

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/126100/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Jefferson, Anneli 2017. Born to be biased? Unrealistic optimism and error management theory. *Philosophical Psychology* 30 (8) , pp. 1159-1175.  
10.1080/09515089.2017.1370085

Publishers page: <https://doi.org/10.1080/09515089.2017.1370085>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



## ***Born to be Biased? Unrealistic Optimism and Error Management Theory***

*When individuals display cognitive biases, they are prone to developing systematically false beliefs. Evolutionary psychologists have argued that rather than being a flaw in human cognition, biases may actually be design features. In my paper, I assess the claim that unrealistic optimism is such a design feature because it is a form of error management. Proponents of this theory say that when individuals make decisions under uncertainty, it can be advantageous to err on the side of overconfidence if the potential gains through success are high and the costs of failure are low. I argue that there are a number of conceptual problems in matching the theory with the existing data. I also show that there is empirical evidence against the error management hypothesis.*

Keywords: Unrealistic Optimism, Evolutionary Psychology, Error Management Theory, Cognitive Bias

Anneli Jefferson

### ***1. Introduction<sup>1</sup>***

It is natural to think that true beliefs are most useful to navigate the natural and social world successfully. However, human cognition is riddled with false beliefs and biased belief forming processes. Evolutionary psychologists have taken up the challenge of coming up with explanations as to why these biases have survived and thrived. They argue that in many cases, we should not see biases as design flaws, but instead should conceive of them as design features. In other words, belief forming mechanisms which seem systematically flawed when one focuses on the accuracy of the beliefs they produce or on their epistemic warrant, may have conveyed some other advantage to individuals in historic environments in which they evolved. One such explanatory strategy is error management theory. According to error management theory, in situations of uncertainty, organisms have evolved to favor the least costly error. Some famous examples are snake over-perception or predator over-perception. Recently, psychologists have attempted to extend this explanation to other psychological processes such as positive illusions generally and unrealistic optimism specifically (cf. Johnson, Blumstein, Fowler, & Haselton, 2013; Johnson & Fowler, 2011), (Haselton & Nettle, 2006; McKay & Dennett, 2009; Nettle, 2004; Sedikides, Skowronski, & Gaertner, 2004).

When individuals make optimistically biased judgments, they systematically underestimate the likelihood of bad things happening to them (either in absolute terms or relative to others) and overestimate the likelihood of good things happening (either in absolute terms or relative to others).

---

<sup>1</sup> Anneli Jefferson is Leverhulme Early Career Fellow at the University of Birmingham.

Why should a belief-forming mechanism that gives me a systematically skewed outlook on the future be advantageous? Shelley Taylor (1988, 1994) and many others have claimed that unrealistic optimism and other positive illusions have psychological benefits as they enhance well-being and improve our ability to cope. However, in this paper I will be interested in a specific proposal according to which unrealistic optimism is an adaptation, which has been selected for because it has been beneficial to the survival and reproduction of individuals who exhibited it. While biological adaptations may be psychologically adaptive, they need not be. It is in principle possible that something that has been beneficial in terms of survival and reproduction was costly in terms of psychological well-being. In addition, it may also be the case that something that was historically adaptive is no longer so. Thus, a theory that pertains to show that a certain trait is an adaptation needs to show that the trait in question was (likely to have been) beneficial in terms of survival or reproduction in the past.

The Error Management account of unrealistic optimism attempts to explain the evolutionary benefits of unrealistic optimism by saying that unrealistically optimistic predictions are beneficial when and because they allow us to make the least costly error. Error management theory (EMT) holds that when making decisions under uncertainty, it is best to form beliefs where the cost of acting on erroneous beliefs is low and the potential gain resulting from a correct belief is high. This can be explained as making the least costly error because the cost of erring on the side of inaction is to miss out on the opportunity of a large gain.

“EMT (error management theory) predicts that if the cost of trying and failing is low relative to the potential benefit of succeeding, then an illusionary positive belief is not just better than an illusionary negative one, but also better than an unbiased belief.” (Haselton & Nettle, 2006 p. 58)

Proponents of EMT argue that unrealistic optimism is such a case, because, provided that the costs of failure to achieve a goal are low and the benefits of success are high, it is better to err on the side of trying to achieve a goal than to miss out on this opportunity. Overconfidence in our chances of success, so goes the argument, makes us more likely to make these attempts.

EMT is not the only evolutionary account of unrealistic optimism. However, it is the one that focuses most specifically on unrealistic optimism and lends itself most readily to empirical evaluation. Error management theory has been quite successful in explaining very specific, local phenomena such as snake-over perception, the Garcia effect, or auditory looming (Haselton & Nettle, 2006). Discovering whether it can be extended to very general biases such as unrealistic optimism is interesting for two

reasons: We potentially gain an explanation as to why we are unrealistically optimistic, and what the benefits of such a seemingly irrational way of thinking about the future are. We also learn more about the explanatory potential of the error management account, specifically, how it fares with beliefs which are sensitive to a lot more diverse information than the classic cases.

## **2. Error Management Theory**

EMT holds that when making decisions under uncertainty, it is sometimes better to systematically err in one direction. This can be illustrated with the time-honored example of the over-perception of snakes: “Because of the dire consequences of being bitten by a poisonous snake, it is better to have a low evidentiary threshold for inferring that long slender objects are snakes, and to identify every snake you encounter, than to require too much evidence and occasionally get a costly surprise.” (Haselton et al., 2009, p.742) In other words, as the cost of missing the odd poisonous snake is far higher than the cost of occasionally getting a scare when you think there is a snake although there is not, you should err on the side of being a bit too ready to perceive snakes.

Another useful analogy that is commonly used to illustrate this model is the smoke detector (Haselton and Nettle 2006, McKay and Dennett 2009). You can either have a smoke detector which is somewhat oversensitive and will go off even when you have just burned the toast. Alternatively, you can set the threshold at which it goes off higher, in which case you run the danger of missing the beginnings of an actual fire. Occasionally having to turn off an overactive smoke detector can be annoying, missing the beginnings of a fire in the house can have disastrous effects. Therefore it is best to err on the side of oversensitivity. Error management theory claims that in situations of uncertainty, we should expect the less costly error to be selected for.

This model is then applied to unrealistic optimism. Assuming that actual probabilities of success are hard or impossible to come by and that we are better off erring on the side of trying and failing, it is good to have an overinflated belief in one’s own chances of success. On this account, unrealistic optimism is adaptive if it occurs in situations where the cost of trying and failing is low, and the benefit of trying and succeeding is high (and correspondingly, the cost of losing out on this benefit is high). This cost-benefit asymmetry can be illustrated by another well known example from the evolutionary psychology literature. Haselton (2003) has argued that men tend to over-perceive women’s interest in them because for them, the cost of missing a potential mating opportunity is higher than that of experiencing a rejection of their advances. Women, on the other hand, are inclined to underestimate a mate’s willingness to commit, because it is very costly to become pregnant and then be without support when raising one’s offspring.

So, depending on the cost-benefit ratio, error management theory predicts either optimism or paranoia. As this is an evolutionary account, the currency of costs and benefits needs to be survival and reproduction, even if the effects on survival and reproduction can of course be indirect. Examples of beneficial paranoia are the snake case or females under-perceiving willingness to commit. Overconfidence and unrealistic optimism can be explained as being adaptive as long as there is an asymmetry between costs and benefits that makes it rational to give something a go even if the chances of success are low.

### *2.1. Unrealistic Optimism as Error Management – Some Clarifications*

With this general picture in hand, there are a few clarificatory comments in order. First, as Ryan McKay and Daniel Dennett (2009) point out, it is not the false beliefs generated by unrealistic optimism that are adaptive on this account. “Smoke detectors biased toward false alarms are no doubt preferable to those biased toward the more costly errors (failures to detect actual fires); but that doesn’t mean that a false alarm is a cause for celebration.” (McKay & Dennett 2009, p.506) Rather, the mechanism of producing optimistically biased beliefs itself is supposed to be adaptive. The cases where the tendency to overestimate one’s chances of success lead us to embark on a course of action that actually ends in success are what make the bias ‘earn its keep’. It is when the belief in a positive outcome is vindicated by individuals trying and succeeding that a belief performs its function. This may be the case when we have a true belief that there is a snake present, but also when we have a false belief in the likelihood of the success of an action. For example, an individual’s estimate regarding the question how likely she is to succeed in a goal may be 40%, when an objective risk assessment would put it at 20%. If she tries and succeeds, this belief will have been useful even if it constituted an unrealistic risk assessment.<sup>2</sup>

The second thing which is important to note for purposes of clarification is that the course of action supported by an erroneous belief is an incidence of rational action and decision making. It is hypothesized that false beliefs subserve a rational decision making process (cf. Marshall, Trimmer, Houston, & McNamara, 2013). What this means is that in principle, we could rationally decide to act in the same way even if we did not make the error. Take the above example of the person who thinks she has a 40% chance of success when the chance is in fact 20%. If her acting as though she will be successful is the best course of action because of the cost-benefit ratio, then a realistic

---

<sup>2</sup> Matters are further complicated by the fact that optimism may itself change the likelihood of success and be self-fulfilling by leading to greater goal persistence, or through more indirect effects, for example by decreasing stress and enhancing well-being. I cannot satisfactorily address the question whether self-fulfilling optimism can still be correctly called ‘unrealistic optimism’ within this paper. For further discussion of these issues see (Jefferson, Bortolotti, & Kuzmanovic, 2017; Shepperd, Pogge, & Howell, 2017)

assessment of the outcome will also support that action. This is entailed by the fact that the decision to have a go when potential benefits are high and costs are low is a rational one. Even if an individual were aware that the likelihood of success is only 20%, it would still be rational for her to try to achieve her goal if the expected benefit were high enough and the cost sufficiently low. Presumably, the role that lack of realism plays is a motivational one, in that it makes it easier for us to take calculated risks because we do not perceive them as quite so risky<sup>3</sup>.

### ***3. What is the explanatory aim of evolutionary accounts and to what extent do they yield testable predictions?***

Before evaluating the account outlined above, I would like to address the question what an evolutionary account aims to explain, to what extent it constrains causal explanations of unrealistic optimism, and what counts as evidence for and against it.

Evolutionary explanations of cognitive biases attempt to show why a bias is adaptive and what the bias is or was 'good for'. Specifically, they try to show that in at least one relevant sense, a cognitive bias such as unrealistic optimism is not a form of irrationality, even if it frequently leads to false beliefs. As Nettle (2004) says:

“The model (...) suggests strongly that positive illusions, though irrational in a deontological sense, could be said to be rational in a consequentialist sense, exactly where the benefits of success outweighs the cost of failure, and there is uncertainty about the true probability of success.” (Nettle 2004, p. 205)

Evolutionary accounts are primarily concerned with the positive effect a certain characteristic has had on survival and reproduction, not with the causal mechanisms underlying that characteristic. However, the fact that the effects are the primary focus does not mean that these accounts provide no constraints on accounts of the mechanisms and causes underlying unrealistic optimism. As Cosmides and Tooby point out “the more precisely one can define an adaptive information processing problem -- the "goal" of processing -- the more clearly one can see what a mechanism capable of producing that solution would have to look like” (Cosmides & Tooby, 1997 p. 16). In other

---

<sup>3</sup> Acting on the basis of an unbiased risk assessment would of course only be rational if the individual's cost-benefit analysis is correct. If the individual over or underestimates the cost of a certain action, a rational assessment of the likelihood for success will not be sufficient to support rational actions. I would like to thank an anonymous reviewer for pointing this out.

words, if we know when the effect was beneficial, this gives us constraints as to when we should expect it to occur.

Consequently, both mechanistic and evolutionary accounts make predictions as to what manifestations of the bias we should expect, what should moderate the bias etc. Evolutionary models frequently make predictions for cognitive architecture, by predicting a modular belief forming process which is fast, self-contained, and insensitive to information available to the agent. This has led to the so-called Swiss-army knife view of the human mind as equipped with a number of self-contained subsystems dedicated to solving a certain type of problem. However, unrealistic optimism is unlikely to derive from a modular process, as Haselton and Nettle explain:

“In one culture the relevant domain for positive illusions might be hunting, in another success in college, and in still another, standing in the local community. The cognitive system leaves open the flexibility for the individual to identify those domains in the environment where success yields benefits, and those where failure is costly.”(Haselton and Nettle 2006, p.60)

In this quote, Haselton and Nettle commit themselves to the claim that *individuals* identify the relevant areas of cost and benefit. This model also allows for a very wide array of possible successes, as it includes ‘success in college’. Normally, adaptiveness is specified within the environment of evolutionary adaptedness, and is characterized as adaptive relative to challenges encountered in that specific environment, for example competition for mates or resources. This leaves open the possibility that what was historically adaptive is no longer adaptive in the current environment. The classic example for this is humans’ predilection for sugary and other high energy density foods. Having such a predilection is advantageous in conditions of scarcity, but leads to health problems in societies with an oversupply of food. EMT is unusual in that it does not restrict the domain of unrealistic optimism to scenarios which were of relevance in the environment of evolutionary adaptedness and leaves identifying relevant domains of success to the individual.

According to EMT, unrealistic optimism has the function of enabling low-cost/high benefit actions across all domains of action. Because Haselton and Nettle hypothesize that optimistically biased beliefs result from an open system, this bias is not domain specific. It deals with all possible negative and positive future developments and is not constrained to a certain domain such as cheater detection or food. The only obvious restriction is that we should only expect unrealistic optimism in low cost, high benefit scenarios, rather than as a general expectation. So the theory needs to be evaluated by looking at what we actually know about factors that influence unrealistic optimism.

A successful evolutionary account provides an ultimate explanation of unrealistic optimism, but this will dovetail with our proximate explanation, with what we know about factors affecting unrealistic optimism, and our best causal explanations of unrealistic optimism. It may also yield hypotheses for further research. So, while we should not expect hypotheses regarding adaptiveness or ultimate explanations to give us a causal mechanism for the bias, we should expect some constraints on what causal stories are compatible with the evolutionary account. In the case of optimism, if our best causal accounts make predictions that do not tally with the predictions of the evolutionary hypothesis, this is a problem for the evolutionary hypothesis. This means that we need to find out whether we actually observe a restriction of unrealistic optimism to low cost/ high benefit scenarios and whether there are proximal explanations of factors that lead to unrealistic optimism which are in tension with the error management hypothesis.

I will spend the next section spelling out what constraints the error management hypothesis puts on manifestations of unrealistic optimism and considering whether this is in fact what we find when we look at the empirical evidence. I will show that unrealistic optimism, at least of the kind evidenced in psychological studies, is ill suited to EMT for two main reasons. First, the link between prediction and potential action is so underspecified, that it is impossible to tell how a given prediction fits the error management paradigm. Furthermore, insofar as we can show what the link between prediction and action is, some of the empirical data undermine the hypothesis.

#### ***4. Unrealistic Optimism as Error Management – The Evidence***

On the face of it, the hypothesis that unrealistic optimism is a species of error management leads to some testable predictions, as it predicts that it makes sense to be optimistically biased under specific conditions, namely when potential benefits of achieving an outcome are high and potential costs are low. However, on closer inspection, it turns out that many of the unrealistic expectations scientists have elicited in the lab (Harris, Griffin, & Murray, 2008; Sharot, Korn, & Dolan, 2011; S. E. Taylor & Gollwitzer, 1995; Neil D. Weinstein, 1980) are not clearly linked to action. For example, a person who believes that they are unlikely to suffer from pest-infestation in the house, may be taking or planning to take active steps to prevent pest-infestation, or may be making this prediction without any plans to avoid this undesirable outcome. How unrealistic optimism relates to action matters, because according to EMT, unrealistic optimism earns its keep through the way it affects action, by motivating us to achieve high benefit results. Proponents of the account therefore need to show that rather than just being a slightly unrealistic expectation that bad things are unlikely to happen to



me while good things are likely to happen to me, unrealistic optimism occurs in such a way as to make low cost/high benefit actions more likely. Some might object that the precise link to action need not affect the evolutionary account, as it provides an ultimate explanation for unrealistic optimism, whereas the causal contributors to unrealistic optimism and its effect on action fall under proximate explanations of the way a certain trait works. This would be a mistake, because, as I outlined above, the proximate explanation needs to be compatible with the ultimate explanation. If we don't know whether and how unrealistic optimism will affect action, we cannot tell whether having an unrealistically optimistic belief will be beneficial in terms of attempting to achieve and achieving low cost/high benefit outcomes.

#### *4.1. Unrealistic optimism and classic cases of error management – Some important differences*

According to EMT, the beneficial effects of unrealistic optimism accrue via their effect on action<sup>4</sup>. However, when individuals exhibit unrealistic optimism, what is observed is an unrealistically optimistic belief regarding a future outcome. This belief does not come with a specified associated action. If I report the belief that I am less likely than the average person to be mugged, I may be taking action to prevent this by avoiding certain areas at certain times, carrying mace, practicing my sprint skills, or I may not be doing anything to prevent this, for example because I have an unrealistically positive view of the safeness of my area.

In contrast, classic cases of error management involve a well-established coupling of a specific inaccurate belief with a corresponding action. Take, for example, the case of predator over-perception by foraging animals. “Ethologists coined the ‘life–dinner principle’ to explain why it is better to err on the side of caution when foraging around predators (it is better to miss a meal than lose your life)” (Johnson et al., 2013 p.3). In these instances, the error management explanation posits that the grazing animals perceive something as a threat and run. The same principle applies to cases such as snake over-perception: a certain perception is reliably coupled with avoidance behaviour. These processes are fast and quasi-reflexive.

---

<sup>4</sup>, This does not mean that we should only expect unrealistic optimism where there is potential for action. The advantage of forming unrealistically optimistic beliefs derives from cases where optimism aids action. In cases where no action is possible, whether an individual is optimistic or not makes no difference to the outcome. Therefore, the evolutionary account makes no prediction as to whether we should expect unrealistic optimism in these scenarios. As it turns out, people are less likely to exhibit unrealistic optimism for events that they perceive to be outside their control (Shepperd, Waters, Weinstein, & Klein, 2015). However, the explanation for this may well be that when individuals realize they can't influence the outcome, their predictions are more strongly guided by base rate information. Another possible explanation is that the illusion of control is what (at least partially) underlies unrealistic optimism, and where there is clearly no scope for control, effects of the illusion of control are mitigated cf. (S. E. Taylor & Gollwitzer, 1995)

Why is this important? Because we can only calculate a cost-benefit ratio when evaluating the costs and benefits of predator over-perception based on the knowledge of what the corresponding action type is. In classic cases, we have a reliable coupling between a belief type and an action type, such as avoidance or flight. Sometimes, we even end up with a more basic mechanism that affects our desires directly without going via specific beliefs. So, for example the Garcia effect leads us to avoid foods we have eaten before suffering from stomach problems, even if we don't believe that the food caused the disease and know we were suffering from a virus. Here, there is a reliable effect on behavior, even in the absence of a belief.

The link to action and the resulting cost-benefit ratio is also clear in a frequently cited model which purports to show that overconfidence was often adaptive:

“Overconfidence is advantageous because it encourages individuals to claim resources they could not otherwise win if it came to a conflict (stronger but cautious rivals will sometimes fail to make a claim), and it keeps them from walking away from conflicts they would surely win. These results conform with previous observations (...) that aggressive strategies (such as ‘Hawk’ in Hawk–Dove games) are favored if the advantages of winning exceed the costs of injury, and that overconfident states can outperform others in an agent-based model of conflict” (Johnson & Fowler, 2011p.317)

Here, the link between overconfidence and action (i.e. engage in conflict) is built into the model itself. However, the cost-benefit ratio and the link to action are unclear in many real life cases of unrealistic optimism. For example, when individuals tell us that they are less likely than the average person to be mugged or to suffer from a pest infestation, we do not know what if any steps they are taking steps to prevent these events.

Unrealistically optimistic beliefs are beliefs about what outcomes are likely, but this belief does not come with a ready made observable action plan. If all that can be observed is a generalized expectation of a positive outcome, the link to action is completely underspecified. The phenomenon people observe in the lab and in real-life is that people have unrealistically positive expectations regarding their future. What we observe is only the positive expectation, not a belief of the kind ‘if I do x, then y will happen’. Sometimes unrealistically optimistic beliefs are coupled with a plan for action, sometimes they are not. To complicate matters further, if optimism is the result of an existing plan for action, this inverts the causal direction posited by the error management account. Rather than optimism making us more likely to take action, our plan to take action is the precondition for optimism.<sup>5</sup> The error management account requires that optimism makes action

---

<sup>5</sup> I would like to thank my anonymous reviewer for drawing this point to my attention.

more likely and that the action in question will be a high benefit/low cost action, and that overall, the beneficial effects will outweigh the costs.

More precisely, the account needs to show that unrealistically optimistic expectations are reliably coupled with a disposition towards a certain kind of action<sup>6</sup> which is beneficial to survival and reproduction. This means that the account can fail in two ways – it can be the case that there is a huge variety of actions associated with a certain expectations, some of which are adaptive, some not. But even if there is only one action type associated with a specific expectation, it needs to be the case that this action type is adaptive for the account to be successful. Importantly, the occasional non-adaptive action resulting from unrealistic optimism would not devalidate the account. All that is needed is that unrealistic optimism leads to adaptive behaviour more often than not. In order to test the account, we therefore need to find out whether there is a close link between specific optimistic expectations and specific actions and if so, whether the actions associated with an expectation are adaptive.<sup>7</sup>

In principle, an optimistic expectation could also lead to high cost actions, or to inaction where inaction is costly. Both of these links between belief and action would constitute counter-evidence to EMT. Whether this is the case needs to be established by looking at the empirical findings and as I will show below, both costly action and costly inaction are observed in connection with unrealistic optimism. In order to establish whether unrealistic optimism is sensitive to cost-benefit asymmetries, we need to establish the link between optimism and action. This permits us to evaluate the hypothesis that unrealistic optimism encourages low cost, high benefit action. In the absence of an established link between prediction and action, we will not be able to establish a cost-benefit ratio. I will illustrate these issues by looking at the evidence that we do have regarding the consequences of unrealistic optimism for action.

#### *4.2. Cost -benefit Sensitivity of unrealistic optimism – some counter-evidence*

In the discussion above, I have outlined some reasons why it is hard to establish whether we predominantly exhibit unrealistic optimism in situations where potential costs are low and potential benefits are high. What we need in order to assess EMT is a reliable link between positive expectations and specific actions. The consequences of unrealistic optimism are insufficiently well researched and more work needs to be done to establish the consequences of unrealistic optimism

---

<sup>6</sup> 'Kind of action' should here be understood broadly to encompass different actions geared towards achieving a goal. So in cases of the prevention of ill health, possible actions could be exercising regularly, eating well, getting enough sleep etc.

<sup>7</sup> I would like to thank an anonymous reviewer for suggesting this way of clarifying this point.

(for a recent review paper, see Shepperd et al., 2017). However, there are some empirical results which support the claim that I have made above, which show that unrealistic optimism may be coupled with a number of behaviors which are incompatible with the error management hypothesis. These are 1) that unrealistic optimism may lead to costly behaviour, 2) that being unrealistically optimistic can lead to costly inaction or risk taking and 3) that known moderators of unrealistically optimistic expectations suggest that the proximal causes underlying this bias are at odds with the evolutionary explanation.

1) Costly action and sunk costs: Upon occasion, unrealistic optimism does lead individuals to engage in extremely costly behaviour because they invest increasing amounts of resources into a project and thereby drive up the cost of their action. This can be seen when we consider the scenario where someone tries to set up his own business and is unrealistically optimistic about his chances of success. Depending on how much energy, time and money he puts into the enterprise, the cost of failure can be very high (Makridakis, 2015). One might object that this is a bad example because it is extremely far removed from the challenges an agent encounters in the environment of evolutionary adaptedness. This is true, but as I outlined above, the error management account does not restrict unrealistic optimism as a strategy to specific evolutionarily relevant scenarios. The danger of overinvestment due to unrealistic optimism arises in any kind of scenario where we can only achieve a positive outcome through a long term, ongoing investment of effort. There are well-established psychological mechanisms which take over when we have committed to a certain goal in order to help us persevere in attaining a goal (S. E. Taylor & Gollwitzer, 1995). These may lead to overinvestment in order to achieve an unlikely outcome. In theory, this point can even be applied to Haselton's sexual interest overperception case. If the male persistently tries to win over a specific female because he is convinced of his chances of success, he pays the opportunity cost of not trying to find a different mate because he focuses his time and energy on this case. The longer he continues, the higher the cost. These examples show that the cost of failure is not static, but changes with the level of investment by the individual. So unrealistic optimism can drive up the cost of failure, because it drives up investments. In such cases, optimism does spur individuals into taking action as hypothesized by the error management account, but in doing so, it also drives up the cost of failure to achieve the desired goal<sup>8</sup>.

---

<sup>8</sup> The possibility of sunk costs on its own does not show that overconfidence is non-adaptive. On Gollwitzer's rubicon model of decision making, people ignore risks once they are implementing a decision. This can be costly in individual cases, when an individual invests more than is justified by the pay-off (either because they are unlikely to succeed or because success only carries minor benefits). However, the fact that this kind of behaviour occurs makes it even more important that we be realistic in the decision making phase (cf. Taylor and Gollwitzer 1995).

2. Complacency and risk taking: A further danger is that the fact that an individual thinks a certain outcome is likely will lead them to be complacent, so rather than leading to action directed at achieving a certain outcome, the positive expectation will lead to inaction. In fact, Haselton and Nettle (2006) discuss a putative counter-example of this kind, which is that people are unrealistically optimistic about their risk of health problems (Haselton and Nettle 2016, p.58). This is indeed a common finding in the literature (Bechtel, 2008; N. D. Weinstein, 1982; N D Weinstein, Marcus, & Moser, 2005). As Haselton and Nettle concede, overconfidence regarding one's health risks is not a good trait if and when it leads to unhealthy behavior. However, they interpret the empirical findings as showing that people are overconfident regarding the likelihood that their health-promoting *actions* will have a positive effect. This is a very interesting move, as it illustrates the problem I have been discussing: the error management account requires a correlation between action and outcome, but there isn't always a course of action linked to individuals' expectations. Nettle and Haselton circumvent this problem by making the success of health-promoting action part of the expectation. However, their interpretation is cast into doubt by a number of studies suggesting that the negative effects of unrealistic optimism can result precisely from the fact that unrealistically optimistic individuals do not see the need to take action to avoid ill health (Dillard, McCaul, & Klein, 2006; Kim & Niederdeppe, 2013). So, for example, Kim and Niederdeppe (2013) found that in a college influenza outbreak, unrealistically optimistic individuals were also the ones who reported lower intentions to take preventative measures such as washing their hands<sup>9</sup>. This finding is relevant because it seems to be showing the opposite effect that Nettle and Haselton predict. Rather than people being optimistic about their positive efforts being effective, they are optimistic that their harmful behavior will be ineffective. It thus appears to show that the effect of optimism on health behaviour is by and large negative. However, it is not entirely clear whether the effect of unrealistic optimism on health behaviour is reliably a negative one. In a famous paper, Taylor and colleagues (1992) report a positive effect on health behaviour of AIDS specific unrealistic optimism in HIV positive men. There are a number of factors which complicate the evaluation of this paper, such as the way unrealistic optimism is assessed and a conflation of unrealistic optimism with the illusion of control. Nevertheless, it is worth pointing out that even within one domain, health, findings on the effects of unrealistic optimism are not entirely conclusive. Unrealistic Optimism has also been linked to unfavorable outcomes in the context of college achievements (Robins & Beer, 2001), or unwanted pregnancies (Pons Salvador, Miralles Díaz, & Guillén Salazar, 2010)

---

<sup>9</sup> Similarly, Weinstein et al. (2005) found that smokers estimated their risk of developing lung cancer as lower than that of other smokers and that their risk estimates did not change much with number of cigarettes smoked per day. However, this finding needs to be regarded with caution, as in the case of smokers, addiction may have a distorting effect on individuals' judgments.

Unrealistic optimism about future outcomes may be tied to plans for action in more than one way. Ideally, it will inspire a can-do attitude and inspire us to be proactive but not in ways that are hugely costly to ourselves. But clearly, this is not the only possible effect. If optimism is a factor that makes us less likely to take action to prevent undesirable outcomes because we do not perceive the need, this runs exactly counter to the error management hypothesis. Even if a person does take action on the basis of their unrealistically favorable expectations of future outcomes, the cost of the action may be very high if many resources are invested into achieving the outcome, as we have seen above in the example of people investing all their time and capital in risky business ventures. In sum, the phenomenon of unrealistic optimism is compatible with a number of different effects on action, and there is evidence that in some domains, such as health, the effect on action is a negative one. The existence of negative effects in individual cases or even individual domains is of course not conclusive, as it is the overall effect of unrealistic optimism that decides whether it is adaptive. Nevertheless, the fact that there is no evidence for positive effects at the decision making stage or even a link to decision making, and some evidence against positive effects, undermines the credibility of the view.

3. Moderators and proximal explanations: So far, we have been approaching our evaluation of the EMT hypothesis by considering whether optimism indeed occurs in low cost/high benefit scenarios. Another important piece of information can be found when we look at the factors we know influence the occurrence of unrealistic optimism. If the causal factors known to affect the optimism bias yield optimistic beliefs in low cost high benefit scenarios, this is good news for the account. If not, this counts as disconfirming evidence.

The experimental literature identifies a number of factors that moderate unrealistic optimism: More information and more accountability tend to lead to less optimistic predictions. (Carroll et al. 2006) The former is to be expected, as more information reduces uncertainty, and less uncertainty should lead to more accurate predictions (assuming that the information the agent has is correct). The latter does not sit particularly well with the error management hypothesis, as it implies that optimism is sensitive to factors which have nothing to do with the costs and benefits of these beliefs of different courses of action<sup>10</sup>. If how optimistic we are depends on whether we think we will later have to justify our expectations and not on the fact that we are faced with a high low cost/high benefit scenario, these data would contradict the error management hypothesis. However, it is in

---

<sup>10</sup> While accountability does not change the cost-benefit ratio of the prospective action, maintaining an unrealistically positive belief when one is held accountable may well drive up the cost of the belief in other ways. It may make justifying the belief more costly in terms of cognitive effort or make one seem less credible in front of others. In particular, being asked to justify themselves may make individuals feel like they are in front of superiors, thereby raising the cost of making non-credible estimates.

principle possible that both effects co-occur. People might be more optimistic in low cost/ high benefit scenarios but this optimism may then be further moderated by accountability. However, in contrast to the moderating effect of accountability, there is no compelling evidence that unrealistic optimism is moderated by the presence of the factors posited by EMT. Furthermore, the fact that agents show less optimism when they know they will have to justify their predictions tends to imply that at some level, a more realistic outlook is in fact available to the agent.

Another moderator of unrealistic optimism is perceived frequency of (negative) events (Harris, Griffin, & Murray, 2008). This factor, too, is orthogonal to the cost-benefit ratio of having an optimistic belief and acting on it. One potentially supportive finding for EMT is that the severity of an undesirable event does not lead to increased unrealistic optimism (cf. Harris et al. 2008 but see Taylor et al., 1992). This is important as it undermines the hypothesis that unrealistic optimism is primarily a mechanism of denial in order to protect the individual from unwelcome truths, which would clearly be incompatible with the error management account.

A well-documented factor that increases unrealistic optimism is the ready availability of stereotypes of individuals who are likely to suffer from an adverse event (Harris et al., 2008; Shepperd, Carroll, Grace, & Terri, 2002). (Overweight smokers in stressful jobs get heart-attacks, I am not an overweight smoker, therefore I am unlikely to suffer a heart-attack.) Once again, the representativeness heuristic as a cause of unrealistic optimism does not directly contradict the EMT hypothesis, but implies that unrealistic optimism is sensitive to factors that have nothing to do with the cost-benefit ratio of trying to achieve a certain outcome.

None of the well-known factors moderating unrealistic optimism provide support for EMT, as the factors that affect level of optimism are unrelated to the cost/benefit ratio. The evolutionary hypothesis requires that the factors which influence unrealistic optimism combine to form a pattern of optimistically biased beliefs that is consistent with the hypothesis that optimism occurs in low cost/high benefit scenarios. The evidence so far suggests that a number of motivational and cognitive factors which explain the occurrence of unrealistic optimism explain the data better than the error management account and that at most, the presence of a low cost/high benefit scenario would be a further moderator in addition to existing ones. In order to establish that what we know about the proximate causes of unrealistic optimism is compatible with EMT, it needs to be shown that familiarity, controllability, expectation of feedback etc. do not exhaustively explain the occurrence of unrealistic optimism and that there are further causes which lead to unrealistic optimism being responsive to the cost/benefit ratio of acting to achieve a desired outcome. As it stands, existing proximal explanations do not support the error management hypothesis, because

they do not provide a mechanism which leads to the pattern of optimistic beliefs predicted by EMT . I have shown that this can only be done if we can find reliable expectation-action pairings and that in as far as we do observe these, they speak against EMT.

## 5. Conclusions

Drawing together the arguments and the empirical evidence on unrealistic optimism and its relation to high benefit/ low cost decision making, we can conclude that the error management hypothesis put forward by Haselton and Nettle is unsatisfactory. First, many of the causal factors moderating unrealistic optimism such as the representativeness heuristic, having to justify one's predictions or commonness of the event predicted are orthogonal to cost-benefit considerations. Second, and more troublingly, when we look at instances of unrealistic optimism, we find a) that people are sometimes unrealistically optimistic in cases where the cost of optimism is very high and b) that unrealistic optimism can lead to complacency rather than spurring individuals to take action. Third, and relatedly, the account suffers from the problem that in predictions, the relation between an expected positive outcome and the action, if any, to achieve this outcome is open. This allows for scenarios where the actions that are undertaken on the basis of an optimistic prediction are extremely costly or, alternatively, that the positive expectation removes the felt need for action, as explained above.

The Error Management theorist may object to my conclusion that many of the problem cases that have been discussed contain scenarios where predictions concern outcomes which are far ahead in the future and are of the type that humans would not have historically encountered. If this could be shown, the proponent of EMT could argue that the reason we presently have cases where people are optimistic in low cost/ high benefit scenarios because of the fact that the problems lie too far in the future. Like a fondness for sweet food, unrealistic optimism may have outlasted its usefulness. Note that this is not the way error management theory currently presents the benefits of unrealistic optimism, which they conceive to be very broad in scope, including such goals as success in college (Haselton and Buss 2006). However, it is in principle possible to provide an argument that the relevant predictions in the environment of evolutionary adaptation were of a more immediate nature where both the action and the cost-benefit ratio were obvious. If individuals showed the relevant cost-benefit sensitivity there, a reworked version of an EMT explanation for unrealistic optimism could be retained. In order to test this, we would need experiments which ask participants to predict outcomes which are both close in the future and where the action they could undertake to achieve these outcomes are clearly specified. We could then make a realistic assessment as to whether individuals are sensitive to cost-benefit scenarios in the way a reworked EMT requires.



However, even if we established the right kind of cost/benefit sensitivity, we would then still need to establish what the effect of unrealistic optimism on behavior is. As Shepperd and colleagues argue in a recent paper (Shepperd et al.2017), this is far from clear. What the attempt to establish the success of EMT as an explanation of unrealistic optimism shows is how hard it is to extend an evolutionary explanation which works well in very specific scenarios to very high level, content-unspecific cognitive processes. It is surely correct that when the cost of trying and failing is low and the potential gain of trying and succeeding is high, we should try to achieve that gain. But it is far from obvious that an optimistic bias is the best way of getting us to do so.

#### References:

- Bechtel, W. P. (2008). *Mental Mechanisms*. New York, London: Psychology Press, Taylor and Francis Group.
- Carroll, P., Shepperd, J. A., & Sweeny, K. (2006). Forsaking Optimism. *Review of General Psychology*, *10*(1), 56-73.
- Cosmides, L., & Tooby, J. (1997). Evolutionary Psychology: A primer. Retrieved from <http://www.cep.ucsb.edu/primer.html>
- Dillard, A. J., McCaul, K. D., & Klein, W. M. P. (2006). Unrealistic Optimism in Smokers: Implications for Smoking Myth Endorsement and Self-Protective Motivation. *Journal of Health Communication*, *11*(sup001), 93-102. doi:10.1080/10810730600637343
- Harris, P. R., Griffin, D. W., & Murray, S. (2008). Testing the limits of optimistic bias: Event and person moderators in a multilevel framework. *Journal of Personality and Social Psychology*, *95*(5), 1225-1237. doi:10.1037/a0013315
- 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*(1), 34-47. doi:[http://dx.doi.org/10.1016/S0092-6566\(02\)00529-9](http://dx.doi.org/10.1016/S0092-6566(02)00529-9)
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuys, W. E., & Moore, T. (2009). Adaptive Rationality: An Evolutionary Perspective on Cognitive Bias. *Social Cognition*, *27*(5), 733-763. doi:10.1521/soco.2009.27.5.733
- 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. doi:10.1207/s15327957pspr1001\_3
- Haselton, M. G., Nettle, D., & Murray, D. (2015). The Evolution of Cognitive Bias. In D. Buss (Ed.), *The Handbook of Evolutionary Psychology* (Vol. 2, pp. 968-987): Wiley.
- Jefferson, A., Bortolotti, L., & Kuzmanovic, B. (2017). What is unrealistic optimism? *Consciousness and Cognition*, *50*, 3-11. doi:<http://doi.org/10.1016/j.concog.2016.10.005>
- Johnson, D. D. P., 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 & Evolution*, *28*(8), 474-481. doi:<http://dx.doi.org/10.1016/j.tree.2013.05.014>
- Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*(7364), 317-320. doi:<http://www.nature.com/nature/journal/v477/n7364/abs/nature10384.html#supplementary-information>
- Kim, H. K., & Niederdeppe, J. (2013). Exploring Optimistic Bias and the Integrative Model of Behavioral Prediction in the Context of a Campus Influenza Outbreak. *Journal of Health Communication*, *18*(2), 206-222. doi:10.1080/10810730.2012.688247

- Makridakis, S. (2015). The Costs and Benefits of Positive Illusions. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.00859
- Marshall, J. A. R., Trimmer, P. C., Houston, A. I., & McNamara, J. M. (2013). On evolutionary explanations of cognitive biases. *Trends in Ecology & Evolution, 28*(8), 469-473. doi:http://dx.doi.org/10.1016/j.tree.2013.05.013
- McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences, 32*(06), 493-510. doi:doi:10.1017/S0140525X09990975
- Nettle, D. (2004). Adaptive illusions: optimism, control and human rationality. In D. Evans & P. Cruse (Eds.), *Emotion, Evolution and Rationality*: Oup Oxford.
- Pons Salvador, G., Miralles Díaz, M. T., & Guillén Salazar, F. (2010). The effect of unrealistic optimism in the intention of the use of the condom as method of prevention of non-wanted pregnancies and sexually transmitted diseases. *Anales de Psicología / Annals of Psychology, 26*(2), 8.
- Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80*(2), 340-352. doi:10.1037/0022-3514.80.2.340
- Sedikides, C., Skowronski, J. J., & Gaertner, L. (2004). Self-enhancement and Self-protection Motivation: From the Laboratory to an Evolutionary Context. *Journal of Cultural and Evolutionary Psychology, 2*(1-2), 61-79. doi:10.1556/JCEP.2.2004.1-2.4
- Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nat Neurosci, 14*(11), 1475-1479. doi:http://www.nature.com/neuro/journal/v14/n11/abs/nn.2949.html#supplementary-information
- Shepperd, J. A., Carroll, P., Grace, J., & Terri, M. (2002). Exploring the Causes of Comparative Optimism. *Psychologica Belgica, 42*, 65-98.
- Shepperd, J. A., Waters, E. A., Weinstein, N. D., & Klein, W. M. P. (2015). A Primer on Unrealistic Optimism. *Current Directions in Psychological Science, 24*(3), 232-237. doi:10.1177/0963721414568341
- Shepperd, J. A., Pogge, G., & Howell, J. L. (2017). Assessing the consequences of unrealistic optimism: Challenges and recommendations. *Conscious Cogn, 50*, 69-78. doi:10.1016/j.concog.2016.07.004
- Taylor, S., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193-210.
- Taylor, S. E., Kemeny, M. E., Aspinwall, L. G., Schneider, S. G., Rodriguez, R., & Herbert, M. (1992). Optimism, coping, psychological distress, and high-risk sexual behavior among men at risk for acquired immunodeficiency syndrome (AIDS). *Journal of Personality and Social Psychology, 63*(3), 460-473. doi:10.1037/0022-3514.63.3.460
- Taylor, S., & Brown, J. D. (1994). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin, 116*(1), 21-27.
- Taylor, S. E., & Gollwitzer, P. M. (1995). Effects of mindset on positive illusions. *Journal of Personality and Social Psychology, 69*(2), 213-226. doi:10.1037/0022-3514.69.2.213
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*(5), 806-820. doi:10.1037/0022-3514.39.5.806
- Weinstein, N. D. (1982). Unrealistic optimism about susceptibility to health problems. *J Behav Med, 5*(4), 441-460.
- Weinstein, N. D., Marcus, S. E., & Moser, R. P. (2005). Smokers' unrealistic optimism about their risk. *Tobacco Control, 14*(1), 55-59. doi:10.1136/tc.2004.008375

