sepw0oD4M

Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior, 26*(1), 88–105.

Nesse, R. M. (2019). *Good reasons for bad feelings: Insights from the frontier of evolutionary psychiatry.* Penguin.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology, 13*, 87–110.

Pascal, B. (1669/2008). *Pensées and other writings*. Oxford University Press.

Pinker, S. (2021). *Rationality: What it is, why it seems scarce, why it matters*. Penguin.

Stefanucci, J. K., & Proffitt, D. R. (2009). The roles of altitude and fear in the perception of height. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 424.

Symons, D. (1979). *The evolution of human sexuality.* Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1121–1131.

van Prooijen, J.-W., & van Vugt, M. (2018). Conspiracy theories: Evolved functions and psychological mechanisms. *Perspectives on Psychological Science, 13*(6), 770–788.