WISHFUL THINKING, FAST AND SLOW

A DISSERTATION PRESENTED

BY

DONAL PATRICK CAHILL

TO

THE DEPARTMENT OF PSYCHOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE SUBJECT OF

PSYCHOLOGY

HARVARD UNIVERSITY

CAMBRIDGE, MASSACHUSETTS

MAY 2015

Wishful Thinking, Fast and Slow

ABSTRACT

Psychologists have documented a panoply of beliefs that are sufficiently skewed towards desirability to arouse our suspicion that people believe things in part because they want them to be true (e.g. "above-average" effects (Alicke & Govorun, 2005; Baker & Emery, 1993; Beer & Hughes, 2010; Dunning, Meyerowitz, & Holzberg, 1989; Svenson, 1981; Williams & Gilovich, 2008), unrealistic optimism (Carver, Scheier, & Segerstrom, 2010; Scheier, Carver, & Bridges, 1994; Sharot, Korn, & Dolan, 2011; Weinstein, 1980), and wishful thinking (Aue, Nusbaum, & Cacioppo, 2011; Babad, 1997; Krizan & Windschitl, 2009; Windschitl, Scherer, Smith, & Rose, 2013)). The ostensible irrationality of these motivated biases poses a deep psychological question: how are such biases generated and maintained by a cognitive system that is presumably designed to accurately track reality? Studies that look at the motivated biases and the biased belief updating that may give rise to them tend to employ rich meaningful stimuli covering different targets of belief that are of every day concern: from your health, intelligence, and attractiveness, to your perfidy, academic performance, marital prognosis and driving ability. The use of such stimuli makes it difficult to account for the prior experience and beliefs relevant to such stimuli that a participant brings to the study as well as inadvertently reinforcing a view that motivated biases emerge through rumination upon specific and relatively sophisticated belief content (Lieberman, Ochsner, Gilbert, & Schacter, 2001).

In this dissertation we changed this methodological emphasis. Over the course of the first three experiments, we demonstrate wishful thinking in a semantically sparse, repeated decision-making task about which participants can have no prior expectations, where the components of the task have no personal relevance beyond the experiment, and where they will be required to update their belief about the current state of affairs based upon a repeated and varying diet of desirable and undesirable evidence. We then situated this bias in the dual-process framework of judgment and decision-making by manipulating the time participants take to make their judgment in our task (Experiments 4a and 4b), by manipulating participants' cognitive load (Experiment 5), and by manipulating participants' thinking style—the weight participants put on the contribution from each type of processing—with an essay writing prime (Experiments 6a and 6b). On the whole, the results show that automatic processes alone are sufficient for wishful thinking. Though controlled, Type 2 processing inhibits the bias when induced to play a role, it does not typically contribute to the bias, either antagonistically or complementarily, absent such an inducement. Far from being an occasional, effortful rationalization that thrives on evidential complexity and uncertain costs, the wishful thinking bias we engendered is a simple, biased, belief updating process that operates automatically and beneath our awareness.

Table of Contents

Acknowledgments

A final PhD dissertation will never fully reflect all the joys and travails of the years that produced it. Throughout those years, it has been my privilege to encounter and befriend a breathtaking abundance of wonderfully interesting and interested people who both shared in those joys and supported me through those travails. As I sit down now, it both humbles and saddens me to realize that I cannot hope to adequately thank all of them here.

I owe very deep thanks to my advisor Josh Greene. As well as providing shelter from early graduate school storms, Josh encouraged me to return to my original interests, and enabled me to find my own identity as a psychologist. I also owe a huge debt of gratitude to the many compadres—past and present—of the Moral Cognition Lab. Josh is a truly decent man, and has populated his lab with truly decent, fun, and warm people.

Thanks are also due to the rest of my committee: Dan Gilbert, Mike Norton, and especially Fiery Cushman whose integrity, generosity, and perspicacity I aspire—but can never hope—to emulate. A special mention is due to Susan Carey for warmly shepherding me through a difficult graduate school period. Additionally, this work has benefitted hugely from the suggestions and encouragement of SueYeon Chung, Neir Eshel, Sam Gershman, Amitai Shenhav, Jordan Suchow, and the good folk at the Moral Psychology Research Lab.

Undoubtedly the greatest boon of my time at Harvard has been the companionship of my fellow graduate students, especially those of my cohort (and slightly beyond): Anna, Jen, Phil, Roman, Steven, and Alek. I have depended hugely on their dear friendship and bottomless fund of intellectual and emotional support. Beyond the psychology department, my experience here has been hugely enriched by those with and for whom I served as either a Dudley Fellow or

GSAS Resident Advisor. Beyond the university, I have been guided and inspired by such fellow academic wanderers as Dermott, Michelle, and Jeffrey.

I would thank my parents, Patrick and Roisin, and my brother Patrick, but that would be like thanking the sun. Anything worthwhile I have ever done with myself is ultimately attributable to their assumed and constant love and support.

Chapter 1

The two puzzles of the motivated biases

Reality has the unfortunate habit of confounding our desires. Our plans get scuppered, our hopes get dashed, and our cherished notions skewered, all upon the unyielding rocks of how things actually are. It is no wonder reality is often described as cold and hard, but rarely, if ever, as warm and soft. Of all the things you can do to bring reality more in line with your desires, not acknowledging it would seem to be one of the most futile. Things don't change simply because you believe or disbelieve them, and having an inaccurate view of things is surely a basis for poor decision-making.

And yet psychologists have documented a panoply of beliefs that are sufficiently skewed towards desirability to arouse our suspicion that people do just this: believe things in part because they want them to be true. These *motivated biases*[1] cover beliefs both about our view of ourselves and our behavior ("above-average" effects (Alicke & Govorun, 2005; Baker & Emery, 1993; Beer & Hughes, 2010; Dunning, Meyerowitz, & Holzberg, 1989; Svenson, 1981; Williams & Gilovich, 2008), *self-enhancement* (Epley & Whitchurch, 2008; Guenther & Alicke, 2008; Sedikides, Gaertner, & Toguchi, 2003), and *overconfidence* (Anderson, Brion, Moore, & Kennedy, 2012; Buehler, Griffin, & Ross, 1994; Johnson & Fowler, 2011; Moore & Healy, 2008)) as well as the state of affairs in the word around us (*unrealistic optimism* (Carver, Scheier, & Segerstrom, 2010; Scheier, Carver, & Bridges, 1994; Sharot, Korn, & Dolan, 2011;

[1] It is hard to find a single term to cover instances of bias towards desirable conclusions that doesn't also imply something about the process which gives rise to them, which is what I'm looking for here. "Motivated bias" doesn't quite escape the difficulty, but it is perhaps the best compromise with common use. I prefer the term "convenient bias" which captures the intuition that these biases seem to be the result of a distorting psychological process, but may not necessarily be so. However I shied away from invoking a new term in a field already crowded with them.

Weinstein, 1980), and *wishful thinking* (Aue, Nusbaum, & Cacioppo, 2011; Babad, 1997; Krizan & Windschitl, 2009; Windschitl, Scherer, Smith, & Rose, 2013)).

The ostensible irrationality of these motivated biases poses two deep psychological questions (see Tinbergen, 1963): how are such biases generated and maintained by a cognitive system that is presumably designed to accurately track reality? And what, if anything, is the evolutionary function of such beliefs?

Research on the first question can be summarized as showing that we can't believe something desirable simply because we want to believe it; we must have *some* supporting evidence. But when it comes to collating that evidence, we behave much like an unscrupulous attorney (to borrow an analogy from Ditto, Pizarro, & Tannenbaum, 2009): distorting its importance, searching only until we find the evidence that we want, ignoring evidence which might undermine it, and even fabricating it.

That we weight favorable and unfavorable evidence differently has been demonstrated in various different ways (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998; Dunning, Leuenberger, & Sherman, 1995; Norton, Vandello, & Darley, 2004; Uhlmann & Cohen, 2005). For instance, Kunda (1987) found that female caffeine drinkers were less likely than male caffeine drinkers and female non-caffeine drinkers to rate evidence that caffeine drinking would lead to poor future health outcomes for women as valid. Similarly, Sharot et al. (2011) found that when the likelihood of a negative event occurring to them in the future was lower than participants initially expected, they adjusted their expectation towards the actual value to a greater extent than when the actual likelihood was higher than they expected. Dunning et al. (1989) found that we self-enhance more readily on more ambiguous personal traits where there is a greater amount of both potentially favorable and potentially unfavorable evidence that can be brought to bear, also implying a greater relative weighting of favorable to unfavorable evidence.

Not only do people treat favorable and unfavorable evidence differently, they search for it with differing levels of assiduity (Sanitioso, Kunda, & Fong, 1990; Windschitl et al., 2013). Ditto and Lopez (1992) required participants to self-administer a medical test for a potentially unfavorable diagnosis. The test involved dipping a paper strip in saliva and observing whether it changed color. Those for whom a change in color diagnosed good health waited longer for the change than those for whom a change diagnosed the medical condition. Taking a different route to a similar conclusion, Dawson, Gilovich, and Regan (2002) had participants categorize themselves as high or low in "emotional lability" and then presented them with a Wason Selection Task that required them to test the hypothesis that high or low emotional lability was associated with early death. The appropriate strategy on the Wason Selection Task is to seek disconfirming evidence, yet notoriously, the default behavior is to seek confirming evidence (Evans, 1984). Dawson et al. found that performance improved substantially (i.e. more search for disconfirmation) when the hypothesis was undesirable.

Even when we cannot muster any favorable evidence, we are capable of generating it ourselves. Quattrone & Tversky (1984) had participants submerge their arms in ice water for as long as they could withstand it, both before and after an aerobic exercise session. Some participants were informed that increased cold tolerance after exercise was diagnostic of future good health, while others were told that it was diagnostic of future health problems. The former group kept their arms submerged longer the second time around, while the latter group took their arms out sooner. Presuming that cold tolerance did not actually change as a result of the exercise, it would seem participants desire to obtain a good prognosis affected how long they kept their arms submerged, even though such interference undermined the validity of the diagnosis.

Similarly, Dana, Weber, and Kuang (2007) invited participants to play the Dictator Game, where "Dictators" have the choice of imposing a more or less equitable split of a monetary endowment upon "Receivers". In this game, choosing the less equitable, or selfish split is has implications for one's moral character that are presumably distasteful to at least some of the people who choose it. In one condition, the selfish choice was set as the default option and would be imposed automatically after an unpredictable amount of time had elapsed unless the participant chose the fair option in the meantime. Importantly receivers were unaware whether any inequitable choices they had to suffer were the result of an active decision by the dictator or the time cutoff being exceeded. As such, any dictator who chose to let the time lapse in favor of merely choosing the selfish option was not doing so in order to temper the negative impression the receivers would have of them. Much like in Quattrone and Tversky (1984), the only reason for the choice was to provide themselves with evidence; in this case that they did not mean to choose the selfish option. Dana et al. found significantly more selfish choices in this condition than in the control condition, a substantial proportion of which were due to letting the time lapse.

These and other studies (Batson, 2007; Chance, Norton, Gino, & Ariely, 2011; Sloman, Fernbach, & Hagmayer, 2010) are instances where the sole reason for an action to be selected seems to be to provide desirable evidence to the actor, which can facilitate a desirable conclusion. The remarkable thing is that selecting the actions on this basis doesn't thereby undermine the credibility of the evidence to the actor. It is hard to successfully deceive someone if they are aware that they are being lied to, and given that they selected the action, one imagines that there must have been some glimmer of awareness. This tension between intention and awareness is the crux of conceptual debates in self-deception (Mele, 1997; Mijovic-Prelec & Prelec, 2009), but regardless of where one might stand on those issues, these experiments stand as testament to the sophistication of our methods to believe what we want to believe, and their careful calibration to the fissures in our cognitive architecture.

So we bias our search for evidence, weight the desirable evidence we find more heavily than the undesirable evidence, and even fabricate desirable evidence if needs be. But, to turn to the second puzzle of the motivated biases, why go to such lengths to hold a distorted perception of reality? What could be the evolutionary payoff to such inferential self-sabotage?

One idea is that believing desirable things protects subjective well-being and sustains motivation, thereby enabling people to persist in the pursuit of desirable goals, especially in the face of negative feedback (Carver et al., 2010; Taylor & Brown, 1988). This motivation would be undermined were we to see things in a perfectly objective manner, as evidenced by the lack of these biases in depressed patients (e.g. Alloy & Abramson, 1979; see Taylor & Brown, 1988) and is potentially facilitated by the avoidance of the chronic costly biological responses to stress we would otherwise endure were we to take negative feedback to heart (Taylor, Kemeny, Reed, Bower, & Gruenewald, 2000).

Sure enough, self-enhancers have been shown to have a milder autonomic response to stress (Taylor, Lerner, Sherman, Sage, & McDowell, 2003a) and at least in the academic domain, self-enhancement is associated with emotional equanimity and predicts future performance (Gramzow, Willard, & Mendes, 2008). Furthermore, positive beliefs do seem to lead to positive real world outcomes. Optimism—simply measured as a person's estimate of their own life expectancy—was positively associated with numerous tangible life outcomes of the kind that natural selection may be expected to care about: from working harder and retiring later, to remarrying and saving (Puri & Robinson, 2007).

Haselton and Nettle (2006) see this idea as falling under the umbrella of Error Management Theory (EMT, see also Johnson, Blumstein, Fowler, & Haselton, 2013). EMT Seeks to explain a host of cognitive and behavioral biases by appealing to likely fitness cost asymmetries in errors of judgment. As a toy example, mistaking a snake for a stick is potentially far more costly than mistaking a stick for a snake. Therefore a bias towards treating snake-or-stick candidates as snakes is likely to pay off over evolutionary time. Under the same logic, if, as suggested, optimistic beliefs cause us to pursue fraught long-term goals that a realist might give up on, an optimistic bias would be expected to evolve if the payoff from the occasional success in achieving those goals was sufficiently high so as to cover the cost of all the failed attempts[2].

Johnson and Fowler (2011) extended this evolutionary logic to the arena of social competition. They modeled the evolution of overconfidence—i.e. the overestimation of your own capabilities using a modified hawk-dove game where two agents had to choose whether or not to claim a resource. If both agents made a claim, a conflict ensues for which the agents pay a cost, and which is resolved in accordance with the relative "fighting" ability of the two agents. Johnson and Fowler found that under a given ratio of the benefit of the resource to the cost of a conflict, natural selection favors an overestimation of one's ability to win the conflict.

However, the idea still has a lot of potential counter-evidence to accommodate and theoretical ground to cover. On the empirical side, there is plenty of evidence to suggest that optimism can also be costly, from wars (Johnson & Tierney, 2011) and business practice (Lovallo

[2] A legitimate and interesting question to be asked of EMT is why biased beliefs are necessary as opposed to merely biased action policies (Pinker, 2011).

& Kahneman, 2003) to risky personal decisions (Carver et al., 2010; Puri & Robinson, 2007). Furthermore, the original inspiration for the line of research—that realism is associated with mental health and motivation problems—has been recently undermined: rather than being symptomatic of a depressed mind, realism is just one stop on depression's journey toward pessimism (Carson, Hollon, & Shelton, 2010; Strunk, Lopez, & DeRubeis, 2006; Szu-Ting Fu, Koutstaal, Poon, & Cleare, 2012). Also, the correlational nature of much of this evidence remains a persistent problem in causally relating these biases to health and life outcomes (Carver et al., 2010).

On the theoretical side, as illustrated by Johnson and Fowler (2011), while EMT shows that these biases should be expected under a certain schedule of costs and benefits, it also implies that the opposite biases should be expected under a different schedule of costs and benefits. The substantive question isn't resolved, so much as it is shifted to a question about the actual costs and benefits ratio of the environment these biases evolved in[3].

An alternative, though not necessarily antagonistic evolutionary hypothesis contends that the cost in inferential accuracy is compensated by the benefits they confer in social interaction. Taking it that the best liars are those who believe their own lies, Robert Trivers argues that self-deception—which on this account would encompass what I have been calling the motivated biases—could have been evolutionarily advantageous insofar as it helped in the deception and manipulation of others (Trivers, 2000; von Hippel & Trivers, 2011, 2013). Successful deception is advantageous to the deceiver and costly to the deceived, resulting in an evolutionary arms race between the capability to deceive and the capability to detect deception. Being able to keep

[3] On this note, McKay and Dennett (2009) argue that psychological traits more phylogenetically ancient than the motivated biases (either a profligate stress response or a conservative motivational system depending upon how you read them) are what determine the cost-benefit ratio.

knowledge of the truth away from the mechanisms that may betray it, without unduly compromising our ability to act upon it, would be advantageous in this race.

While the logic is clear, and even compelling, the theory suffers from a dearth of empirical support. However, there is evidence that self-enhancers are better liked than non-self-enhancers (Taylor, Lerner, Sherman, Sage, & McDowell, 2003b), and that overconfidence is associated with a behavioral signature of competence which can ultimately lead to a higher status in group tasks (Anderson et al., 2012). There is also an indirect report from von Hippel and Trivers themselves that participants who knew they would soon have to convince another about the sugar content of a drink, waited longer for evidence that would have been consistent with what they would have to argue for than they did for evidence that would have been inconsistent with what they would have to argue for (von Hippel & Trivers, 2013).

All in all, I hope the foregoing brief review suffices to demonstrate both the depth of the psychological conundrums at the heart of the phenomenon of motivated bias, as well as the empirical and theoretical ingenuity that has been brought to bear upon them. Nonetheless, despite such ingenuity, we feel there are some constraints inherent in the broad approach that has been taken to the issue thus far. As we have seen above, most studies that look at the motivated biases and the biased belief updating that may give rise to them tend to employ rich meaningful stimuli covering different targets of belief that are of every day concern: from your health, intelligence, and attractiveness, to your perfidy, academic performance, marital prognosis and driving ability. The use of such stimuli may inadvertently promote a view of the motivated biases as only operating on specific and relatively sophisticated belief content, e.g. my conception of my traits, qualities, and social worth, or my forecast of future life outcomes. We may be thus biased away from the view of motivated biases as the result of a more simple, domain general process, one that operates without regard to the particular content of the desired belief but merely with regard to its bare desirability.

A further constraint that comes with using semantically rich stimuli is the difficulty in accounting for the prior experience and beliefs relevant to such stimuli that a participant brings to the study, giving rise to the long running worry that what can appear to be a motivated bias may actually be a rational belief given the history of evidence the participant has encountered (Dunning et al., 1995; Kunda, 1987). Such stimuli also make it difficult to flexibly manipulate a motivated bias paradigm, vary it parametrically, describe it with quantitative models and port it to different populations, cultures, and species. All of this is to say that they entail a degree of sacrifice of experimental control and flexibility, something which may ultimately limit our ability to precisely decompose the different processes that may undergird the motivated biases.

Agreeing with Greenwald (2012, p. 99) that “there is nothing so theoretical as a good method”, we aim to bring fresh impetus to the study of motivated bias by changing its methodological emphasis and thereby relieving it of the constraints discussed above. In Chapter 2, over the course of three experiments, we will develop a semantically sparse wishful thinking paradigm about which participants can have no prior expectations, and where they will be required to update their belief about the desirability of the current state of affairs based upon a repeated and varying diet of desirable and undesirable evidence. In Chapter 3, we will then take the first steps in exploiting this paradigm, aiming to situate wishful thinking in the dual-process framework of judgment and decision-making.

Chapter 2

Wishful Thinking in a Simple, Repeated, Decision-Making Task

In this chapter, we aim to investigate wishful thinking in a task where we can control both the prior evidence a participant brings to bear on the task and the provision of desirable and undesirable evidence thereafter. In wanting to assess how people react to more or less favorable news, our aims are in line with Sharot et al. (2011). In their study, they first required participants to estimate the odds of a hypothetical event occurring to them (e.g. car theft, or Parkinson's Disease). They then informed participants of the average probability of those events happening to someone similar to them before having them make the estimation a second time. Participant's second estimates shifted further towards the base rate when the base rate was more desirable than their initial estimate than when the base rate was less desirable. Participants swallowed the good news more readily than the bad news.

This approach has many virtues. It specifically tackles how desirable and undesirable information gets incorporated into current beliefs. The large number and range of different future events makes it less likely that any particular participant is on average actually better off than the population with respect to all of the stimuli. It also means that the updating of desirable and undesirable information can be looked at within each participant, allowing for the control of individual variance in updating in general, i.e. how much a participant might weight new information relative to their previous estimates regardless of desirability. Finally, it also provides a sufficiently repeated measure to enable an investigation of the neural basis of the phenomenon.

However, it remains difficult to separate the role of participants' prior knowledge from a bias in updating per se in contributing to the asymmetric updating. For instance, let's imagine the case of estimating one's chances of contracting gum disease, and let's assume a participant's estimate of a future event happening to them is a combination of their estimate of how likely it is to happen to the average person, and their estimate for how likely it is to happen to them relative to the average person. Now let's imagine, whether justified or not, that I am very confident that I floss more than the average person. The larger my perceived flossing advantage, the more likely I am to rate my chances of getting gum disease as lower than the actual base rate, even when I overestimate the actual base rate. Rating my chances as lower than the actual base rate is what Sharot et al. (2011) categorize as a bad news case. Bad news cases are thus likely to include not only all the cases where I either wildly or mildly underestimate the actual base rate, but even cases where I mildly overestimate the base rate. Good news cases on the other hand, are likely to only include cases where I wildly overestimate the base rate. This means that the average estimation error in bad news cases will be smaller than in good news cases, and as such, even rationally updating my base rate estimate in line with the actual revealed value will result in larger average update values for good news cases than bad news cases. That this may be explaining some of the effect Sharot et al. found, is suggested by the data in Garrett and Sharot (2014) where participants were also asked to estimate the base rates as well as their individual likelihood. When news is classified as good or bad relative to the base rate estimations, the asymmetric updating effect is attenuated.

Nonetheless, it is unlikely to explain the whole effect and the paradigm has proved very fruitful, enabling the assessment of biased updating in depression (Garrett et al., 2014; Korn, Sharot, Walter, Heekeren, & Dolan, 2014), across the life span (Chowdhury, Sharot, Wolfe, Düzel, & Dolan, 2014; Moutsiana et al., 2013), as well as its neural basis (Sharot, Kanai, et al., 2012; Sharot, Guitart-Masip, Korn, Chowdhury, & Dolan, 2012; Sharot et al., 2011).

Eil and Rao (2011) also looked at how people respond to good or bad news—in this case, about one's personal attributes—and were specifically concerned to control for participants' priors. Eil and Rao ran two versions of their game, one where the target attribute was attractiveness—assessed by a speed-dating game—and another where the target attribute was intelligence—assessed by performance on an abbreviated IQ test. Participants were recruited in groups and initially played either the speed dating game or took the IQ test, from which their ranking within the group on that attribute was established.

Next, participants had to estimate their likelihood of being ranked in each particular rank, allowing the experimenters to estimate the distribution of their prior beliefs. Participants were then informed whether they ranked higher or lower than a randomly chosen group member, and had to estimate their ranking again. Participants did this three times in all. Eil and Rao found that while participants updated quite rationally when they found out that they ranked higher than another participant, they failed to update at all when they found out that they ranked lower.

Both Eil and Rao (2011) and Sharot et al. (2011) demonstrate a biased updating of belief depending upon the desirability of the evidence received and, in doing so, extend a challenge to theories that endeavor to explain how we learn and make decisions, which typically assume unbiased updating. Given that the phenomenon seems contrary to a prevailing assumption, it may be thought surprising that motivated bias hasn't been treated in that domain. While some of the reason is no doubt due to a natural lack of cross talk between two different research fields, I suspect some of it is also due to the previous lack of experimental paradigms that both include an inducement to bias, and also involve quantifiable feedback that varies over time. By quantifying the extent of the belief updating, studies like Sharot et al. (2011) and Eil & Rao (2011) have started to bridge the gap between the two areas. But both these studies still use stimuli conceptually richer than the button pressing and primary reinforcers that are the typical components of a learning study. This leaves room for the skeptic to assume the bias is somehow the product of dealing with such “higher-level” stimuli—perhaps it is an innocuous indulgence, or a strategic gambit, or even a rational belief if you assume some particular life history—and thereby avoid the potential conclusion that this is a simple bias in lower-level learning mechanisms.

We aim to tackle this latter idea by adapting the approach of learning theorists (Dickinson, 1980) and testing for wishful thinking in a paradigm where a participant has no relevant prior beliefs to bring to bear on the task, where the components of the task have no personal relevance beyond the experiment, and where we can vary and quantify the evidence provisioned within the task. This will be the project of Chapter 2.

Experiment 1

To achieve the aims described above, we turned to the reversal learning paradigm commonly used in investigations of learning and decision-making (Cools, Clark, Owen, & Robbins, 2002; Hampton, Bossaerts, & O’Doherty, 2006; Hornak et al., 2004). Like all other forms of reward learning tasks, a participant in a reversal learning task is required to select from among a set of actions where each action will return some level of reward. Through trial and error the participant can come to learn which of these actions returns more reward more frequently. The peculiar twist to a reversal learning task is that every so often the reward contingencies of each action will switch, or reverse, so that what was previously the optimal choice is now the suboptimal choice, and what was previously the suboptimal choice becomes the optimal choice. Typically, such a switch in reward contingencies (or *state*) will leave the average expected reward across all actions unchanged. If I was getting 4 pellets of food from Lever A before the switch, I should still expect to get 4 pellets per choice after the switch; I will just have to adjust my response to Lever B to obtain it. One has no reason to prefer being in one state over the other save for the inconvenience of having to relearn the contingencies.

In order to provide an inducement to engage in wishful thinking in our task, we made one state generally rewarding and the other generally punishing. Simply by virtue of being in the rewarding state you could expect a greater amount of reward than the punishing state, regardless of the choice you made. Each reverse still meant a switch in optimal response, so that in order to maximize payout, you still had to adjust your choice. But now there was also reason to hope that you would find yourself in one state rather than another, even though there was nothing you could do about it.

Specifically, our task presented the participant with a choice between a "High-Range Lever", which would return either 150 or -150 tokens, and a "Low-Range Lever", which would return either 120 or -120 tokens (tokens were converted to a monetary bonus at the end of the study). When the task was in the rewarding state, each lever would mostly return a positive amount of tokens, but would occasionally return a negative amount, and when the task was in the punishing state, each lever would mostly return a negative amount of tokens, but would occasionally return a positive amount. As such, even though it was better for a participant's token balance to be in the rewarding state, regardless of what they chose, in order to fully maximize their total reward they needed to choose the optimal lever for that state: the High-Range Lever for the rewarding state, and the Low-Range Lever for the punishing state.

Like in a typical reversal learning task, the best way of doing this is to infer the task's current state from the recent history of returns—if it has mostly been returning positive amounts it is more likely to be in the rewarding state, and if it has mostly been returning negative amounts, it is more likely to be in the punishing state—and then choose the appropriate lever. We expected wishful thinking to bias this inference towards the rewarding state, and thus bias choice towards the High-Range Lever, which is the optimal lever for that state.

The potential returns from both levers in our task were symmetrical about zero. If you had no information about what state the game was in, the expected value of both levers would be the same: zero. But since the range of return from each lever differed, an agent with a risk preference might still be biased towards one or the other lever.

Since people are typically considered to be risk averse (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009; Kahneman & Tversky, 1984), we would expect, if anything, a bias towards the Low-Range Lever, but the possibility remains that there is something about this particular decision-making context that induces a preference for risk. We sought to control for this by running an extra session of our game that retained the same levers with the same payoffs, but dispensed with the idea of a rewarding and punishing state that determined the likelihood of each outcome. Instead the likelihood of a positive outcome on each trial was clearly displayed to the participant, for a range of likelihoods from 0 to 100% (see Method).

This alternative version of the task deconfounds the estimate of the current task state from the likelihood of choosing a particular lever given a current task state, allowing us to measure a participant's bias in response due to their idiosyncratic risk preference, and separate it from any choice bias resulting from a bias in their estimate of the current task state, i.e. wishful thinking.

Method

Participants. 366 participants were recruited from the Amazon Mechanical Turk online labor marketplace (*www.mturk.com*). Participants were paid a show up fee of $2 as well as any money they earned as a bonus during the course of the task. Participation was restricted to those who were resident in the United States and who had a 95% or greater approval rate on their previous tasks on Mechanical Turk. We also restricted participants to those who had not taken any tasks we had run in this research program by filtering participation based upon worker ID codes (https://uniqueturker.myleott.com).

Procedure. Participants were invited to play "The Casino Game" where they could earn a bonus payment on top of their participation fee. Participants first underwent extensive training before taking both the reversal learning task and the risk preference task in a randomly assigned order.

Training. Participants underwent extensive training in order to ensure they were familiar with the structure of the game. The full wording of the training is provided in Appendix A. The main points of the instructions were: the possible outcomes from choosing each lever; that the game can be in either a "Positive Mode" or a "Negative Mode" and what the outcome contingencies were for each mode; what the optimal choice was given a particular mode; that the switches between modes occurred at random; and that though the current mode of the game would be hidden from them, they could try to determine it through reverse inference over the last few outcomes. Participants answered multiple choice questions on each of these main points. If a participant answered any of these questions incorrectly, they were posed again. If the participant answered incorrectly 3 times in a row, then the correct answer was displayed onscreen. The questions served both to reinforce the most important points about the structure of the task, as well as a comprehension filter to enable the potential elimination of participants who were paying insufficient attention to the task.

Participants also played a practice game of 30 trials. This was identical to the reversal learning task except that the current mode of the game was displayed on screen and participants knew they wouldn't earn any bonus from the game. The training session gave participants experience of the game interface. It also gave them experience of the frequency of mode switching and of the different outcomes in each mode.

The training described so far was mostly relevant to the reversal learning task. Prior to taking the risk preference task participants were informed that for that task, the likelihood of a positive outcome would be displayed on screen for each trial. Participants then answered a multiple choice question to ensure they realized the number represented the likelihood of a positive outcome and not a negative outcome.

Reversal learning task. In the reversal learning task, participants had 150 trials to earn as large a bonus payment as they could. Participants had a choice of two "levers" (see Figure 1 for trial display and timing), one of which always returned either 120 tokens or -120 tokens (the Low-Range Lever), and the other of which always returned either 150 tokens or -150 tokens (the High-Range Lever). An image representing each lever was displayed on screen, with each lever labeled by the outcomes associated with it. Participants chose each lever using the left and right arrow keys. After their choice, the chosen lever's handle would depress and the participant would see the cogwheel graphic rotating for 800 ms before the outcome was shown on screen. The outcome was displayed on screen for 1000 ms before the beginning of the next trial.

As described earlier, the game would oscillate between a generally rewarding and generally punishing state. These states were labeled "Positive Mode" and "Negative Mode" in the context of the game. In Positive Mode the participant would receive the positive outcome associated with their chosen lever with a probability of 0.67, and a negative outcome otherwise. In the Negative Mode, these contingencies were reversed. The task alternated between modes with a probability of 0.167 on every trial (i.e. an average of 1 in every 6 trials).

Participants were compensated with 1 cent for every 100 tokens they earned. As such, participants weren't playing for a large amount of money per trial. The game was couched in terms of tokens rather than cents in an attempt to make this fact less salient or obvious to participants. Participants started the game with an endowment of 10000 tokens, and were provided with a running tally on the screen throughout the task. The initial endowment was calculated so as to make it extremely unlikely that a participant could end up with a negative amount of tokens. For that to be the case, assuming that the participant chose the High-Range Lever throughout the task, they would have to receive 109 or more negative outcomes from the 150 trials, the probability of which was 1.3e-08. Though they only stood to win a small amount per trial, both their expected overall bonus and the expected variance of that bonus were substantial relative to their participation fee, and as such we expected the task to be motivating.

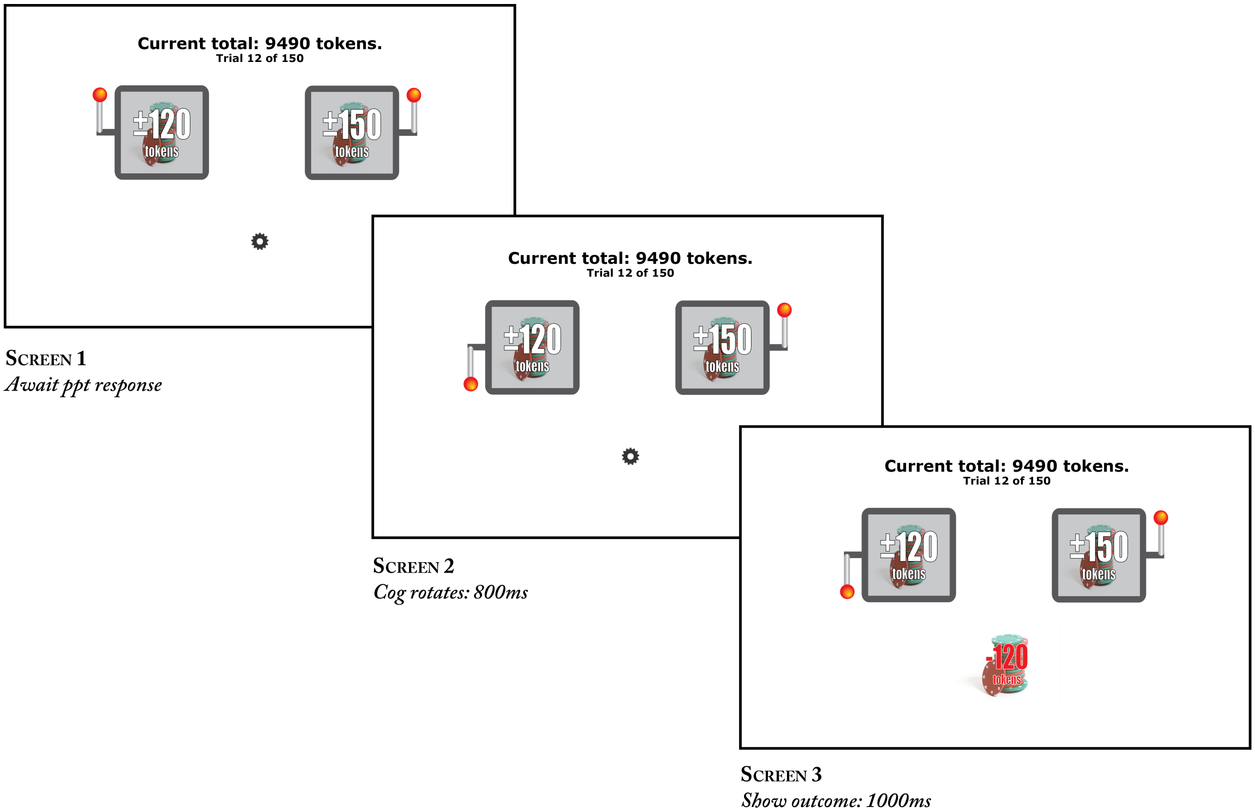


Figure 1. Screenshots and timing for Experiment 1

Risk preference task. In the risk preference task, participants had 53 trials to earn as large a bonus payment as they could. Just like in the reversal learning task, the participant had a choice of a High-Range Lever which would return either 150 or -150 tokens and a Low-Range Lever which would return either 120 or -120 tokens. In this task, the likelihood of a positive outcome on each trial was displayed in the center of the screen between the two levers. A common set of likelihood values was used for all participants. This set included 2 trials each of 0% and 100%, with the remaining 49 values clustering symmetrically and more densely around 50%[4]. These values were shown in random order with no dependency between trials. As such there was no reverse inference available to or required of participants in order to make an optimal choice. They

[4] The full list of likelihoods used (in %): 0, 0, 13, 25, 29, 32, 34, 36, 37, 38, 39, 40, 41, 42, 43, 44, 44, 45, 46, 46, 47, 48, 48, 49, 49, 50, 50, 50, 51, 51, 52, 52, 53, 54, 54, 55, 56, 56, 57, 58, 59, 60, 61, 62, 63, 64, 66, 68, 71, 75, 87, 100, 100

simply had to choose whether or not to take the riskier, i.e. the High-Range Lever given the likelihood of receiving a positive outcome.

Participants started with 2500 tokens and the running tally was displayed on screen. The final tally was included as part of their overall bonus payment.

Exit survey. Participants also completed a brief exit survey where they were asked for feedback regarding their motivation; their understanding of the task and its purpose; their age, gender, ethnicity and education level; and their fiscal and social conservatism. The explicit wording of these items is available in Appendix B.

One open-ended item asked participants whether or not they used a strategy in the game, and if so, what it was. The task was full of uncertainty and quite difficult. We thought it would not be an unreasonable strategy for a participant to set up a decision rule for themselves: e.g. "I only switch when I receive two of the opposite outcome in a row". The costs to inferential accuracy that such a strategy might entail could be offset by the relative ease of implementing it. Our worry was not that such a strategy would be suboptimal, but that such rules would be so easily implemented that they would not allow sufficient wiggle room for any bias—to which the participant would otherwise have been susceptible—to creep in.

We coded participants as strategizers if their response to the item above volunteered an unambiguous (e.g. "2 in a row" not "2 or more in a row") decision rule that applied equally to positive and negative outcomes.

We took the opportunity to append other trait item measures to the task as pilot testing for separate studies. Those measures are not reported here.

Dependent measures.

Reversal learning task. Our aim is to measure any bias towards choosing one lever over the other. It is not immediately apparent what the best way of measuring this is. Under the assumption that on average, participants will have as much reason to choose the Low-Range Lever as choose the High-Range Lever, perhaps the simplest idea would be to simply test whether the High-Range Lever was chosen significantly more often than the Low-Range Lever. However, this would mean ignoring the context in which each particular decision was made—e.g. treating a High-Range choice after a string of positive outcomes the same as a High-Range choice after a more ambiguous sequence of outcomes—which entails sacrificing sensitivity. This method also denies us the ability to examine whether biased decisions are dependent upon evidential context, i.e. are participants more biased when things are good or when things are bad, or are they more biased when the situation is ambiguous? Finally, this method would make it more difficult to assess individual differences in bias. Different participants are likely to experience different levels of fortune during the task than others—in the sense of spending more time in Positive Mode—and simply looking at their mean choice of lever would confound this fortune with their bias.

To circumvent these difficulties we took two different, but related approaches with the hope and expectation that each would confirm the other. Both these approaches essentially aimed to model the effect of the evidence the participant had received upon their choice, thereby separating bias from fortune while also enabling us to look at bias across the range of evidence.

Mirrored proportions bias measure. Our first method compares how a participant responds to evidence that the game is in Negative Mode—i.e. a recent preponderance of negative outcomes—with how they respond to the equivalent evidence that the game is in Positive Mode. To do this, we bin each trial according to the history of outcomes that preceded it and determine the proportion of High-Range Lever choices by each participant for each bin. Each bin will have its mirror image: another bin defined by the opposite sequence of positive and negative outcomes. An unbiased participant would be as likely to choose the High-Range Lever in one bin as they are unlikely to choose it in the bin's mirror image. By comparing the proportion of High-Range choices in each mirrored pair, we can assess each participant's bias.

Since we are calculating over proportions of lever choice in each bin, our measure should not be affected by how often a participant visits each bin, i.e. how lucky they have been during the task. Furthermore, by comparing mirror-imaged bins we effectively cancel out variation in the strength of the relationship between outcome and response. That is, each participant is likely to vary in how enthusiastically or tepidly they respond to recent outcomes (in the sense of likelihood to choose one lever over the other), but regardless of how strong this relationship is, it should be equally strong for positive and negative evidence in an unbiased participant. By comparing responses across symmetrically defined bins, we effectively use each participant as their own control.

We have a choice in how deep an outcome history we want to use to define each bin. In using longer sequences we trade specificity for sensitivity; the more outcomes we use to define each bin, the more bins we will end up with but the fewer trials per participant we will have per bin. We calculated our measure for all of one, two and three trial history depths, but set a two trial depth as our target going into the analysis, seeing it as the safest compromise.

We calculated our measure per participant by subtracting the proportion of Low-Range Lever choices in one bin from the proportion of High-Range Lever choices in the mirror bin (it makes no difference which bin of the pair is counted as the mirror). Resulting values greater than zero indicate a bias towards the High-Range Lever and values less than zero indicate a bias towards the Low-Range Lever. Thus we get a value for each pair of trial bins per participant. Since our expectation is that any bias that might emerge is more likely to show up in the ambiguous cases, we chose to use the most ambiguous bin pair as the value that best represents a participant's bias. In the two trial history depth analysis, this refers to the pair of bins defined by either a positive-negative or a negative-positive sequence of outcomes (hereafter the *YX* bin).

The advantage of this measure is that it is relatively simple and intuitive. But while it serves as a good approximation of behavior in our task, it is not as sensitive as it could be. Each proportion is calculated from a differing number of trials. This means that each proportion is a more or less reliable estimate of lever choice in that bin. By simply using proportions as out measure, we fail to account for this structure in our data, and thus sacrifice sensitivity. A further disadvantage is that by collapsing across oppositely valenced bins we are not able to determine whether any bias we elicit occurred selectively when times were good or when times were bad.

N-Back Outcome Models. Our second approach also modeled lever choice as a function of previous outcomes. Rather than averaging over the choices a participant made, in order to capture the full structure of our data, we needed to estimate the dependency between choices that arises as a result of their being made by the same participant. Parametric repeated measure approaches were not available to us given that we had an imbalanced design—participants were not guaranteed to receive an equal amount of positive and negative outcomes—and a binary dependent variable (lever choice). To get around this we used a mixed model logistic regression approach (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013; Skatova, Chan, & Daw, 2013; Winter, 2013), modeling participants as random effects and outcomes as fixed effects, with lever choice as the dependent variable.

The intercept value of such a model tells us the log odds ratio of choosing the High-Range Lever when the predictor variables are set to zero. Thus we coded our outcome variables with 1 as a positive outcome and -1 as a negative outcome so that the intercept of the model would constitute an estimate of the likelihood of choosing the High-Range Lever independent of the effect of previous outcomes. An unbiased participant would be as likely to choose the High-Range Lever as the Low-Range lever all else being equal. This would equate to an odds ratio of 1, the log of which equals 0. As such, an intercept value greater than zero represents a bias towards the High-Range Lever, and a value less than zero represents a bias towards the Low-Range Lever.

As before, we had to decide how many outcomes back to include as predictors in our analysis. The more we use, the more variance is likely to be explained by the model, but at a cost of greater model complexity, with attendant difficulties in both estimating and interpreting the model. We settled on using both the two most recent outcomes for our main analysis. However we also planned to use both only the most recent trial and the last three trials in subsequent exploratory analysis in order to gain a more complete picture of how this analysis approach behaved with respect to this paradigm.

We planned to include the interactions of outcome predictors in our models since this would allow us to test for potential ambiguity effects on bias. For instance, when using the two most recent outcomes as predictors, since outcomes were coded with 1 and -1, the interaction of the two previous outcomes would be negatively signed for trials where the previous two outcomes differed, and positively signed for trials where the previous two outcomes were the same.

Model analyses were performed a mixed effects analysis using R statistical software (R Development Core Team, 2014) with the lme4 package (Bates et al., 2014). In doing this, our approach was guided by Barr et al. (2013), who advocate "maximally" modeling the random effects structure. Thus for every analysis we planned to begin by including both random intercepts and the random slopes for each of our predictor variables. But given that we were using a logistic regression, we anticipated difficulties in getting complex models to converge. In such an eventuality we planned to take the advice of Barr et al. (2013) to eliminate the least important components—in terms of relevance to the question we were asking—of the random effects structure in turn until we found a model that converged. We use a simple normal approximation to determine p-values, considering it the most appropriate approach for logistic mixed effects regression models with large samples (Faraway, 2005).

Risk preference task. Risk preference was measured by regressing lever choice on the positive outcome likelihood of each trial. Since this was a repeated measure, we again used a mixed effects modeling approach (for exact details see Results). This analysis allowed us to estimate risk preference at the population level. To estimate risk preference for each participant we extracted the by-participant random slope estimates for likelihood that were estimated as part of that analysis.

Results

9 participants suffered technical problems and a further 27 failed to complete the task for unknown reasons, leaving 330 participants for analysis (M_{age} = 31.8, 48% female). Though we collected measures of comprehension and engagement, no participants were excluded on this basis. Self-reported comprehension and motivation was high (M_{comp} = 6.75; M_{try} = 6.56; both measures scored on a 7 point scale), and participant's performance on the comprehension check measures was strong ($M_{incorrect}$ = 0.28). Indeed, we see the same pattern across all the studies in this dissertation.

Mirrored proportions bias measure. Participants showed a bias towards choosing the High-Range Lever in the ambiguous trials (YX) bin (M_{YX} = 0.07, $t(329)$ = 3.52, p < 0.001). An exploratory hierarchical regression of our bias measure on both XX and YX bins with participant entered as a random factor showed that YX trials showed more High-Range Lever bias than XX trials (β = 0.04, $t(329)$ = 2.69, p = 0.007) and that XX trials were themselves not significantly biased (β = 0.03, $t(438.5)$ = 1.89, p = 0.06).

Outcome Model Measure. A multi-level mixed effects logistic regression analysis was employed to model the effect of the outcomes from the previous two trials (*Outcome t-1* and *Outcome t-2*) on Lever Choice (hereafter the *Two-Back Outcome Model*). Participant was entered as a random factor with both random intercepts and random slopes estimated with for both outcome variables. Variables were coded as described in Dependent Measures. The model returned a significantly positive intercept (β = 0.17, z = 2.93, p = 0.003), implying that participants were biased towards the High-Range Lever independent of the effect of previous outcomes. Previous outcomes predicted likelihood to choose the High-Range Lever, though the interaction of the previous outcome with the outcome two trials preceding did not predict the choice of lever (see Table 1[5]), implying that there was no difference in response due to the ambiguity of the previous outcomes.

The same pattern of results was observed when we modeled response using the previous three trial outcomes (see Table 2). In this case, random slopes were not estimated due to convergence difficulties.

[5] For ease of formatting and so as to not overly disrupt the flow of text, all tables and figures will be presented at the end of the section in which they are referred to.

Risk Preference Task. To analyze risk preference in the risk preference task we again used a mixed effects model approach to assess the relationship between the likelihood to receive a positive outcome (*likelihood*) on each trial and lever choice. Likelihood was centered at the equal likelihood point (0.5) and entered as the sole fixed effect, with the random effects structure incorporating both random intercepts for participant as well as a by-participant slope for likelihood. Participants were significantly risk averse as indicated by a negative intercept estimate ($\beta = -0.23$, $z = -2.13$, $p = 0.033$), i.e. participants were more likely to choose the Low-Range Lever when equally likely to receive a positive or negative outcome. Interestingly this estimate does not correspond with their average choice on the specific trials where they were equally likely to receive positive or negative outcomes (see Figure 3), implying that though participants may be risk seeking on those particular trials, they were on the whole risk averse on all trials either side of these trials. That is, they were less likely to choose the High-Range Lever on trials were they were likely to receive a positive outcome than they were to choose the Low-Range Lever on trials were they were likely to receive a negative outcome.

Reversal learning task performance controlling for risk preference. To test whether a preference for risk seeking could explain the bias we found in the main task, we took the random intercept estimates from our mixed effects analysis of the risk preference task and entered them as an additional regressor in our Two Back Outcome Model. Convergence issues meant that we had to drop the by-participants random slope estimate for both outcomes. The results are shown in Table 3 and indicate that while risk preference was a significant predictor of lever choice in the task ($\beta = 0.17$, $z = 5.96$, $p < 0.001$), participants still showed a bias towards the High-Range Lever ($\beta = 0.19$, $z = 3.99$, $p < 0.001$) independent of that influence.

Order Effect. We checked for an effect of task order by adding task order as a fixed effect to the Two-Back Outcome Model, dropping the random slopes from the random effects structure to avoid convergence difficulties. The results showed no effect of task order ($\beta = 0.05$, $z = 0.55$, $p = 0.58$).

Reaction Time. The Reaction Time results are shown in Figure 4. We assessed the effects of lever choice—coded as -1 for Low-Range and 1 for High-Range—and previous outcome on log-transformed reaction time in the main task by using a mixed effects model with the full interaction of lever choice and the previous two outcomes entered as fixed effects. Random intercepts were estimated for participant, as well as a by-participant random slope for lever choice. Since this was not a logistic mixed effects regression model, unlike most other analyses here, p-values were calculated with a Satterthwaite approximation for the degrees of freedom using the lmerTest R software package (Kuznetsova, Brockhoff, & Christensen, 2014). The results are shown in Table 4.

Bias across trials. In order to test whether the bias in the task was due to participant's prior expectations about how they would fare in the task, we looked at the extent of the bias across trials, hypothesizing that if the bias was solely due to participant priors, then they should converge towards no bias as the experiment progressed. We tested for effects of trial by including trial as a linear predictor of lever choice in a mixed effects model. To do this we first centered trial at the midpoint. We also squared this predictor to enable us to test for curvilinear effects across the task. We included both of these predictors in the model along with previous outcome as fixed effects, and modeled the random effects structure with random intercepts for participant and random slopes for the effect of both trial terms on participant. The results showed a significant positive quadratic trial effect ($\beta = 0.09$, $z = 4.14$, $p < 0.001$) implying that the bias initially dropped off but subsequently rose again after the midpoint. The full results are shown in Table 5.

Demographics and task engagement measures. We also assessed the effect of the demographic data we gathered on lever choice by including age, gender, education, income, and social and fiscal conservatism in a mixed effects logistic regression model. We also included previous outcome as a fixed effect and interacted it with each of the demographic variables. Random intercepts were estimated by participant. The results show a significant effect of gender ($\beta = 0.1$, $z = 2.22$, $p = 0.027$, full results are shown in Appendix C). We also modeled the effect of our comprehension and engagement measures on lever choice by adding them to our Two-Back Outcome Model (dropping the random slopes from the random effects structure to enable convergence). No significant effects were found, as shown in Appendix D.

Table 1

Experiment 1: Effect of previous trial outcomes on likelihood to choose High-Range Lever.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.17	0.06	2.93	0.003
Outcome t-1	1.29	0.06	22.16	< 0.001
Outcome t-2	0.89	0.04	20.53	< 0.001
Ot-1 x Ot-2	-0.01	0.02	-0.62	0.537

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

Table 2

Experiment 1: Effect of previous three trial outcomes on likelihood to choose High-Range Lever

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.15	0.05	3.08	0.002
Outcome t-1	0.96	0.01	84.03	< 0.001
Outcome t-2	0.62	0.01	55.45	< 0.001
Outcome t-3	0.28	0.01	25.46	< 0.001
Ot-1 x Ot-2	-0.02	0.01	-2.43	0.015
Ot-1 x Ot-2	-0.01	0.01	-0.65	0.517
Ot-1 x Ot-2	-0.02	0.01	-1.71	0.087
Ot-1 x Ot-2 x Ot-3	--0.04	0.01	-3.28	0.001

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

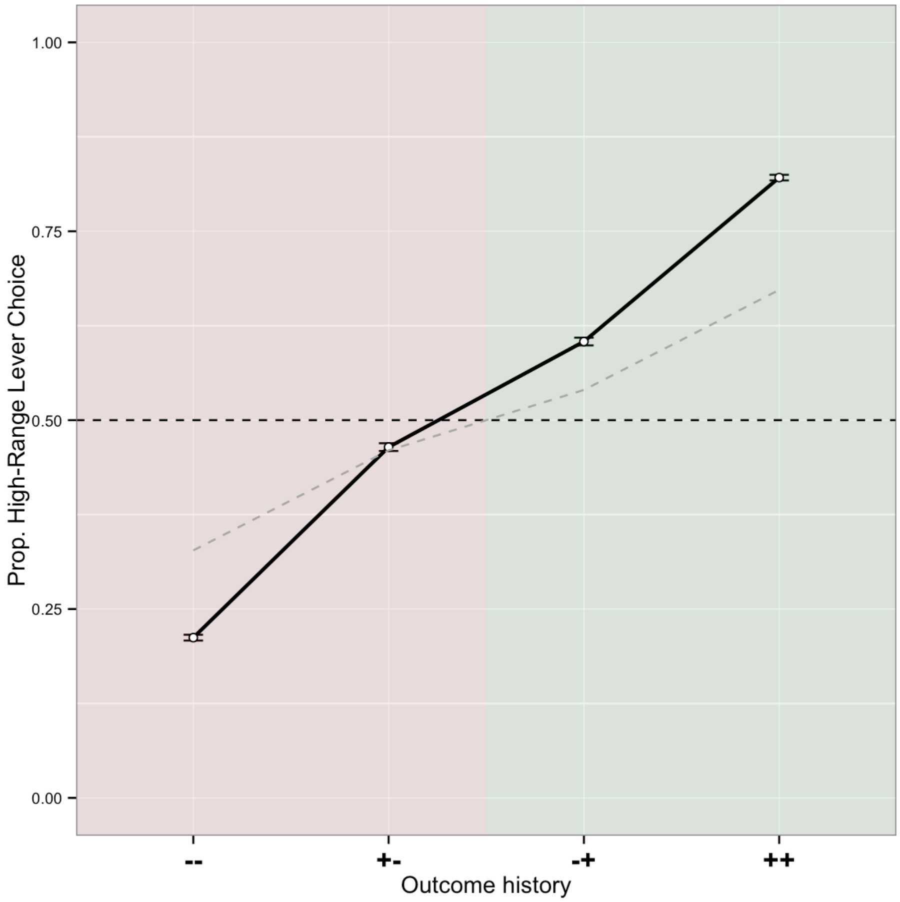


Figure 2. Choice behavior for Experiment 1. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of High-Range Lever choices for each bin, with error bars calculated using the within-participant correction outlined in Morey (2008). The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose the High-Range Lever (green) or the Low-Range Lever (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line.

Table 3

Experiment 1: Effect of previous trial outcomes and risk preference on likelihood to choose High-Range Lever

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.19	0.05	3.99	< 0.001
Outcome t-1	0.96	0.01	84.39	< 0.001
Outcome t-2	0.64	0.01	57.14	< 0.001
Risk Pref.	0.17	0.03	5.96	< 0.001
Ot-1 x Ot-2	-0.03	0.01	-2.53	0.01

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

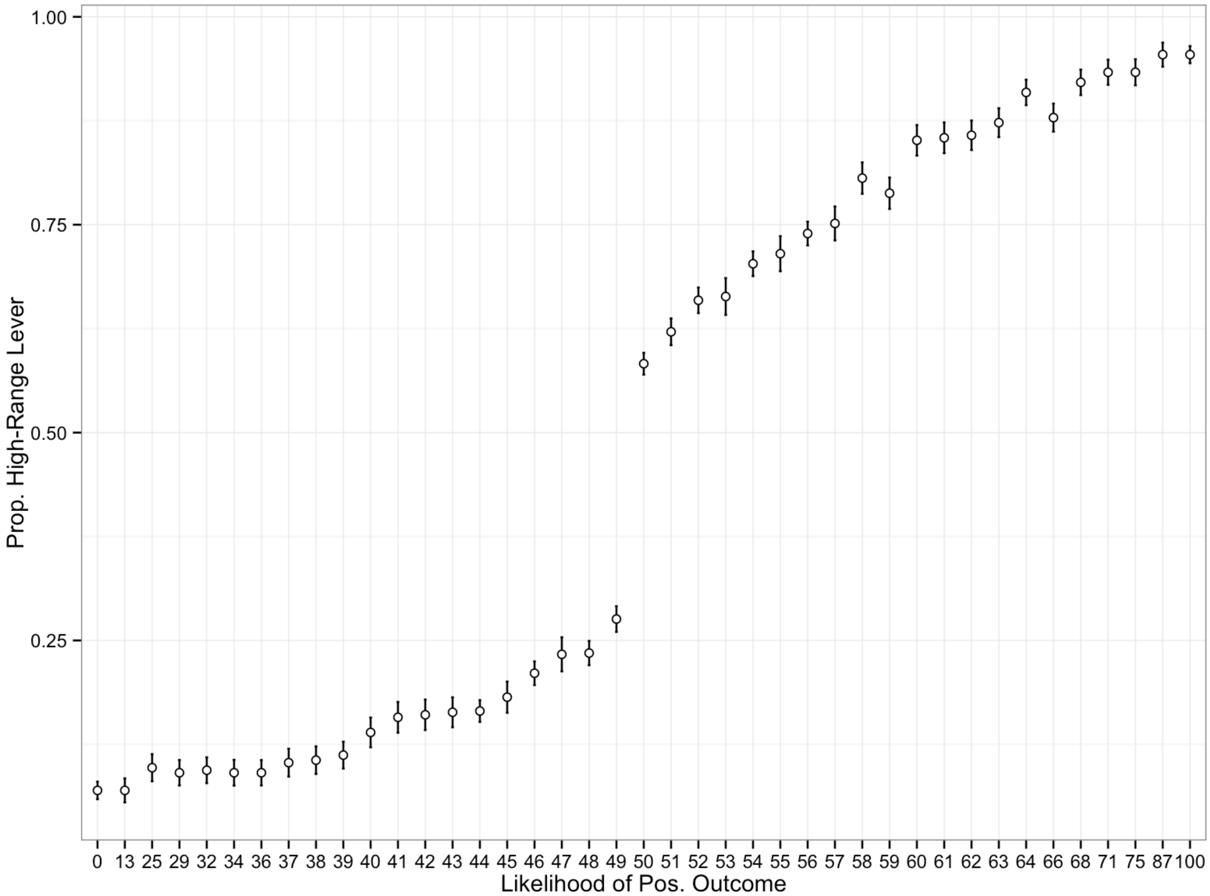


Figure 3. Choice behavior in the risk preference task. for Experiment 1. Error bars calculated using the within-participant correction outlined in Morey (2008).

Table 4

Experiment 1: Effect of previous trial outcomes and lever choice on reaction time.

Fixed Effects	*Estimate*	*Std. Error*	*t value*	*p*
Intercept	6.11	0.03	214.781	< 0.001
Choice	0.008	0.005	1.71	0.088
Outcome t-1	0.038	0.004	10.07	< 0.001
Outcome t-2	0.001	0.004	0.2	0.837
Choice x Ot-1	-0.017	0.004	-4.39	< 0.001
Choice x Ot-2	0.01	0.004	2.56	0.01
Ot-1 x Ot-2	-0.002	0.004	-0.59	0.556
Ch. x Ot-1 x Ot-2	-0.001	0.004	-0.15	0.881

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as -1 for Low-Range Lever and 1 for High-Range Lever. Reaction times were log transformed.

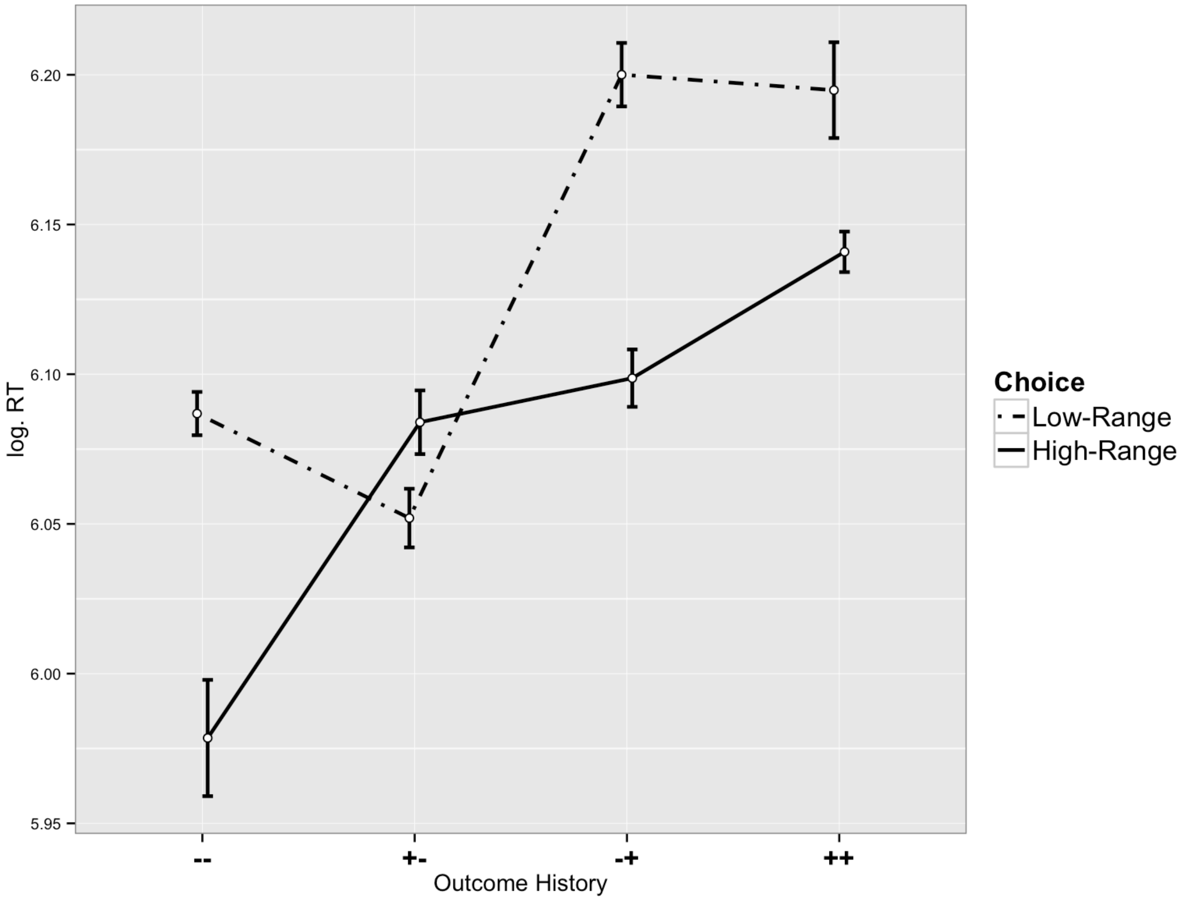


Figure 4. Reaction times for Experiment 1. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in (Morey, 2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Table 5

Experiment 1: Effect of previous trial outcome and trial on likelihood to choose High-Range Lever.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.06	0.06	1.16	0.247
Outcome t-1	0.93	0.02	56.41	< 0.001
Trial	-0.04	0.02	-1.6	0.109
Trial Quadratic	0.09	0.02	4.14	< 0.001
Ot-1 x Trial	-0.02	0.01	-1.69	0.092
Ot-1 x Trial Q.	0.02	0.01	1.8	0.071

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Trial regressors centered at session midpoint. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

Discussion

Experiment 1 showed a wishful thinking bias that was independent of participant's risk preference. The bias did not disappear as the task went on, ruling out the possibility that the bias is explained by any priors the participant may have had about the task at the beginning of the experiment. As such, we have broadly succeeded in our aim of eliciting wishful thinking in a highly controlled task.

However, some specifics of our results merit further reflection. For one thing, our mirrored proportions measure implied that the bias was stronger for ambiguous trials but the outcome model measure showed no such effect. What explains the discrepancy?

We suspect this is an artifact of how the mirrored proportions measure is calculated. We calculated the proportion of High-Range Lever choices for each trial bin. Each value we obtain will inevitably be partly determined by sampling error. That sampling error is more likely to be biased the closer the values come to either the floor (never choosing the High-Range Lever) or the ceiling (always choosing the High-Range Lever). In those cases, the clipping of the sampling error at one or other extreme will bias the proportion away from the floor or the ceiling. Since there is a general bias towards the High-Range Lever in our task, there are likely to be more ceiling cases than floor cases, and therefore a general measurement bias away from the High-Range Lever. This measurement bias suppresses the measurement of the true bias in our task, and this suppression will be greater the more likely values are to be closer to the extremes, i.e. on the unambiguous trials. As such, the apparent effect of ambiguity that we see in our mirrored proportions measure is really a suppression of the bias on the unambiguous trials.

Since the outcome model measure assesses participant variability at the trial level, this is not a concern on this measure, and as a result becomes our more trusted measure and our primary dependent measure for the rest of the studies in this project.

All of this means that there was no effect of ambiguity on our bias, a finding that is contrary to findings in work on self-enhancement and self-deception (Dunning et al., 1989; Mazar, Amir, & Ariely, 2008; Sloman et al., 2010). This could be simply because in our task even the unambiguous trials in are relatively ambiguous. Alternatively, it could be that risk preference is interfering with our ability to measure it—a conjecture that is given support by the result in Table 3 that shows an ambiguity effect once risk preference is controlled for. Ideally, we would be able to measure wishful thinking without interference from risk preference to test this more directly.

Our reaction time data are also puzzling. Our mixed effects analysis showed a significant effect of the most recent outcome such that people are slower to respond when times are good. It is unclear why this should be. This effect isn't explained by looking at the reaction time data from the risk preference study where there was no main effect of likelihood to be rewarded.

We also saw an interaction of previous outcome and lever choice such that participants are slower to choose the less appropriate lever given the recently received outcome, i.e. slower to choose the Low-Range Lever after a rewarding outcome and vice versa. This is the opposite interaction to that found in the risk preference task where reaction times were quicker when choosing the sub-optimal lever. This difference can probably be attributed to the uncertainty in the main task—where the current task state is uncertain, and relatively longer reaction times for suboptimal choices are probably due to indecisiveness. This uncertainty doesn't exist in the risk preference task, and the faster reaction times for suboptimal choices are probably due to simple inattention or mindlessness.

We also saw an effect of gender such that males were more biased than females. This effect held when we controlled for risk preference ($\beta = 0.1$, $Z = 2.19$, $p = 0.028$), implying that the effect of gender reflects a gender difference in wishful thinking. This is a curious and unexpected result that merits further investigation.

Given that we assessed wishful thinking as a bias in lever choice, and given that each lever returned a reward or punishment, it might be wondered whether our bias can be explained by the history of reinforcement the participant experienced. Various reward-learning models can be brought to bear on our data. Niv, Edlund, Dayan, and O'Doherty (2012) found evidence for different learning rates for rewards than for punishments in reward learning. If learning rates were higher for rewards than for punishments, this might explain a bias towards the High-Range Lever. However, this is unlikely to be the case. In their sample, Niv et al. (2012) found that learning rates were actually higher for punishments than rewards. Indeed, they offered this as a potential explanation of risk aversion. There is no reason to suspect that the balance of learning rates would be different in our task, especially given that we also found participants to be risk averse, and that that risk preference didn't explain our wishful thinking bias.

Perhaps a better candidate for potentially explaining our data is found in the model of mood instability devised by Eldar and Niv (2015). In their model mood is defined as the average of recent prediction errors: if things are better than expected, my mood gets better; if things are worse, my mood deteriorates. Furthermore, mood modulates perceived reward. That is, if I am in a good mood, the same objective reward is more valuable to me than if I was in a bad mood.

Applied to our task, this model would predict that a participant's mood would oscillate as our task's state oscillates. This would mean that the difference in reward between the levers would be exaggerated in Positive Mode—because I'm in a good mood—and attenuated in Negative Mode—because I'm in a bad mood. This difference in reward differences would correspond to a greater likelihood to choose the High-Range Lever when times were good, relative to the likelihood of choosing the Low-Range Lever when times were bad, which is essentially how we are defining a wishful thinking bias.

This is an interesting alternative explanation, especially given the association between wishful thinking and subjective well being described in Chapter 1, and might be taken to imply that the wishful thinking bias is not so much a functional feature of the mind as it is a mere by-product of the operations of reward learning mechanisms. In order to test this idea, it would be helpful to rule out the role of reinforcement.

Finally, we note that even though it didn't entirely explain the bias we found, risk preference did still predict likelihood to choose the High-Range Lever. This is perhaps unsurprising, but also somewhat unfortunate since it could prove to be problematic for manipulating wishful thinking using this paradigm. Any manipulation designed to affect wishful thinking would be in danger of affecting risk preference too.

Given all of the above: the difficulty of separating risk preference from wishful thinking when interpreting ambiguity, reaction time, and demographic effects; the potential to explain the data as arising from reinforcement history; and the difficulty in using the paradigm for future manipulations given the potential effects of any such manipulation on risk preference, we sought to redesign our paradigm in Experiment 2 so as to remove the potential confounds of risk preference and reinforcement history. Nonetheless, Experiment 1 has proven enlightening in providing an initial demonstration of wishful thinking in a semantically sparse and highly controlled task, elucidating the relationship between risk preference and wishful thinking, and providing a paradigm that may be amenable to both testing on non-human animals and modeling the interaction of wishful thinking and reward learning.

Experiment 2

Our aim for Experiment 2 was to alter our paradigm so that we could remove the confound of risk preference entirely, and also divorce our measure of wishful thinking from reinforcement history in case it was reinforcement history that actually explained our effect.

In Experiment 1 we expected our participants to assess the current task state and then make a reward maximizing choice given that assessment. In measuring wishful thinking we are solely interested in that assessment of task state, but we could only measure it indirectly by looking at the choices they made, and of course these choices were also potentially influenced by risk preference and reinforcement history. As such, in Experiment 2 we sought to measure wishful thinking more directly. We did this by simply asking them what their assessment of the task state is.

In Experiment 2 participants only had one lever to choose. Pulling this lever would result in a positive or negative outcome, the likelihood of which was determined by whether or not the game was in the rewarding or punishing state just like in Experiment 1. Indeed the outcome contingencies and state transition frequency were all retained from Experiment 1. After the participant witnessed the outcome, we asked them to judge whether the machine was in positive mode or negative mode, with a bonus amount offered for accuracy. Importantly, participants didn’t receive feedback on their accuracy until after the task was over, so there was no reinforcement on these judgments. Equally as important, there is no difference in risk associated with this judgment; you stand to win just as much for correctly judging that you are in either mode. With this simple modification, we are now in a position to measure wishful thinking—operationalized as a bias in this mode judgment—directly, without any response bias confounds.

Method

Participants. We recruited 254 participants from the Amazon Mechanical Turk online labor marketplace (*www.mturk.com*), paying them 70 cents for their participation and any money they earned as a bonus during the course of the task. All other recruitment details were the same as Experiment 1.

Procedure. Participants were invited to play "The Casino Game" where they would make a series of choices for a potential monetary reward.

Training. Training began with a practice session of 5 trials. For each trial in this first practice session, participants simply had to press the down arrow on their keyboard to pull the lever and witness the outcome. The machine returned a reward or punishment image at random, and these rewards and punishments did not contribute to their bonus. The purpose of this first practice session was simply to familiarize the participants with the interface of the game, so that subsequent instructions could be more easily put in context.

Following the first practice session, we introduced participants to the idea that the machine could be in one of two modes: Positive Mode and Negative Mode, where Positive Mode mostly, but not always, returned positive outcomes and Negative Mode mostly, but not always, returned negative outcomes. We also told them that we would ask them to make their best guess as to what mode the machine was in after each outcome, and that they could increase their bonus by guessing correctly. We specifically stated that while the bonus they earned from the machine was due to luck—just like at a casino—the bonus they earned from these judgments was up to them.

Participants then had a second practice session of 20 trials where they not only pressed the down arrow to pull the lever, but also had to press either the left or right arrow to indicate which mode they thought the machine was in. For this practice session, and unlike in the main task, the present mode of the game was displayed on screen throughout the session. The purpose of this session was to both familiarize the participants with the full task, as well as giving them some experience of the frequency of mode transitions and the payoff contingencies in each mode.

After the second practice session, we instructed participants, in turn, that the machine switched modes at random, i.e. that the switches were not pre-programmed, nor—crucially—were they dependent upon anything the participant did; that they could try determine what mode the machine was in through reverse inference, i.e. if the machine had been mostly rewarding them, then it was more likely to be in Positive Mode, and vice versa for Negative Mode; and that since the machine didn't solely administer rewards in Positive Mode nor solely administer punishments in Negative Mode, that they should monitor outcomes over the last few trials rather than merely looking at the last trial.

Similarly to Experiment 1, we interspersed 4 multiple choice comprehension check questions throughout the training period covering 1) what they can expect in Positive Mode; 2) that the switches between modes occurred at random intervals; 3) that in order to determine which mode the machine was in, they should pay attention to what the machine does; and 4) when deciding what mode the machine is in, that they should pay attention to the last few outcomes, with each question presented immediately after the explanation of the corresponding point. Just like in Experiment 1, the questions were posed again if a participant answered incorrectly up to a maximum of 3 times after the correct answer was displayed onscreen.

The full text of the instructions is in Appendix A.

Casino game task. Participants played 150 trials of the casino game. On each trial, participants saw a "one-armed bandit machine" on the screen in front of them, labeled with the two potential rewards it could confer: 120 and -120 tokens (see Figure 5). To begin the trial participants would press the down arrow key to pull the lever. Once they did so, the lever on the machine would depress and the cogwheel graphic would animate for 600 ms. The outcome of the trial (120 or -120 tokens) would display on screen for 400 ms before a text prompt would appear on screen requiring the participant to use the left and right arrow keys to decide whether they thought the machine was in Positive Mode or Negative Mode. The outcome would remain on screen until the participant made their choice.

The two modes operated in much the same way as Experiment 1. In Positive Mode, the participant would receive a positive outcome with a probability of 0.67, and a negative outcome otherwise. In the Negative Mode, these contingencies were reversed. The task alternated between modes with a probability of 0.167 on every trial.

As well as the tokens returned by the bandit machine, the participant earned 30 tokens for every correct judgment they made regarding the current mode of the game. Their total bonus payment was a combination of the reward received from the machine and the amount earned from their judgments. Like Experiment 1, participants started with a tally of 10000 tokens so as to avoid their ever reaching a negative tally; every 100 tokens earned in the game was worth 1c; and participants were provided with a running tally on the screen throughout the task, though in this case only the portion of the tally that was due to the machine (and not their judgments) was displayed.

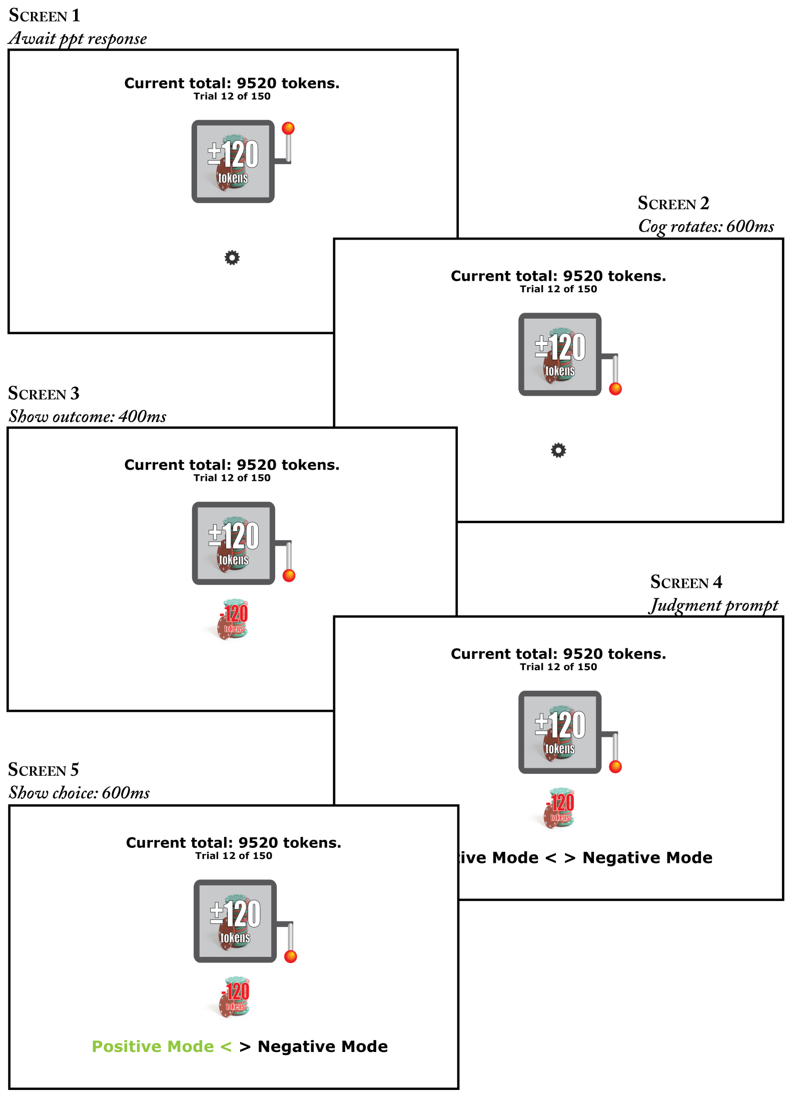


Figure 5. Screenshots and timing for one trial in Experiment 2.

Exit survey. The exit survey items were the same as in Experiment 1.

Dependent measure. Given the potential limitations of our mirrored proportions measure of bias discussed in Experiment 1, we decided to focus entirely on the Two-Back Outcome Model measure for Experiment 2. The logic of the measure was unaffected by the changes to the paradigm. The only difference is that the dependent variable now represents *mode choice*, rather than lever choice.

Results

7 participants suffered technical problems and a further 12 failed to complete the task for unknown reasons, leaving 330 participants for analysis (M_{age} = 33.1, 37% female). No participants were otherwise excluded.

Two-Back Outcome Model Measure. To measure bias we used the Two-Back Outcome Model as per Experiment 1 (substituting mode judgment for lever choice). The model failed to converge with the full random effects structure, so we were forced to drop the random slopes estimates. The results showed a significantly positive intercept term (β = 0.27, z = 8.06, p < 0.001) indicating that participants were biased towards guessing Positive Mode. Full results are shown in Table 6.

Reaction Time. We modeled reaction time for mode judgment (choice) as per Experiment 1. The results showed an effect of choice on reaction time (β = -0.07, t = -7.69, p < 0.001) as well as significant interactions with previous outcomes (see Table 7 and Figure 7 for full details). Considering that a participant's belief that the game was in positive mode might have engendered greater enthusiasm to pull the lever on the next round, we were also interested to run the same model for lever pull reaction times. The results showed no effect of previous choice on lever pull reaction time (β = -0.01, t = -1.66, p = 0.099).

Bias across trials. We assessed the progression of the bias across trials using the same model as in Experiment 1 (again, substituting mode judgment for lever choice). The results showed no significant change in bias across trials (see Table 8).

Demographics and task engagement measures. Finally, we tested for demographic effects using the same model as in Experiment 1. Again, the effect of gender was significant ($\beta = 0.07$, $z = 2.28$, $p = 0.022$) such that males were more likely to choose Positive Mode. The full results are listed in Appendix C. We also tested for the effect of task engagement using the same measures as Experiment 1. This time we found a significant negative effect of being counted as a strategizer ($\beta = -0.07$, $z = -2.26$, $p = 0.023$), implying that strategizers managed to suppress their bias to some degree.

Table 6

Experiment 2: Effect of previous trial outcomes on likelihood to choose Positive Mode

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.27	0.03	8.06	< 0.001
Outcome t-1	1.17	0.01	78.24	< 0.001
Outcome t-2	1.04	0.01	70.02	< 0.001
Ot-1 x Ot-2	-0.0001	0.01	-0.01	0.993

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode.

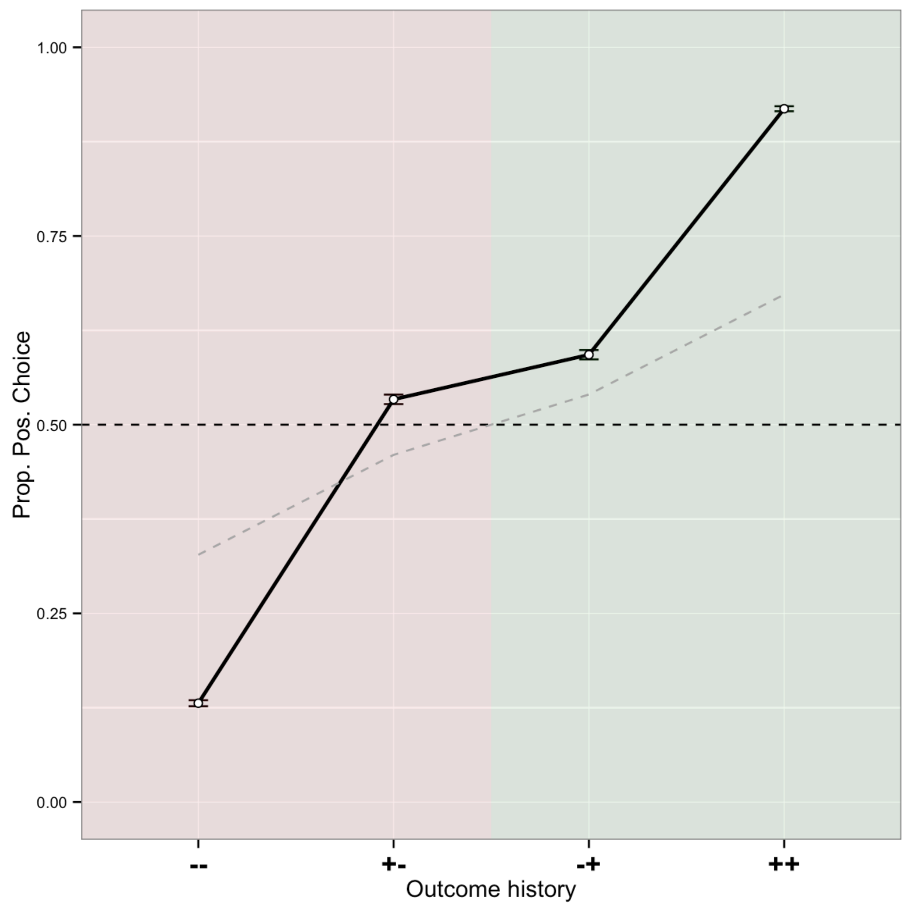


Figure 6. Choice behavior for Experiment 2. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line.

Table 7

Experiment 2: Effect of previous trial outcomes and choice on reaction time.

Fixed Effects	*Estimate*	*Std. Error*	*t value*	*p*
Intercept	6.14	0.036	168.17	< 0.001
Choice	-0.07	0.009	-7.69	< 0.001
Outcome t-1	0.03	0.006	5.09	< 0.001
Outcome t-2	0.02	0.006	3.48	< 0.001
Choice x Ot-1	-0.14	0.006	-23.9	< 0.001
Choice x Ot-2	-0.06	0.006	-9.9	< 0.001
Ot-1 x Ot-2	-0.09	0.006	-15.89	< 0.001
Ch. x Ot-1 x Ot-2	-0.02	0.006	-3.44	< 0.001

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as -1 for Negative Mode and 1 for Positive Mode. Reaction times were log transformed.

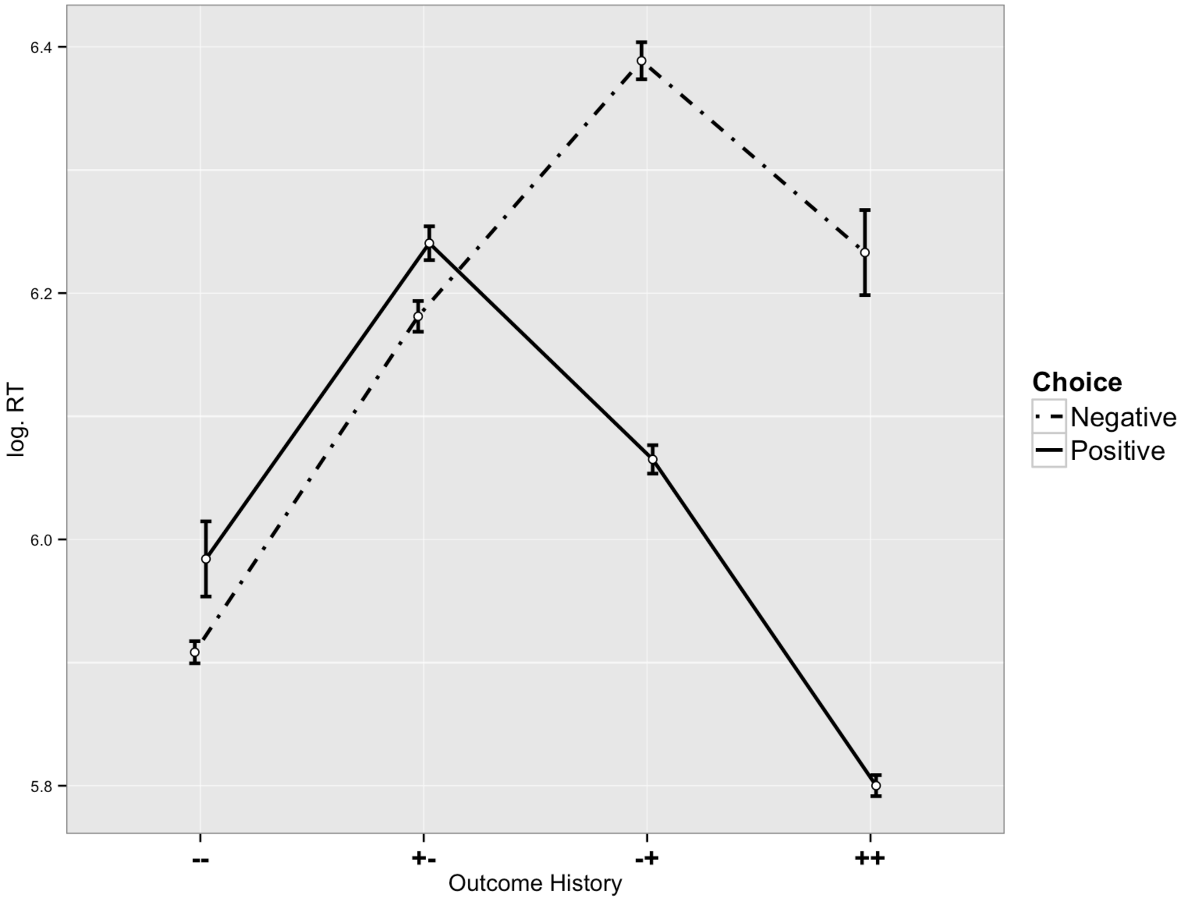


Figure 7. Reaction times for Experiment 2. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Table 8

Experiment 2: Effect of previous trial outcome and trial on likelihood to choose Positive Mode.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.2	0.04	4.8	< 0.001
Outcome t-1	0.95	0.02	50.47	< 0.001
Trial	0.04	0.03	1.44	0.15
Trial Quadratic	0.04	0.03	1.51	0.13
Ot-1 x Trial	-0.04	0.01	-3.13	0.002
Ot-1 x Trial Q.	0.06	0.01	4.15	< 0.001

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Trial regressors centered at session midpoint. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

Discussion

Experiment 2 more clearly demonstrates wishful thinking than Experiment 1. Not only is the bias estimate larger than in Experiment 1, it also cannot be attributed to risk preference or reinforcement history.

Why have we seen a stronger bias in Experiment 2? A potential explanation is that due to changes in the reward structure, participants might have felt a stronger inducement to indulge in wishful thinking. But this explanation can be ruled out: not only was the difference in expected reward between states—independent of participant behavior—slightly smaller than in Experiment 1[6], but the reward for accuracy was greater[7].

A more likely explanation is that, freed from the countervailing influence of risk aversion, the true extent of wishful thinking has now emerged. However, it is worth noting that even when we controlled for risk preference in Experiment 1, the estimate of bias barely changed ($\beta_{NO\_RP\_CONTROL} = 0.17$, $\beta_{RP\_CONTROL} = 0.19$), and it certainly didn't match the magnitude we see in Experiment 2. While risk aversion may have tempered the bias to some degree, it doesn't seem sufficient to explain the full extent of the difference.

An alternative explanation is that in Experiment 1, participants were shown the outcome of their choice for 1000 ms before being able to make their next choice, whereas in Experiment 2, the outcome was only on screen for 400 ms before participants were able to make their judgment. It is possible that the longer delay in Experiment 1 tempered the wishful thinking bias.

[6] 90 tokens in Experiment 1, 80 tokens in Experiment 2

[7] The amount of reward per trial that was contingent upon a participant's choice was the same in both experiments: 30 tokens. However, in Experiment 1, due to the stochasticity of the outcomes, the appropriate choice could still result in a suboptimal outcome. The average amount you could expect to earn by making the appropriate choice on every trial in Experiment 1 was 5 tokens.

This idea is consistent with the reaction time results we see in Experiment 2. These results require some unpacking. First up is the effect that the greater the evidence that a participant is in Positive or Negative Mode, the faster the corresponding judgment will be and the slower the opposite judgment will be. We can think of this as a confidence effect, and it is described by the significant interaction of outcome and choice in our mixed effects model of the reaction time data. This pattern breaks down at the extremes where a participant chooses the response that is relatively unlikely to be true, e.g. choosing positive after two negative outcomes. As can be seen in Figure 7, these judgments are quicker than the confidence effect might lead you to expect. We attribute this—much like in the reaction time data for the risk preference task—to a greater proportion of quick mindless responses to slow, mindful, indecisive responses in these bins.

But independent of the effect of confidence, there is a strong and significant effect of choice on reaction time. On the whole, participants are faster choosing Positive Mode than choosing Negative Mode.

Since Positive Mode is associated with greater reward, it might be suspected that this is due to some sort of reward facilitation effect. But this doesn't appear to be the case. It cannot be due to facilitation from a recent reward because the effect is controlled for outcome[8]. It cannot be due to an expectation of reward from the judgment since—though there is a reward for judgment accuracy—there is no reward associated with simply judging positive versus negative. Nor can it be due to facilitation from the expectation of future rewards from the machine since there is no effect of choice on the time taken to pull the lever.

[8] Interestingly, just like in Experiment 1, there is an effect of Outcome on judgment reaction time, such that participants are slower after a positive outcome. We still have no good idea why this might be.

Rather we think the effect of choice on reaction time provides a clue about the cognitive processes that underlie wishful thinking. Given that processing speed is a hallmark of automaticity (Evans, 2008; Shiffrin & Schneider, 1977), faster positive choices are consistent with the idea that wishful thinking is an automatic process.

Some other aspects of our data are worth noting. Again we saw no effect—linear or quadratic—of trial on the bias. The sign of the coefficient of the linear effect was positive, implying if anything that the bias got stronger as the task progressed, but this effect was not significant. This rules out the alternative explanation that participants start the experiment with a high prior expectation that the task will be generous to them, and that prior—rather than any biased updating throughout the task—is sufficient to explain the bias we find. Given the lack of aspects of the task about which one could have pre-established ideas[9], and the thoroughness of the instructions, that alternative explanation is undermotivated to begin with. If it were true however, it would predict that the bias would decrease over time as the hypothetically rational participants came to learn the unbiased nature of the task. As we can see, that is not the case in our data[10].

We saw the same effect of gender as we did in Experiment 1, where the bias towards judging Positive Mode was stronger in males than females. This was an entirely unexpected effect in Experiment 1, but is more difficult to write off now that we have seen it replicated, and now that risk preference has been entirely removed as a potential explanation. It would not be

[9] The best candidate we can muster in this regard is a possible prior belief that psychologists are both generous and misleading, so that, despite what they may say, the experimenters are likely to have fixed the task to be rewarding. We will leave it to the reader to decide which of these attributes is more likely to be true of a psychologist.

[10] Interestingly, the lack of a quadratic effect here might be taken to imply that the quadratic effect of trial that we saw in Experiment 1 may have been due to varying risk preference over the course of the task. That is, participants may have started out relatively risk seeking, became relatively risk averse as the task wore on, before becoming relatively risk seeking again as the end of the task neared.

difficult to start to speculate as to the possible reasons for this effect, but given that will be repeating this paradigm throughout the rest of the dissertation anyway, it is prudent to hold off on speculation until all the data is in.

Finally, we also saw an effect of being a strategizer on choice, such that strategizers were less likely to judge that they were in Positive Mode. This result isn't too surprising. If anything, it is perhaps more surprising that being a strategizer only attenuates the bias, and doesn't abolish it outright. Our main concern in deciding to categorize participants as strategizers was that they might effectively select themselves out of our test for wishful thinking. Though significant, the effect here doesn't seem dramatic enough to be worried about. Indeed, as it transpired, there was no significant effect of being a strategizer in all but one of the other experiments in this dissertation (see Appendix D).

Though our new design sidesteps the potential for risk preference or reinforcement history to explain our results, there remains one potential alternative explanation: biased memory encoding. If participants paid more attention to the task when times were bad than when times were good, they might be quicker to realize that things are getting better than that things are getting worse. This would mean that they would be more likely to judge Positive Mode simply because of a biased encoding of information, rather than a biased updating of information.

The hypothesis has an intuitive plausibility. It makes sense that we might devote more attention to things when times are bad—that's when they are most in need of our attention. Furthermore, negative stimuli are known to be attention grabbing (Armony & Dolan, 2002). On the other hand, the expectation of a negative stimulus might actually be attentionally aversive (Isaacowitz, 2005). In Experiment 3, we included memory probes in our paradigm in order to test this alternative hypothesis.

Experiment 3

Experiment 3 aimed to rule out the potential memory-encoding alternative explanation of the wishful thinking bias found in Experiment 2. That hypothesis predicts that outcomes witnessed when times are bad will be more likely to be encoded than when times are good. We did this by interspersing memory probes throughout the game, asking participants to recall the outcomes from previous trials and rewarding them for accuracy. Assuming objective likelihood to be in a particular state to be a reasonable proxy for subjective belief that the game is in that state, memory accuracy for our probes should be higher when the game was more likely to be in the punishing state than when it was more likely to be in the rewarding state if the memory encoding hypothesis is correct.

We probed participant's memory for outcomes from 2 and 4 trials ago. We did this partly to increase the variety and unpredictability in the task, but in so doing we opened up an interesting opportunity to explore the interaction of wishful thinking and memory. There are two reasons to think there might be a bias towards memories of positive outcomes, each of which predicts a different effect of such a bias on more versus less recent trials.

On the one hand, insofar as your memory of a particular outcome is uncertain, you may be expected to bias your recall towards your estimate of the task state at that time. Since participants' state estimates are biased towards Positive Mode, this may bias their memory towards positive outcomes. In this case we would expect a stronger memory bias for more distant trials relative to more recent trials.

On the other hand, insofar as what has happened recently has implications for what is likely to happen now, we might expect more of a bias for recent trials relative to more distant trials, since the diagnosticity of an outcome for current task state will diminish with every trial that passes.

In Experiment 3, as well as asking whether memory accuracy for outcomes is better depending upon the likely task state at the time of encoding, we will assess memory for bias towards positive outcomes, and whether any such bias is stronger for more or less recent trials.

Method

Participants. We recruited 230 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). All other recruitment details were the same as Experiment 2.

Procedure. The procedure was the same as Experiment 2 except for the differences noted below.

Training. As well as being instructed as per Experiment 2, participants also received training on the memory task component of the game. We informed them that every so often they would be prompted to remember what happened some amount of trials ago and that they would receive 100 extra tokens for every correct answer. We explained one hypothetical example which emphasized that they were to include the most recent outcome when counting back to the target trial. Participants then took a training session of seven trials during which they were prompted to remember the outcome from both 2 trials and 4 trials ago. Finally we posed a comprehension question asking them to imagine that they had just received a certain sequence of three outcomes and to tell us what the correct answer would be if we asked them for the outcome two trials ago. Though it sounds like a trivial question, it was important to ensure participants understood that "two trials ago" referred to two outcomes ago, and not three outcomes ago which would be a reasonable interpretation if a participant counted the most recent outcome as being the current trial, and as such didn't include it when counting back.

Casino game task. We occasionally probed participants' memory during the task. There were 30 memory probes in total: 15 asking the participant to identify the outcome from 2 trials beforehand (i.e. two outcomes ago), and 15 asking for the participant to identify the outcome from 4 trials beforehand. Each probe was initiated by the message "[2/4] trials ago?" appearing in place of the cogwheel at the center of the screen (see Figure 8). It was followed 600 ms later by the text "120 tokens < > -120 tokens" which was the prompt for participants to indicate their response. The response choice was highlighted on screen for 400 ms before the game continued.

Since it would have been—however weakly—diagnostic of the current game state, no feedback was given in response to participants' answers on the probes. There were no probes in the first 5 trials but otherwise memory probes were randomly interspersed throughout the task. As such, it was possible that two separate probes could identify the same target trial (if a 4-trial probe occurred two trials after a 2-trial probe). It was also possible that a probe could identify an earlier target than the probe preceding it (if a 4-trial probe occurred immediately after a 2-trial probe).

In response to participant feedback, we quickened each trial in order to make the task a bit snappier and more enjoyable. We did this by shortening the duration that the cog rotated after the participant pulled the lever from 600 ms in Experiment 2 to 400 ms in the current study. We also shortened the time that the participant's choice as to whether the game was in positive or negative mode was highlighted on screen before the next round commenced from 600 ms to 400 ms. Neither of these changes were expected to make any difference to participant's behavior during the task.

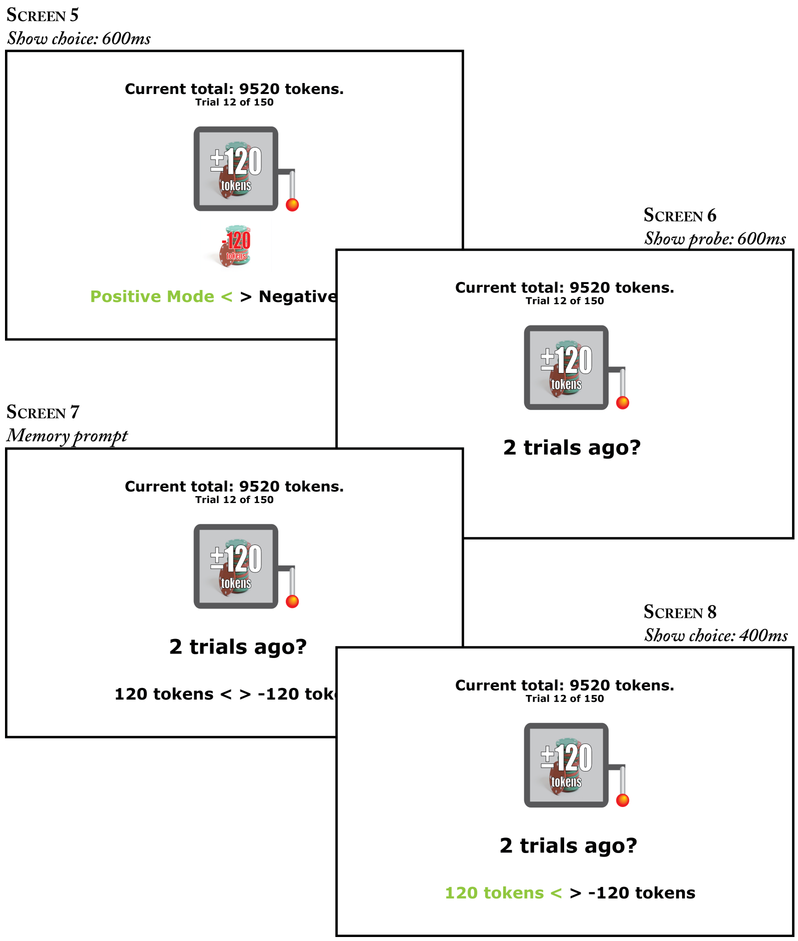


Figure 8. Screenshots and timing for the memory task component of Experiment 3.

Exit survey. We added a measure to our exit survey to gauge whether participants were using external workarounds, such as writing the outcomes down, to help with the memory task. See Appendix C for wording.

Results

3 participants suffered technical problems leaving 227 participants for analysis (M_{age} = 32.7, 48% female). 4 participants reported writing the outcomes down during the task but were not excluded from analysis. No participants were otherwise excluded.

Two-Back Outcome Model Measure. To measure bias we used the Two-Back Outcome Model as per Experiments 1 and 2. As in Experiment 2, the model failed to converge with the full random effects structure, forcing us to drop the random slopes estimates. The results showed a significantly positive intercept term ($\beta = 0.28$, $z = 5.61$, $p < 0.001$), replicating the bias from previous studies. Unlike previous studies, there was a significant outcome interaction ($\beta = 0.03$, $z = 2.07$, $p = 0.039$), with the positive parameter estimate implying that participants were more likely to choose positive mode after two congruent previous outcomes (regardless of valence) than two incongruent previous outcomes, all else being equal. This means that participants were *less* likely to exhibit the bias on more ambiguous trials. Full results are shown in Table 9.

Memory Probes. Performance on the memory component of the task is plotted in Figure 10. We tested for the asymmetric memory encoding hypothesis by using previous outcomes at the time of the target outcome as proxies for perceived likelihood to be in positive mode at the time. We constructed a mixed effects model designed to test the relationship between the memory performance and perceived current mode. Our fixed effects included the full interaction of the two outcomes previous to the target outcome and the distance from the target outcome at the time of the probe (2 trials vs. 4 trials). For random effects we included random intercepts for participants. Target distance was coded with more distant trials (4 trials back) as the reference level. The model intercept was significant ($\beta = 0.76$, $z = 13.96$, $p < 0.001$), implying that participants showed better than chance accuracy even on the more distant trials. There was no effect of previous outcomes, but there was a significant effect of target distance ($\beta = 0.75$, $z = 12.84$, $p < 0.001$, for full details see Table 10).

We also tested to see if there was any bias in participants' memory, and whether this differed depending on the memory target distance. We did this by calculating bias using the non-parametric signal detection based measure in Mueller and Zhang (2006), encoding positive outcome as signal present. This measure returns a value, *b*, the natural logarithm of which should not significantly differ from zero if there is no memory bias.

Two participants had to be dropped from the analysis for returning uninterpretable values. We found no evidence of memory bias for either the 2-trial probes ($M = 0.02$, $t(224) = 0.68$, $p = 0.5$) or the 4-trial probes ($M = 0.05$, $t(224) = 1.56$, $p = 0.12$). Nor was there any difference in bias between both trial distances ($M_{T2-T4} = -0.03$, $t(224) = -0.69$, $p = 0.49$).

Reaction Time. Reaction time for mode choice was modeled as per Experiment 2. We saw a very similar pattern of results to Experiment 2 including a replication of the main effect of choice ($\beta = -0.04$, $t = -4.58$, $p < 0.001$, see Table 11 and Figure 11 for full results).

Bias across trials. We were still concerned to show that the bias did not disappear by the end of the task, and attempted to model its progression using the same mixed effects model as Experiment 2. That model didn't converge so we tried two alternative models. The first dropped the outcome regressor from the Experiment 2 model, leaving both trial and the quadratic trial regressor as the two fixed effects. The random effects structure incorporated both random intercepts for participant as well as by-participant random slope estimates for both trial and the quadratic trial factors. This model returned a significantly positive effect of trial ($\beta = 0.05$, $z = 2.25$, $p = 0.024$). The second model retained the outcome regressor from the Experiment 2 model but dropped the random slopes form the random effects structure. This model also returned a significantly positive effect of trial ($\beta = 0.06$, $z = 4.59$, $p < 0.001$; full details of both models in Tables 12 and 13). Together, both these results imply that the bias increased over the course of the game.

Demographics and task engagement measures. Demographics were tested as in both Experiments 1 and 2. No significant main effects of any of our demographic measures were found (see Appendix C). We likewise tested for task engagement effects using the same analysis as Experiments 1 and 2. No significant effects were found for any of our measures.

Table 9

Experiment 3: Effect of previous trial outcomes on likelihood to choose Positive Mode.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.28	0.05	5.61	< 0.001
Outcome t-1	1.37	0.02	84.68	< 0.001
Outcome t-2	0.98	0.02	61.48	< 0.001
Ot-1 x Ot-2	0.03	0.02	2.07	0.039

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode.

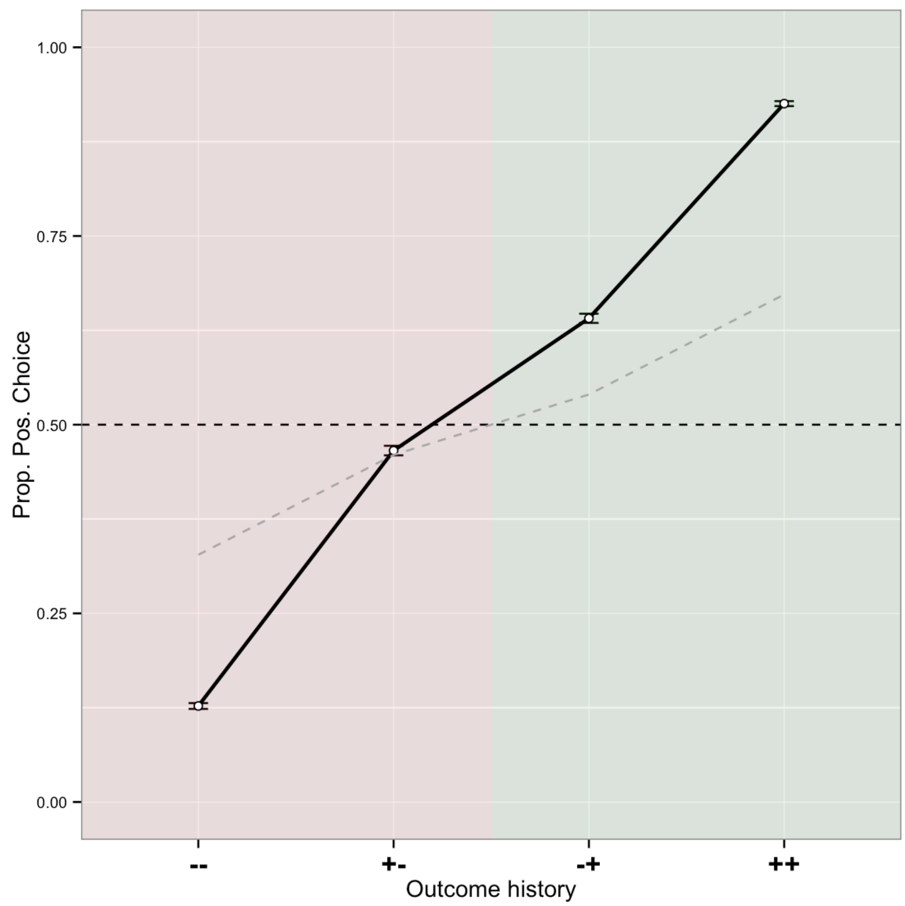


Figure 9. Choice behavior for Experiment 3. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line.

Table 10

Experiment 3: Effect of previous trial outcomes and target distance on memory probe accuracy.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (accuracy)	0.76	0.05	13.96	< 0.001
Outcome t-1	-0.04	0.04	-1.11	0.265
Outcome t-2	0.04	0.04	0.94	0.349
2 trials back	0.75	0.06	12.84	< 0.001
Ot-1 x Ot-2	0.01	0.04	0.24	0.813
Ot-1 x 2 trials	0.08	0.06	1.37	0.172
Ot-2 x 2 trials	0.02	0.06	0.31	0.756
Ot-1 x Ot-2 x 2 tr.	0.12	0.06	2.13	0.033

Note. Outcome t-1 represents most recent outcome before the target. Outcome regressors coded as 1 for reward and -1 for punishment. Accuracy coded as 0 for incorrect and 1 for correct.

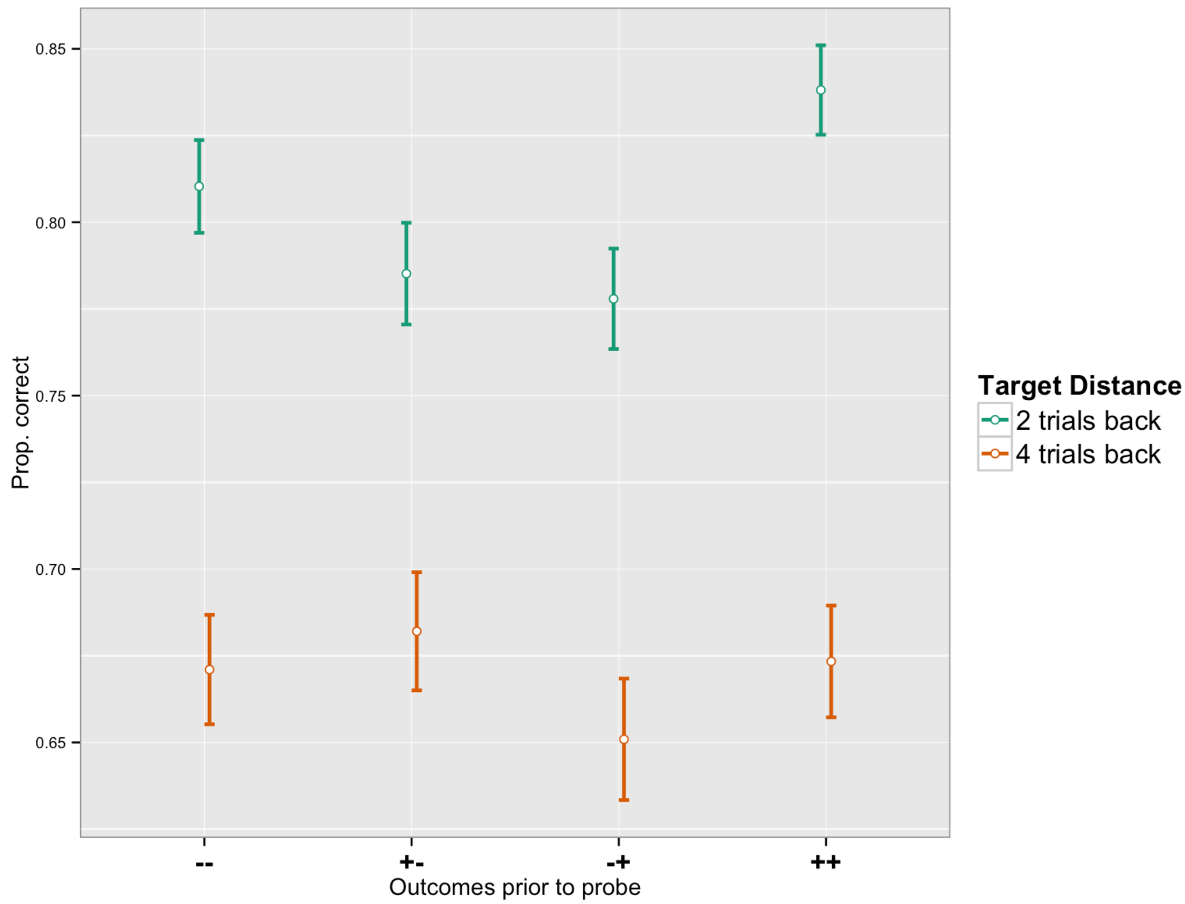


Figure 10. Performance on the memory probe task plotted by trial outcomes prior to memory target. Outcome history is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. Error bars calculated using the adjustment in Morey (2008).

Table 11

Experiment 3: Effect of previous trial outcomes and choice on reaction time.

Fixed Effects	*Estimate*	*Std. Error*	*t value*	*p*
Intercept	6.5	0.041	158.24	< 0.001
Choice	-0.04	0.009	-4.58	< 0.001
Outcome t-1	0.01	0.006	1.66	0.096
Outcome t-2	0.008	0.006	1.34	0.181
Choice x Ot-1	-0.15	0.006	-24.86	< 0.001
Choice x Ot-2	-0.04	0.006	-6.54	< 0.001
Ot-1 x Ot-2	-0.11	0.006	-19.73	< 0.001
Ch. x Ot-1 x Ot-2	-0.007	0.006	-1.22	0.222

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as -1 for Negative Mode and 1 for Positive Mode. Reaction times were log transformed.

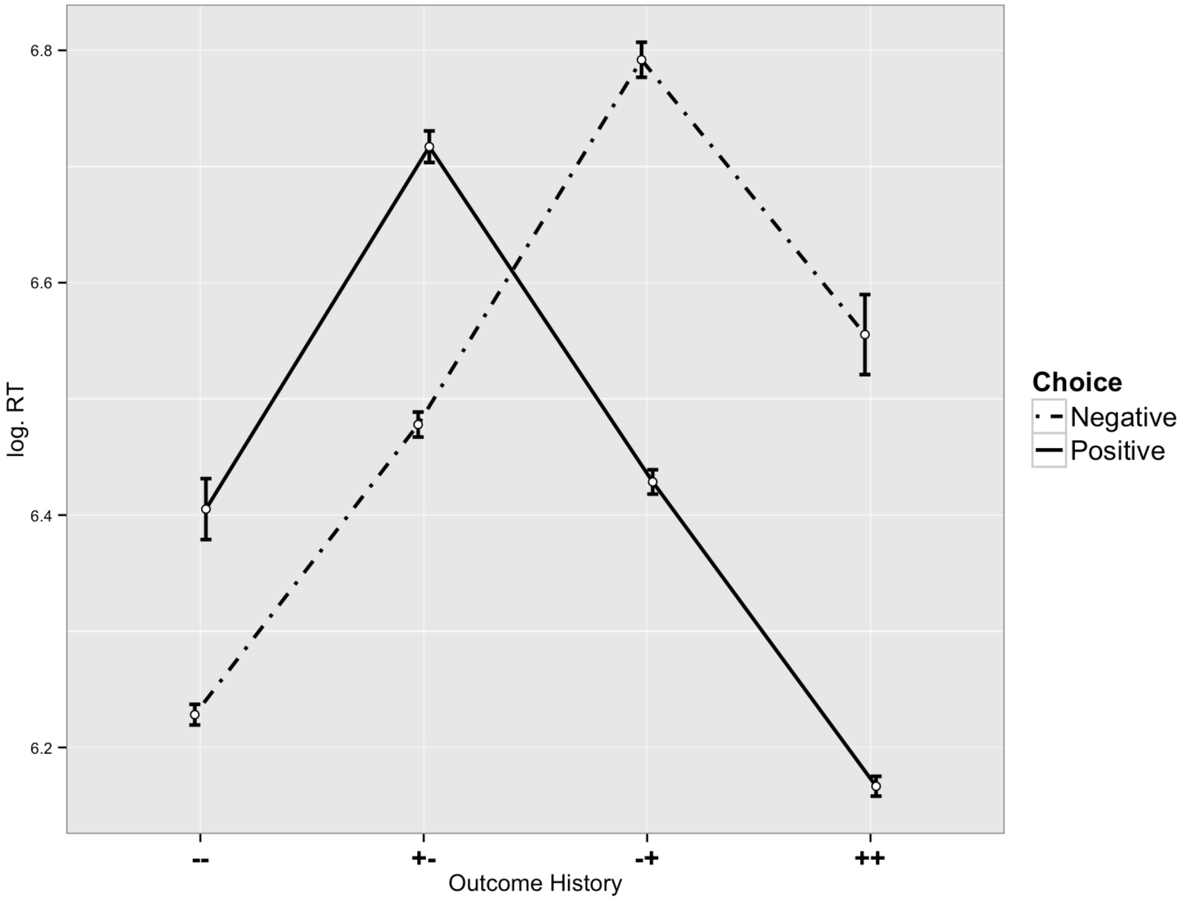


Figure 11. Reaction times for Experiment 3. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Table 12

Experiment 3: Effect of trial on likelihood to choose Positive Mode.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.2	0.05	4.15	< 0.001
Trial	0.05	0.02	2.25	0.024
Trial Quadratic	-0.04	0.03	-1.37	0.169

Note. Trial regressors centered at session midpoint. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

Table 13

Experiment 3: Effect of previous trial outcome and trial on likelihood to choose Positive Mode.

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.25	0.05	5.37	< 0.001
Outcome t-1	1.15	0.02	57.57	< 0.001
Trial	0.06	0.01	4.59	< 0.001
Trial Quadratic	-0.02	0.01	-1.6	0.109
Ot-1 x Trial	0.01	0.01	1	0.318
Ot-1 x Trial Q.	0.04	0.01	2.53	0.011

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Trial regressors centered at session midpoint. Choice coded as 0 for Low-Range Lever and 1 for High-Range Lever.

Discussion

Experiment 3 both replicated the wishful thinking bias of Experiment 2 and ruled out asymmetric memory encoding as a potential alternative explanation. There was no effect of recent outcomes at the time of the probe upon memory accuracy. Excitingly, we also replicated the effect of choice on reaction time, further strengthening the idea that the bias may be the result of an automatic process. However, there were some mild differences between the results of Experiment 2 and Experiment 3 that warrant scrutiny.

As is apparent even with a visual comparison of the choice results from Experiment 2 and Experiment 3, participants seemed to place more weight on the most recent outcome relative to the two-back outcome in Experiment 3 than in Experiment 2. This can be seen by the greater separation between the responses in the ambiguous bins in this experiment. We attribute this to the load placed upon the participants as a result of the memory task which may have reduced working memory capacity and disrupted participants' ability to hold less recent outcomes in mind. Interestingly, this explanation implies that participants make their judgment by considering a range of recent outcomes held in working memory, as opposed to, for instance, a running estimate that gets updated on every trial. As we shall see, this effect anticipates Experiment 5 where we manipulate cognitive load.

We also saw a significant effect of ambiguity. That is, trials that are characterized by a mixed sequence of prior outcomes showed a significantly different gravitation towards positive judgments than trials characterized by a repeated sequence of prior outcomes. Surprisingly, this effect tells us that the ambiguous trials showed *less* bias towards positive judgments than the unambiguous trials. However, since we didn't see it in Experiment 2, and since the effect is both mild and in the unexpected direction, we are inclined to dismiss it.

Another difference is the significant effect we now see in the extent of choice bias over the course of the task. Our goal in looking at the trend in bias was to rule out the possibility that participants were beginning the task with optimistic expectations and that this alone—rather than any bias in updating—would explain the bias that we found (see Experiment 2 Discussion). We were mostly concerned to see that the bias didn't diminish over time. Given the results of Experiment 2—where we saw a non-significant trend towards greater bias over time—and now, Experiment 3—where we see a significant trend towards a greater bias over time—it seems safe to rule that possibility out. This is welcome given that eliminating any opportunity for participants' prior beliefs to influence behavior was a chief aim in designing the task.

But this also forces us to wonder about this increasing effect of trial on bias. Is it a genuine effect? And if so, how do we explain it? Since the effect of trial on bias will not be a crucial question for the experiments in Chapter 3, it is perhaps best to skip ahead now and review this effect across all of our studies rather than reporting this result piecemeal, experiment by experiment. The collated results for this analysis for all experiments can be seen in Appendix F. We had the same difficulties with model convergence for all experiments in Chapter 3 as we did in this experiment so we performed the same analysis strategy: estimating the effect of trial on choice controlling for outcome but not including the trial slope estimate as part of our random effects structure—a potentially anti-conservative strategy (Barr et al., 2013)—and modeling the effect of choice on trial including the random slope estimates but no longer controlling for the effect of outcome—a potentially less sensitive strategy.

As can be seen in Appendix F when we use the first analysis strategy we see a frequent, though not entirely consistent positive effect of trial on choice, in line with the result we see in Experiment 3. These results are strongly attenuated when we employ the second analysis strategy, so that effects are only trending, if they remain present at all. This might seem to undermine the results from the first anti-conservative strategy but it could also be that the variance that we fail to absorb from our outcome condition in the second strategy is swamping the effect.

In any case, the results are all consistently pointing in the same direction. While it is far from a certain effect, it is at least a prevalent enough trend to suspect it might be worth further targeted investigation. Should the effect hold, an intriguing, and speculative explanation, is that the bias itself is reinforcing. Wishful thinking is rewarding, and the more you indulge in it, the more likely you are to indulge in it again. Though speculative, this idea might align with recent models of executive control that consider cognitive operations, and not just actions, to be potential targets for reinforcement (e.g. Hazy, Frank, & O'Reilly, 2007; see also Mijovic-Prelec & Prelec, 2009). A different, though not necessarily incompatible explanation is that participants grow weary or less attentive to the task over time, and as they do, they are more likely to rely upon automatic processing, which promotes wishful thinking. This is perhaps a simpler explanation and would be consistent with the evidence from the reaction time data, but confirmation will have to await future investigation.

Finally we turn to our memory task. Participants did commendably well on the task, performing significantly above chance even when the target trial was four trials ago. However, besides the unsurprising difference in accuracy due to target distance, there seems little to be gleaned from our memory data. As can be seen in Table 10, there is a significant interaction of both outcome regressors and target distance. This is visually evident as the 'U' shape pattern in accuracy levels across outcome for two trials ago that is not present at four trials ago, showing a slight increase in accuracy for the unambiguous trials (those at either extreme of the x-axis).

Unexcitingly, we suspect that this effect is merely due to the fact that you can afford a greater inaccuracy in recalling a particular trial outcome without hurting memory accuracy when the neighboring trial outcomes are homogenous—as is the case for the unambiguous trials.

Perhaps disappointingly, we saw no bias towards either positive or negative outcomes in participants' responses to the memory probes. This result is similar to that found in Sharot et al. (2011) who also found no difference in memory accuracy for desirable or undesirable information in their belief updating task. However, our task was not optimized to find such a result, and even though there were no significant differences, the pattern of results was suggestive of a more positive bias for more distant trials. Perhaps with more probes and more distant targets, a memory bias could be brought out. We retain our sense that the relationship between wishful thinking in the present and nostalgia for the past may prove an interesting avenue of inquiry.

Concluding Remarks

In Chapter 2, we developed a paradigm that elicited wishful thinking in a simple, repeated decision-making task. Unlike other paradigms aiming to assess motivated bias, our task did not ask participants to indulge in an elaborate role play nor to consider the likelihood of personally concerning but causally opaque events. Instead we placed the participants in a world where they were aware of all the moving parts, and where the nothing in their experience prior to taking our task could inform them about how it would pan out. Only the evidence they encountered during the task could help them. In this sense, we saw wishful thinking laid bare.

Though we settled on a paradigm that asks participants to explicitly judge the favorability of the state they are in, our first experiment shows that the bias is not just an explicit process, but gets implicitly integrated with other decision variables, i.e. risk preference, in value-based decisions. Indeed, along the way we saw suggestions that biased updating of beliefs may be an automatic process. This is most clear in reaction time results that show a quicker response for positive judgments than negative judgments. But there were some other suggestions too: we saw a stronger effect of bias in Experiments 2 and 3, where participants were less delayed than in Experiment 1 before getting to make their decisions, and we saw a hint that the bias increases over the course of the task which is potentially attributable to a greater reliance on automatic processing as fatigue or mindlessness creeps in.

Our hope is that our paradigm will prove a useful tool in probing the processes behind wishful thinking and biased updating. These hints suggest a potentially fruitful first application. In Chapter 3 we aim to investigate whether wishful thinking is the product of automatic or controlled cognitive processing.

Chapter 3

Wishful Thinking: Dual-process manipulations

The studies in Chapter 2 demonstrated wishful thinking in a simple repeated decision-making context where both priors and evidence were well controlled. In an attempt to decompose the processes underlying wishful thinking, the next suite of experiments aims to embed it in the dual-process framework of judgment and decision-making (Evans, 2008; Evans & Stanovich, 2013; Stanovich & Toplak, 2012).

The dual-process framework has its roots in the many different areas of cognitive science that all invoked a roughly similar distinction between two different classes of cognitive processes. The distinctions made in areas such as reasoning and belief formation (Evans, 1984; Gilbert, 1991; Kahneman, 2003; Sloman, 1996; Stanovich & West, 1998), perceptual attention (Shiffrin & Schneider, 1977), reward learning (Hall & Pearce, 1979), philosophy (Fodor, 1983), and social cognition (Chaiken & Trope, 1999; Gilbert & Hixon, 1991; Lieberman, 2007), all bear a family resemblance to one another, but not an exact resemblance, and one aim of the dual-process framework has been to make sense of the similarities and differences. The result is a division of cognitive processes into those that operate automatically, rapidly, beneath conscious awareness, and which impose little if any demand upon working memory, and those that are conscious and controlled, temporally extended, and demanding of working memory resources. These two classes are often referred to as System 1 and System 2 respectively, but such terminology has been misunderstood as positing exactly two distinct cognitive or neural systems, as opposed to two different classifications of processes that may be common across many different cognitive systems. In response, we follow Evans and Stanovich (2013) in relabeling these classes as Type 1 and Type 2 processes.

The second goal of the dual-process framework is to describe how these two classes of processes interact to determine the judgments and decisions that we make. While some accounts consider both processes to operate in parallel (Sloman, 1996), more recent accounts describe a *default-interventionist* view (Evans & Stanovich, 2013; Kahneman, 2003; Morewedge & Kahneman, 2010) where Type 1 processes are considered to automatically and rapidly generate intuitive judgments which may be intervened upon by Type 2 processes in order to review and potentially revise those judgments. Such an intervention is cognitively effortful and imposed more or less sparingly (Stanovich & West, 1998), with the result that systematically inaccurate intuitions can often slip by.

Given that characterization, our reaction time results from Chapter 2—where judgments that participants were in positive mode were found to be quicker than judgments that they were in negative mode—seem to suggest that wishful thinking may be the result of Type 1 processing. Further support can be gathered from scattered studies on related forms of bias. Beer and Hughes (2010) found that putting participants under time pressure increased self-enhancement when rating themselves against others on personality traits. Sanitioso et al. (1990) found that participants were faster to generate autobiographical memories that exemplified either introversion or extraversion depending upon which of the two traits had been rendered as more desirable by the researchers. Lieberman, Ochsner, Gilbert, and Schacter (2001) found that putting participants under cognitive load failed to disrupt cognitive dissonance based updating, which on some accounts should be considered a motivated process (Aronson, 1997; Sherman & Cohen, 2006). Epley and Whitchurch (2008) showed that participants were quicker to recognize their own face when the image of their face was partially morphed with an attractive face than when it wasn't morphed at all, though it is not clear that this isn't simply due to an already biased representation of one's face, rather than a bias in processing during recognition (see Pinker, 2011; von Hippel & Trivers, 2011).

Under the view that non-human animals and non-adult humans are less governed by Type 2 processes, some developmental and comparative cognition results might be considered to lend some support. Biased updating about the likelihood of undesirable future life events is stronger in adolescents (Moutsiana et al., 2013), and dissonance based updating has been shown in monkeys and children (Aronson & Carlsmith, 1963; Egan, Santos, & Bloom, 2007; see also Peterson et al., 2003). Even rats show some evidence of interpreting ambiguity favorably (Harding, Paul, & Mendl, 2004).

In contrast to the above findings, Valdesolo and DeSteno (2008) found that putting participants under cognitive load eliminated their hypocrisy, i.e. their tendency to rate their own actions as less morally blameworthy than the same actions committed by others. Also, Kahan, Peters, Dawson, and Slovic (2013) found that numeracy increased participants' susceptibility to the confirmation bias when presented with somewhat complex numerical evidence, implying that their capacity to reason about the evidence facilitated the safe arrival at the preferred conclusion.

Perhaps the most difficult evidence to evaluate within this framework is the work on motivated reasoning first described in Chapter 1. Stimuli in that research often seem to require explicit (Dunning et al., 1995; Kunda, 1987), rule based (Dawson et al., 2002; Kahan et al., 2013), and temporally extended (Ditto & Lopez, 1992) evaluation. However, a continual theme of motivated reasoning research is that biased reasoning must always stay within the bounds of plausibility and maintain the façade of objectivity (Kunda, 1990). This suggests the bias is pitted against a process that is charged with reviewing judgments for legitimacy but which does not do so with perfect thoroughness; a characterization that fits with Type 2 processes in the default-interventionist model.

What's more, the bias inherent in motivated reasoning doesn't appear to be under conscious control. Indeed, if we look at the studies on self-deception described in Chapter 1 (Dana et al., 2007; Quattrone & Tversky, 1984) and the attendant controversy about its very possibility (Mele, 1997; Mijovic-Prelec & Prelec, 2009), it seems crucial that the bias avoids entering awareness or else risk failure. I cannot deceive myself if I am aware of so doing.

This idea is consistent with results from Batson's studies of moral motivation (Batson, 2007) where participants had to decide who—either themselves or another—would get to perform a potentially rewarding task, and who would be stuck with a seemingly dull task. Participants would avail of the opportunity to flip a coin in order to make the decision if one was provided, but would remain as selfish as they were without one even so. Much like in Dana et al. (2007), the decision to use the coin could only have been made in order to convince oneself that you were being fair—it didn't alter participants' allocations, there were no other witnesses, and participants who did avail of the coin reported their decisions as being more moral. Interestingly however, such self-deception (and the selfish allocation of tasks) vanished when participants made the decision when facing a mirror. Plausibly, heightened self-awareness engendered by the mirror made for more vigilant Type 2 processing, which unmasked the charade.

Given that the default-interventionist model arose out of evidence from tasks where accuracy was the yardstick and the presumed goal of the participant (Evans, 1984; Kahneman, 2003; Stanovich & West, 1998), Type 2 processes had little to do but correct the errors of Type 1 processes. However, motivated reasoning and self-deception open up the possibility that some of these processes could be used in service of biasing judgments towards a desirable conclusion, namely through a subconscious hijacking of working memory resources—an idea that admittedly, can be difficult to articulate precisely without falling into conceptual paradox (Kurzban & Aktipis, 2007; Mele, 1997).

From all of the above we can conclude first, that the evidence that can be currently brought to bear on the role of Type 1 and 2 processes in wishful thinking is scattered and at best only suggestive, and second, that the question is not only of interest insofar as it explains wishful thinking, but also serves as an interesting test case for the default-interventionist account of thought. In this chapter we aim to bring our wishful thinking paradigm to bear on this question.

In Experiments 4a and 4b we directly follow on from the reaction time findings of Chapter 2 and manipulate the time participants take to make their judgment in our task, with the idea that taking longer will increase the contribution of Type 2 processing. In Experiment 5, we will manipulate participants' cognitive load, with the idea that this will disrupt the contribution of Type 2 processing. Finally, in Experiments 6a and 6b we aim to manipulate thinking style—the weight participants put on the contribution from each type of processing—with an essay prime.

Experiment 4a

For our first experimental attempt to manipulate the operation of Type 1 and Type 2 processes, we chose to manipulate response time. Processing speed is one of the key characteristics that separates Type 1 and Type 2 processing (Evans, 2008; Stanovich & West, 2000), and manipulating the time in which participants make their decision—either by forcing a speeded response or by imposing a delay—has proven an effective way to unmask the differing contributions of each type of processing to various types of decisions (e.g. Beer & Hughes, 2010; Paxton, Ungar, & Greene, 2012; Rand, Greene, & Nowak, 2012).

We decided to impose a delay on participant's response times with the idea that this would increase the likelihood of Type 2 processes contributing to the decision. We expected this manipulation to increase the likelihood of each decision coming under review and rational adjustment and thereby to reduce bias (Morewedge & Kahneman, 2010; Stanovich & Toplak, 2012). Alternatively, and more in line with motivated reasoning research, this manipulation could provide a longer window for processes such as evidence reweighting, biased search, and decision justification which might strengthen the bias (Ditto & Lopez, 1992; Kunda, 1990).

We imposed the delay by manipulating the time between the appearance of the outcome and the response prompt (see Figure 5). We chose a within-participants manipulation design rather than a between-participants design primarily so as not to confound task length across conditions, and we chose to block the delayed and non-delayed trials into two separate sessions since preliminary testing on an interleaved design, where delayed and non-delayed trials were mixed together, proved less engaging and more frustrating.

The converse option would have been to force participants to respond quicker than they would otherwise do by constraining the time available to make their decision. We decided against this option because response times in both Experiment 2 and Experiment 3 were already quite quick (Mdn_{RT_Exp2} = 456 ms, Mdn_{RT_Exp3} = 580 ms), and we feared that attempting to speed them yet further would result in a large amount of missed trials and complicate our analysis. Furthermore, given our response time results from Experiments 2 and 3, our prior expectation was that there already is a high automatic processing component to the bias, and that we were more likely to affect response by imposing the engagement of Type 2 processes, rather than amplifying the contribution of Type 1 processes.

Nonetheless, we did allow participants to respond sooner than they could in Experiments 2 and 3. In both those experiments there was a delay of 400 ms from when the outcome appeared on screen until participants were able to respond. We shortened this duration to 100 ms on our non-delayed trials to give the participants to respond more quickly if they were so capable and willing. It was also thought that the decreased delay might encourage faster responding generally.

Method

Participants. We recruited 269 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). All other recruitment details were the same as Experiment 2.

Procedure. The procedure was the same as Experiment 2 except for the differences noted below.

Casino game task. Participants completed two sessions of the game, one of which prompted them for their judgment 100 ms after the outcome was displayed (the *no-delay* condition), the other of which prompted them for their judgment 700 ms after the outcome was displayed (*delay* condition). The timing value in the no-delay condition was shorter than in Experiments 2 and 3 (400 ms) in an attempt to encourage or entrain participants to quicker reaction times than they might otherwise have made. The choice of delay length in the delay condition was chosen after an informal inspection of reaction times in Experiment 2. We wanted to balance our aim of delaying for longer than responses would naturally take with our concern that if trials were too long, participant enjoyment and engagement with the task would suffer, and be confounded across conditions.

The order of the sessions was randomized across participants. Each session was 100 trials long and began with a fresh endowment of 10000 tokens on each session. Participants were paid the bonus from both sessions.

Like Experiment 3, the duration of the cogwheel rotation (300 ms) and the time for which the participant's choice was highlighted on the screen (400 ms) differed from Experiment 2. These subtle alterations were made to make the game run a bit smoother in response to participant feedback from Experiment 2.

Results

6 participants suffered technical problems and a further 30 failed to complete the task for unknown reasons, leaving 233 participants for analysis (M_{age} = 35.7, 54% female). No participants were otherwise excluded.

Two-Back Outcome Model Measure. To assess the effect of bias across conditions we adapted the Two-Back Outcome Model from Chapter 2. Now that condition was to be included as a factor, the model became too complex to converge with a maximal random effects structure. As such we settled on using the full interaction of the two previous outcomes and condition as fixed effects and by-participant random intercepts and slopes for condition as our random effects structure. Condition was coded with the no-delay condition as the reference level. The results showed a significantly positive intercept (β = 0.32, z = 5.87, p < 0.001) implying a wishful thinking bias in the no delay condition and replicating the main result from Chapter 2. There was no main effect of condition (β = -0.05, z = -1.06, p = 0.288) but condition did interact with both outcome variables (see Table 14 for full details).

Inspecting the results visually (see Figure 12) implied that the interaction with outcome may have been driven primarily by a difference in condition when the game was more likely to be in negative mode. To test this we ran the same model again but recoded the most recent outcome with 0 for a negative outcome and 1 for a positive outcome (whereas before it was -1 for a negative outcome and 1 for a positive outcome). Since the intercept estimate of our model represents mode choice when our outcome variables are zero, this meant that the intercept estimate would now test for likelihood to choose positive mode specifically on trials that were preceded by a negative outcome. Similarly our estimate of the main effect of condition would be measured for these trials. In short, this analysis tested whether condition affected choice when the game was more likely to be in negative mode. Results showed a significant effect of condition in this analysis ($\beta = -0.18$, $z = -3.23$, $p = 0.001$).

Reaction Time. Reaction time results are shown in Figure 13. To test for the effect of choice and condition we performed a similar analysis to those in Experiments 1-3 with the addition of condition as a fixed effect and including it as part of the random effects structure. To be fully specific, we ran a mixed effects model with the full interaction of the previous two outcomes, mode choice, and condition as fixed effects and with by-participant random intercepts and slopes for condition and choice and their interaction as the random effects structure.

Results showed an effect of choice in the no-delay condition ($\beta = -0.04$, $t = -6.05$, $p < 0.001$) such that participants were quicker to make positive mode choices, and a significant condition by choice interaction ($\beta = 0.02$, $t = 3.67$, $p < 0.001$) implying that this effect of choice was attenuated in the delay condition (see Appendix G for full model details).

Demographics. We tested for demographic effects on mode choice in a similar manner to Experiments 1-3. We included age, gender, education, income, and social and fiscal conservatism in a mixed effects logistic regression model. We also included previous outcome, condition and their interaction as fixed effects and interacted them with each of the demographic variables. Due to model complexity, we only included random intercepts for participants as random effects. The full results are shown in Appendix C. Notably, we saw no effect of gender ($\beta = 0.004$, $z = 0.11$, $p = 0.916$) but did see an effect of education ($\beta = -0.09$, $z = -2.28$, $p = 0.022$). Interestingly, the main effect of condition was significantly negative in this model ($\beta = -0.04$, $z = -2$, $p = 0.045$). This could have been due to controlling for our demographic variables, something which is given credence by the significant interaction of education and condition ($\beta = 0.1$, $z = 4.31$, $p < 0.001$) which implied that the effect of delayed responding was diminished the more educated a participant was. Alternatively this result could have been due to the removal of the condition slopes from the random effects structure compared to our Two-Back Outcome Model (Barr et al., 2013). To test for this we ran a simpler model that included only education out of our demographic variables and fully interacted it with condition and previous outcome. We also included a by-participants random slope estimate for condition, as well as by-participant random intercepts. The results showed no main effect of condition ($\beta = -0.06$, $z = -1.32$, $p = 0.187$) but both the effect of education ($\beta = -0.11$, $z = -2.17$, $p = 0.03$) and its interaction with condition retained their significance ($\beta = 0.11$, $z = 2.3$, $p = 0.02$).

Table 14

Experiment 4a: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode.

Fixed Effects	*Estimate*	*Estimate with Outcome t-1 recoded*
Intercept (bias)	0.32***	-0.89***
Outcome t-1	1.21***	2.42***
Outcome t-2	0.95***	0.95***
Delay	-0.05	-0.18**
Ot-1 x Ot-2	0.001	0.002
Ot-1 x Delay	0.13***	0.25***
Ot-2 x Delay	0.14***	0.14***
Ot-1 x Ot-2 x Delay	-0.002	-0.01

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. * $p < .05$, ** $p < .01$, *** $p < .001$

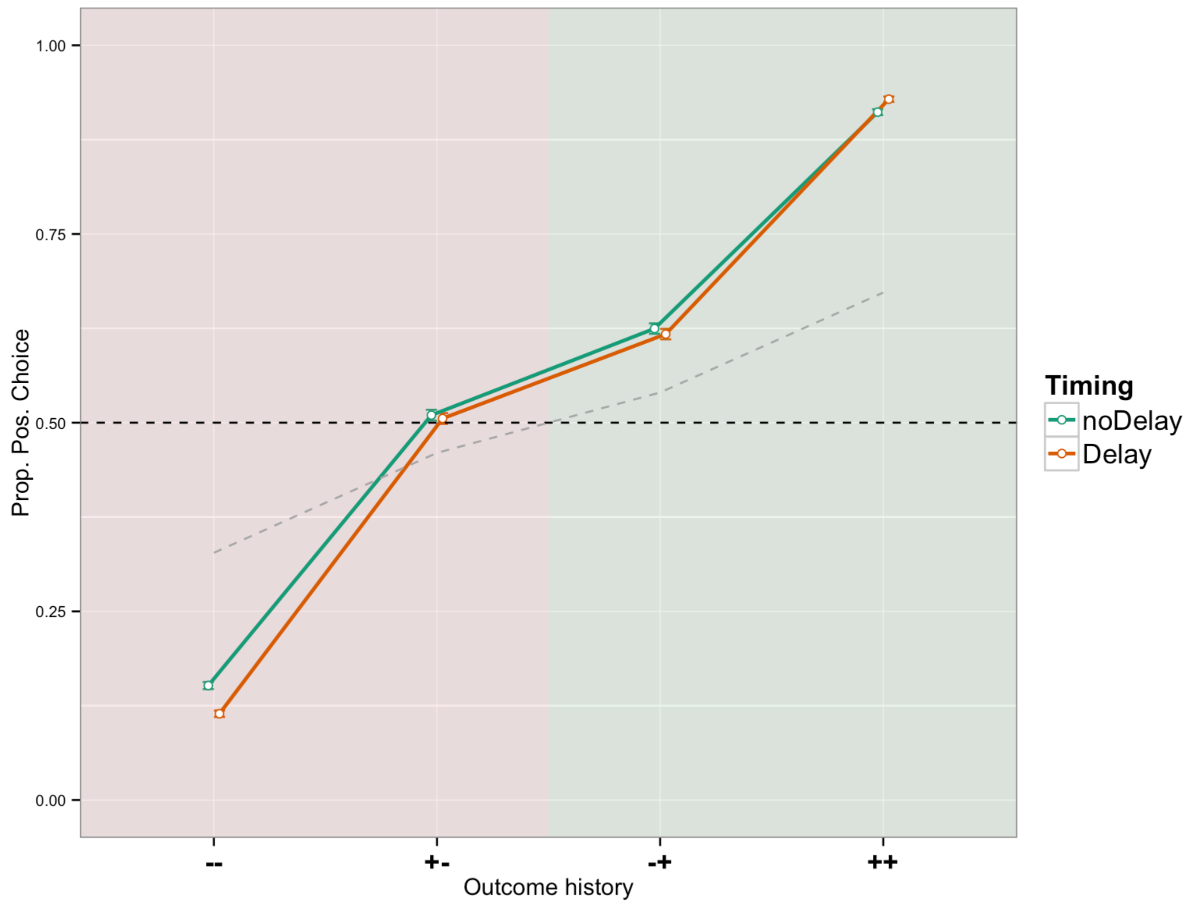


Figure 12. Choice behavior for Experiment 4a. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line. The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

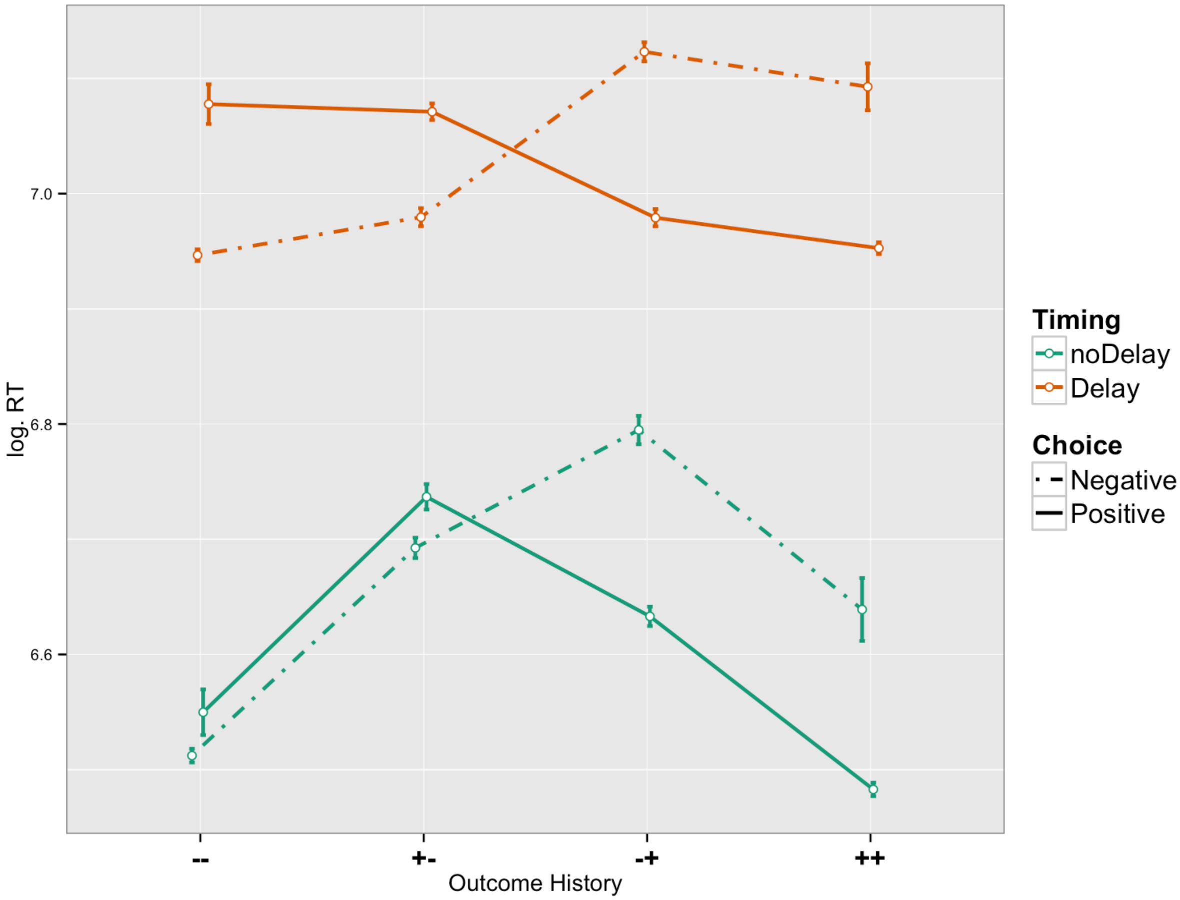


Figure 13. Reaction times for Experiment 4a. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Discussion

Though we didn't see a main effect of condition, we did see an intriguing interaction of condition and outcome such that delay might have an effect specifically when the evidence implies the task is in Negative Mode (this effect is perhaps more visually apparent in three-back plot in Appendix H). When we constrained our analysis to those trials, we saw a significant effect of condition such that delay causes relatively less bias. This is in line with the idea that wishful thinking is an automatic, intuitive process that Type 2 processes might intervene upon and correct.

This is also consistent with our reaction time results from Experiments 2 and 3 where negative mode choices were associated with longer reaction times. Indeed, we see that same effect again in our no delay condition, but see it attenuated in the delay condition, confirming that this difference in response time mediates wishful thinking.

Interestingly, reaction times in the no delay condition were not much different than in Experiments 2 and 3 (Mdn_{RT_Exp4a} = 689 ms)[11] implying that our original timing conditions did not inadvertently impose any delay upon participants.

Though these results were encouraging, we had not predicted this exact pattern. As such, we decided to recalibrate our expectations and attempt a replication. This will be the goal of Experiment 4b.

Finally, we saw an interesting effect of education such that not only was more education associated with relatively less bias, but education also tempered the effect of delay. To explain this, one might posit that education could inculcate a stronger default tendency to mistrust one's

[11] Though this initially appears larger than in Experiments 2 and 3, it should be noted that RTs are anchored from when the response prompt appears. If you anchor the response time from when the outcome appears, the differences are far smaller, (Mdn_{RT_Exp2} = 856 ms, Mdn_{RT_Exp3} = 980 ms, Mdn_{RT_Exp4a} = 789 ms).

intuition, rendering our manipulation redundant on the more highly educated. However, our trust in the effect has to be tempered by the recognition that we have not seen it in any of our other experiments.

Experiment 4b

Method

Participants. We recruited 407 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). All other recruitment details were the same as Experiment 2.

Procedure. The procedure was identical to Experiment 4a with the sole exception that we lengthened the amount of time between showing the outcome on screen and allowing the participant to indicate their judgment from 700 ms to 1200 ms in the delay condition.

Results

10 participants suffered technical problems and a further 54 failed to complete the task for unknown reasons, leaving 343 participants for analysis (M_{age} = 32.1, 45% female). No participants were otherwise excluded.

We ran identical analyses to Experiment 4a with the expectation this time that the effect of condition upon mode choice would be limited to trials with a negative outcome. This expectation was borne out with a significantly negative effect of delay upon likelihood to choose positive mode on trials with a negative outcome ($\beta = -0.17$, $z = -3.65$, $p < 0.001$, see Table 15 for full details). Running the more typical Two-Back Outcome Model with all outcome predictors centered again showed a non-significant main effect of condition ($\beta = -0.06$, $z = -1.61$, $p = 0.107$). Thus our bias measure results replicated the results of Experiment 4a.

Our reaction time results (displayed in Figure 15) likewise replicated Experiment 4a, showing a significant negative effect of choice in the no-delay condition ($\beta = -0.03$, $t = -5.22$, $p < 0.001$) implying participants were quicker to make positive mode choices in that condition, and a significant condition by choice interaction ($\beta = 0.02$, $t = 3.52$, $p < 0.001$) implying that this effect of choice was attenuated in the delay condition (see Appendix G for full details).

Our analysis of demographic effects again showed an education by condition interaction but in the opposite direction ($\beta = -0.1$, $z = -5.2$, $p < 0.001$). Again, this analysis showed a significant main effect of condition ($\beta = -0.06$, $z = -3.42$, $p < 0.001$), but this effect did not survive the simpler model analysis from Experiment 4a that included a by-participant random slope for condition ($\beta = -0.07$, $z = -1.77$, $p = 0.076$). Full details of the demographic results can be seen in Appendix C.

Table 15

Experiment 4b: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode

Fixed Effects	*Estimate*	*Estimate with Outcome t-1 recoded*
Intercept (bias)	0.3***	-0.77***
Outcome t-1	1.06***	2.12***
Outcome t-2	0.94***	0.93***
Delay	-0.07	-0.17***
Ot-1 x Ot-2	0.01	0.02
Ot-1 x Delay	0.1***	0.2***
Ot-2 x Delay	0.06**	0.08**
Ot-1 x Ot-2 x Delay	-0.02	-0.04

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. * $p < .05$, ** $p < .01$, *** $p < .001$

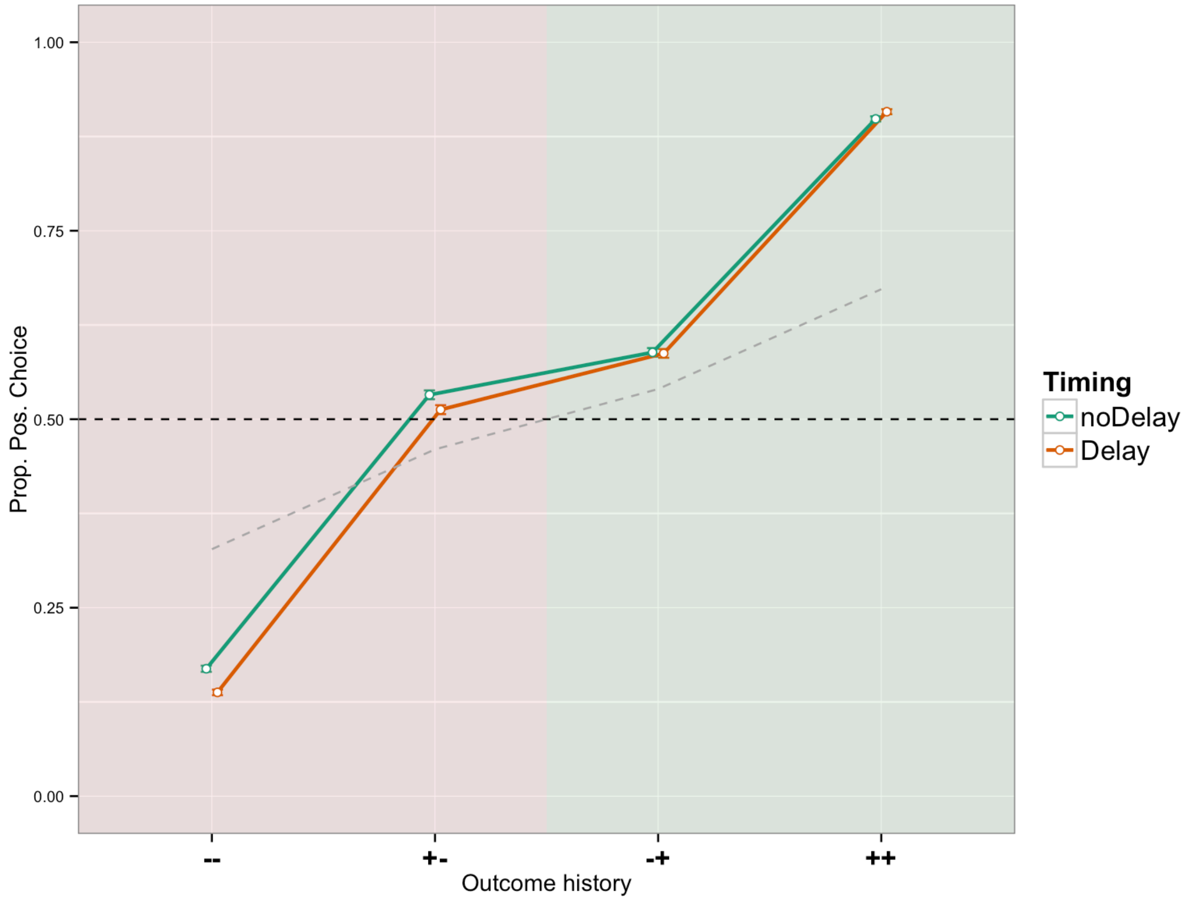


Figure 14. Choice behavior for Experiment 4b. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line. The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

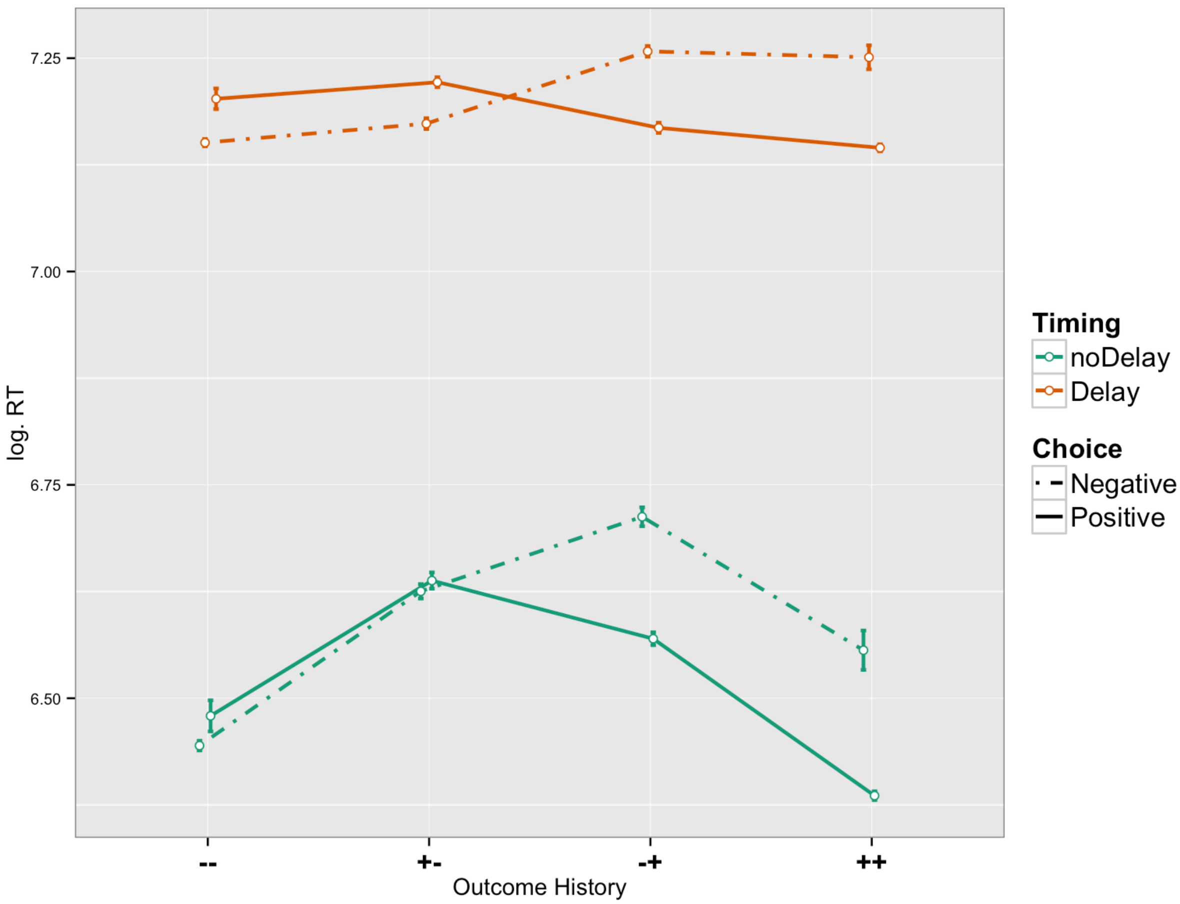


Figure 15. Reaction times for Experiment 4b. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Discussion

Our results exactly replicated those from Experiment 4a, confirming that delay reduces wishful thinking specifically when the current evidence implies that the game is in Negative Mode. This is strong evidence that wishful thinking is an automatic process, but one that can be overridden by Type 2 thinking.

It might be wondered why the effect is specific to wishful thinking in the face of undesirable evidence, and why delay doesn't also restrict any bias that might be present when the task seems to be in Positive Mode. It could be that there is less wishful thinking to override when the task seems to be in Positive Mode. Given our current task set up, we are not well placed to evaluate this: in order to test for variation in wishful thinking across different levels of evidence we require a non-linear or interaction effect, or otherwise the variation in bias will get confounded with the linear effect of evidence. This is why we can test for ambiguity effects—since they can be characterized by the interaction of previous outcomes—but we cannot compare different levels of bias simply for desirable versus undesirable evidence, since any such variation in bias would simply get absorbed into the one-back outcome predictor. In order to compare the magnitude at these two levels we would need a control slope, potentially provided by a similar state estimation task that requires the same evidence integration, but which doesn't have any wishful thinking component, e.g. by equating the mean reward in the two states.

While that explanation posits a difference in the extent of bias across positive and negative outcome trials, and as such a difference in opportunity for Type 2 processes to intervene, an alternative explanation is that there is a difference in the likelihood of Type 2 intervention itself across positive and negative outcome trials. Type 2 intervention has been shown to inversely correlate with subjective confidence (Thompson, Prowse Turner, &

Pennycook, 2011). Since biased choices are aligned with the evidence when the evidence implies that things are good and not aligned with the evidence when the evidence implies things are not good, subjective confidence is presumably lower for biased choices made in the face of undesirable evidence, putting them under greater danger of suffering a Type 2 intervention.

Our inclination would be to side with this latter explanation. Looking at our reaction time data from Experiment 2, and for the no-delay condition in Experiments 4a and 4b we can see a very close alignment of reaction time for positive and negative choices when the evidence is undesirable (admittedly less so for Experiment 3). This is not the case in the delay condition where positive choices now take longer than negative choices when the evidence is undesirable. This might be taken to indicate a difference in confidence in the judgment (see Experiment 2 Discussion), and insofar as it does, would provide support for this second explanation.

However, fully answering these issues will require further work. In any case, we can congratulate ourselves on having landed our first clear blow in the struggle to place wishful thinking in the dual-process framework.

Experiment 5

For Experiment 5 we wanted to test the effect of cognitive load on wishful thinking. Imposing cognitive load is a highly effective method of disrupting the operation of working memory—perhaps the only necessary defining feature of a Type 2 process (Evans & Stanovich, 2013; Gilbert & Hixon, 1991; Logie, Gilhooly, & Wynn, 1994).

Experiments 4a and 4b showed a relative decrease in bias given a relative increase in Type 2 processing. But we can still wonder what the contribution of Type 2 processing is to wishful thinking when decision time is unrestricted. It is possible that there is some level of restraint even without the enforced delay. On this view, by restricting working memory capacity we might release wishful thinking from whatever default level of constraint it usually operates under and see an increase in bias in our task.

This describes a view where the default level of wishful thinking we see is the outcome of a tussle between facilitating Type 1 and inhibitory Type 2 processing. But it might be the case that wishful thinking doesn't typically operate under any Type 2 constraint at all but is rather the result of an automatic Type 1 process calibrated so as to avoid censure from Type 2 thinking. On this view, restricting working memory capacity should not affect the strength of the wishful thinking bias.

Bias might be expected to decrease under cognitive load if working memory facilitates it. Given our results in Experiment 4, we don't expect this to be the case (see also Lieberman et al., 2001), but some evidence that it could be so comes from Kahan et al. (2013) who found that numeracy—presumably closely related to working memory capacity—predicted how much a participant would succumb to the confirmation bias when considering numerical evidence for a favored hypothesis.

The cognitive load component of the task began with the participant being presented with a letter at the beginning of the session. They were required to internally update that letter with the succeeding letter in the alphabet on every trial until—some unpredictable amount of trials later—they were prompted to report it. After which, the cycle began anew. Performance on the load task was rewarded to keep participants motivated.

We required participants to update the letter on each trial—as opposed to merely retaining it in memory—in order to help ensure that they were not writing it down or otherwise storing it externally when performing the task. We chose letters in preference to numbers as the memory target just in case numbers might have interfered with the processing of reward outcomes.

Method

Participants. We recruited 210 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). All other recruitment details were the same as Experiment 2.

Procedure. The procedure was the same as Experiment 2 except for the differences noted below.

Training. As well as being instructed as per Experiment 2, participants also received training on the memory task component of the game. Before the relevant session, we informed participants that they would be asked to hold a letter in mind as they played the game and that they would be asked to tell us what the letter was every so often. For every correct answer they would get 100 tokens. The task was broken down into three steps and explained in greater detail, including that they would initially be prompted with the letter "A", and that they should increment the letter on every trial until they were prompted to return it. Participants then played a practice session of 7 trials duration that included the cognitive load component and familiarized them with its instantiation in the game. Participants were asked a multiple choice comprehension question regarding when they should increment the letter. It was also requested, just prior to beginning the session, that they do their best to keep the letter in mind and refrained from writing or typing it down.

Casino game task. Participants took two sessions of the game: a *no-load* and a *load* session. Each session was 75 trials long. The no-load session proceeded as per Experiment 2. The load session included the cognitive load component described below. The order of sessions was randomized across participants.

The load session began by presenting participants with a letter “A” in the middle of the screen. This was the signal to begin holding the letter “A” in mind. Participants pressed the spacebar to continue and thereafter proceeded to play the game as normal (i.e. as in Experiment 2). Participants were tasked with updating the target letter with the next letter of the alphabet for every pull of the lever they made during the game. After a certain amount of trials, participants saw a “?” on the screen. This was the prompt to guess the target letter. They made their guess using the corresponding key on their keyboard and were rewarded with 100 tokens if they were perfectly accurate and 50 tokens if they were only one letter off the correct letter. We rewarded them in this graded manner so that they wouldn’t give up early in the sequence if they started to feel unsure about the target letter. Participants were given immediate feedback after their choice which was displayed on screen for 800 ms before the beginning of the next trial.

On the next trial they were again prompted with the letter “A” and the sequence started afresh. In total, there were 10 sequences—from prompt to probe—in the task. The shortest sequence duration was 3 trials long, the next shortest was 4 trials long, and so on up to 12 trials. Thus the total duration of all sequences was 75 trials, i.e. the total length of the session. The sequences were randomly ordered. The variable length of the sequences and their random order prevented participants from accurately forecasting the timing of the next probe and thereby being able to calculate the target letter in advance.

As per Experiment 3, and in contrast to Experiment 2, the duration of the cogwheel rotation was 300 ms and the time for which the participant’s choice was highlighted on the screen was 400 ms.

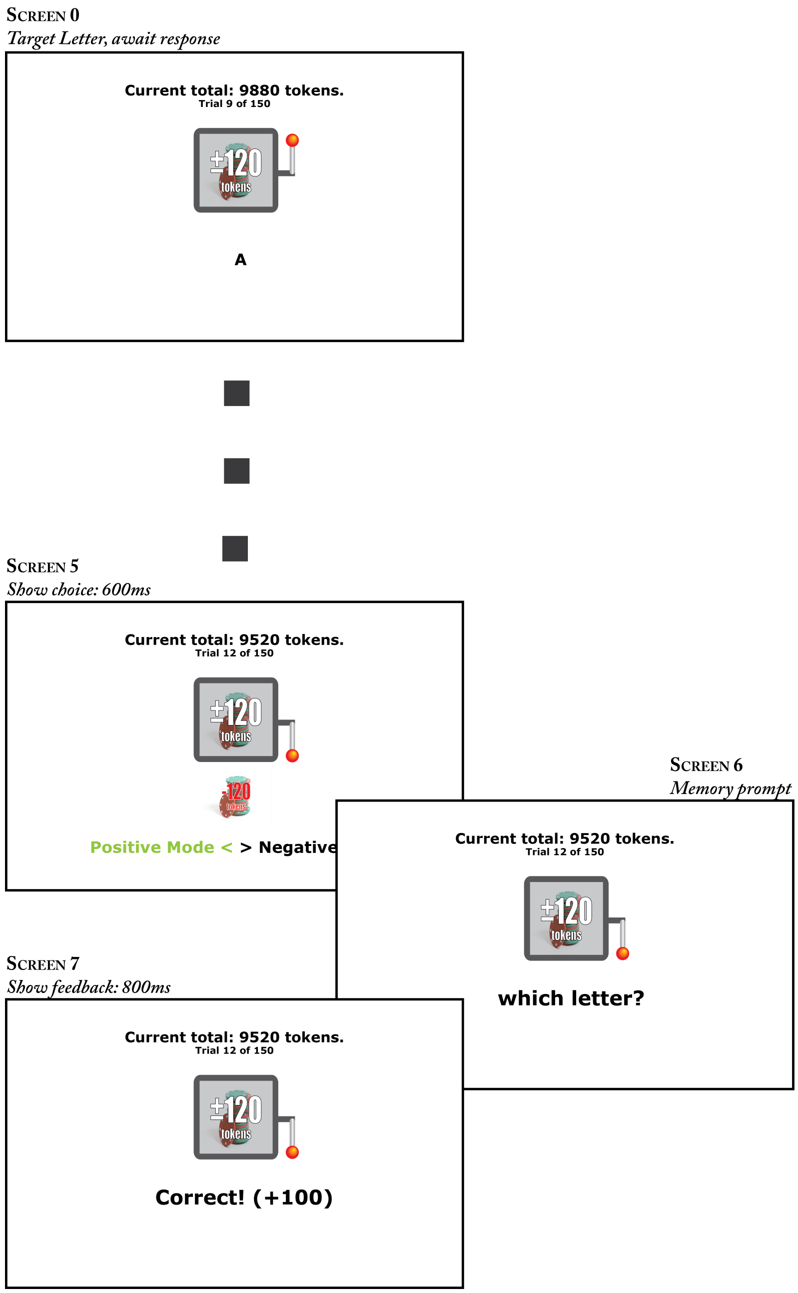


Figure 16. Screenshots and timing for load task component of Experiment 5.

Exit Survey. Just like in Experiment 3, we added a measure to our exit survey to gauge whether participants were using workarounds in the cognitive load task. See Appendix C for wording.

Results

6 participants suffered technical problems leaving 210 participants for analysis (M_{age} = 33.7, 54% female). 2 participants reported writing down the letter for the cognitive load task. These participants were not excluded. No participants were otherwise excluded.

Two-Back Outcome Model Measure. We used the model described in Experiment 4a—with the reference level for condition set to no-load—to test for the effect of cognitive load on bias. The results showed no main effect of condition ($\beta = 0.01$, $z = 0.32$, $p = 0.75$), but did show an interaction of condition with the two-back outcome variable ($\beta = -0.11$, $z = -3.50$, $p < 0.001$). The negative sign of the parameter estimate implies that the two-back outcome had less effect on choice in the load condition. This can be seen in Figure 17. Full details of the model are in Table 16.

Since the load condition appeared to specifically affect the ambiguous trials, we tried focusing our analysis only on those trials. Using a mixed effects logistic regression model to look at the relationship of choice and condition only on trials that differed in their previous two outcomes, we entered previous outcome and condition as fixed effects and by-participant random intercepts and slopes for condition as random effects. Even when focusing our analysis in this way, there was no effect of condition ($\beta = 0.01$, $z = 0.14$, $p = 0.886$).

Load task analysis. On the whole participants performed well—though not perfectly—on the load task ($M_{correct}$ = 7.42, $SD_{correct}$ = 2.91, see Figure 18). To look at the effect of load task performance on choice we ran a variation of our Two-Back Outcome Model where we focused specifically on the load condition and entered the number of load probes correct as an extra fixed effect. This meant we had the full interaction of both previous outcomes and load performance as fixed effects. Convergence difficulties meant that we only had random intercepts for participant entered as random effects. The results showed no effect of load performance on choice ($\beta = -0.01$, $z = -0.31$, $p = 0.76$), but showed an interaction of load performance with both one back outcome ($\beta = -0.02$, $z = -3.1$, $p = 0.002$) and two back outcome ($\beta = 0.05$, $z = 5.85$, $p < 0.001$), such that participants who performed better seemed to rely less on the more recent outcome and more on the less recent outcome than those who performed less well. This is the opposite pattern to the effect of condition generally, implying that while load pushed people towards a greater reliance on more recent outcomes, those who could handle the load task better were not as affected. The full results are in Table 17.

Reaction Time. We analyzed reaction time (see Figure 19) using the same model as in Experiment 4a. Results showed the usual effect of choice on reaction time ($\beta = -0.03$, $t = -2.93$, $p = 0.004$). There was also a significant effect of condition ($\beta = 0.34$, $t = 9.92$, $p < 0.001$) such that reaction times were longer in the load condition. Though reaction time was longer in the load condition, there was no choice by condition interaction ($\beta = -0.02$, $t = 1.16$, $p = 0.248$). In this respect the results differed from Experiments 4a and b. Appendix G has the full details.

Demographics. We tested for the effect of demographic factors as per Experiment 4a. No significant main effects of any demographic factor were found.

Table 16

Experiment 5: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.24	0.05	5	< 0.001
Outcome t-1	1.22	0.02	52.55	< 0.001
Outcome t-2	1.01	0.02	43.89	< 0.001
Load	0.02	0.06	0.32	0.75
Ot-1 x Ot-2	0.02	0.02	0.88	0.378
Ot-1 x Load	0.04	0.03	1.3	0.195
Ot-2 x Load	-0.11	0.03	-3.5	< 0.001
Ot-1 x Ot-2 x Load	0.002	0.03	0.07	0.942

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode.

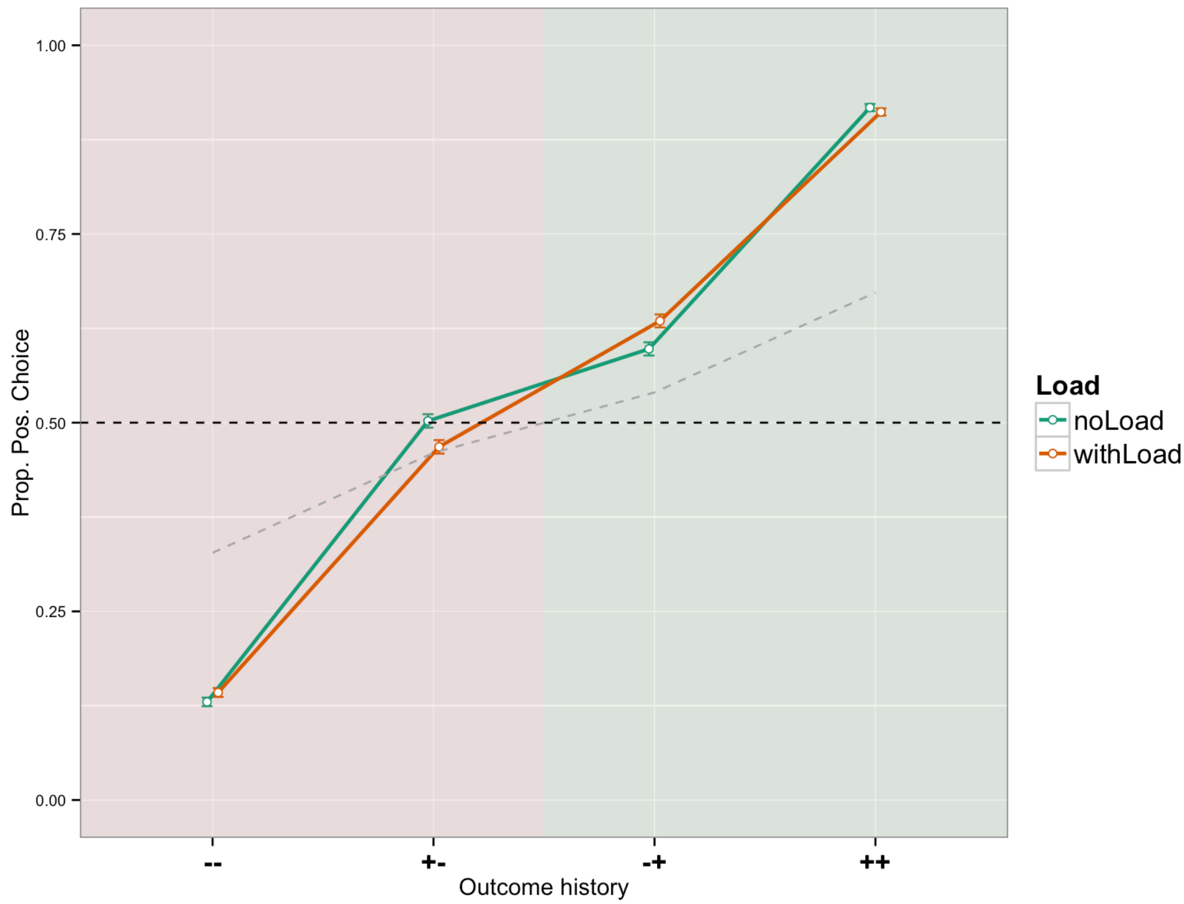


Figure 17. Choice behavior for Experiment 5. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line. The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

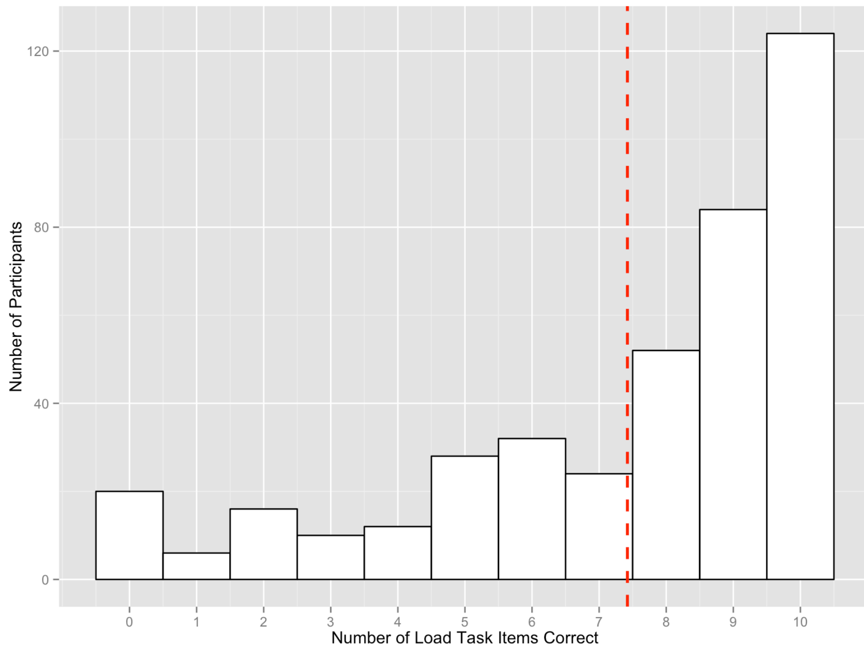


Figure 18. Histogram of performance on load items. The red dotted line indicates mean performance.

Table 17

Experiment 5: Effect of previous trial outcomes and load task performance on likelihood to choose Positive Mode

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.3	0.14	2.21	0.027
Outcome t-1	1.45	0.06	23.19	< 0.001
Outcome t-2	0.57	0.06	9.24	< 0.001
Load Performance	-0.01	0.02	-0.31	0.76
Ot-1 x Ot-2	-0.02	0.06	-0.37	0.711
Ot-1 x Load Perf.	-0.02	0.01	-3.1	0.002
Ot-2 x Load Perf.	0.05	0.01	5.85	< 0.001
Ot-1 x Ot-2 x Ld Perf.	0.01	0.01	0.8	0.426

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode.

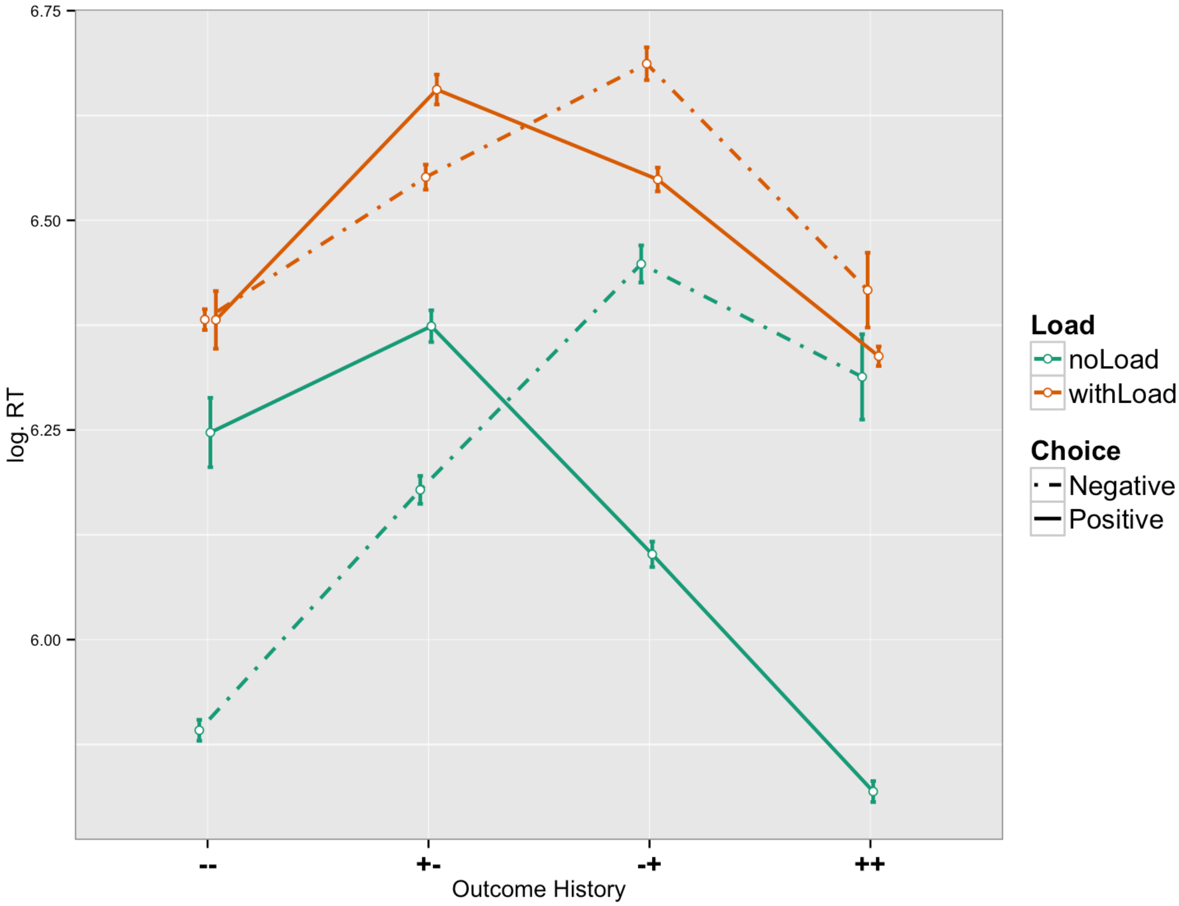


Figure 19. Reaction times for Experiment 5. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Discussion

Both choice data and reaction times show that our load manipulation had an effect on behavior. In terms of choice, it seems that when under load, participants' decisions were less affected by the outcome from two trials ago which is consistent with restricted working memory. In terms of reaction time, decisions made under load took longer, again consistent with restricted working memory capacity. Also, though participants did well on the load task, they didn't perform perfectly well, implying both that they engaged with the task, but didn't find it trivially easy.

However, despite the efficacy of the manipulation, we saw no effect at all on bias. It could be argued that though our manipulation did enough to affect performance, it was not sufficiently potent enough to disrupt bias. A stronger load manipulation might yet affect it. Perhaps, but it should also be noted that the bias did not vary with load task performance. Even those for whom the task was most difficult did not change their level of bias.

The implication is that Type 2 processes are not actively involved in wishful thinking, neither to enhance nor diminish it. This accords with the idea mentioned earlier that wishful thinking is calibrated to stay inconspicuous to Type 2 processing. However it should be noted that our task has certain attributes that might facilitate wishful thinking's success in this regard: first, the inducement to bias one's thinking is not as strong in our task as in, for instance, a task that requires you to make a judgment with implications for your future health and well-being (Kunda, 1987; Quattrone & Tversky, 1984). Second, the evidence upon which you base your decision in this task is simple and unidimensional. Where judgments both require integrating multi-faceted evidence, and where there is a lot on the line, there may be a more active dynamic between Type 1 and Type 2 processes in wishful thinking. This is an interesting avenue for future research.

Experiment 6a

Our final manipulation designed to separate the contributions of Type 1 and Type 2 processes to wishful thinking was to prime different thinking styles by giving participants a short writing assignment, with the idea that different thinking styles would affect the balance of contribution from the different types of processes. We chose three different thinking styles: using your gut or intuition; reflecting; or taking your time. The first two were chosen as previously successful manipulations of dual-process contributions to decision-making (Rand et al., 2012; Shenhav, Rand, & Greene, 2012) with *intuition* assumed to promote the contribution of Type 1 processes at the expense of Type 2 processes, and *reflection* assumed to promote the Type 2 processes at the expense of Type 1 processes.

Taking your time (*take time*) was included as a supplemental condition chosen on the basis of unpublished results from a different experiment with similar aims (M. Ho, personal communication, August, 2014). In that experiment, reflection was primed via instructional guidance to take one's time when making decisions. We were curious to know if a prime to take one's time would behave more similarly to an intuition or to a reflection prime. This was very much an exploratory manipulation. Intuition and reflection served as our primary focus.

As mentioned, previous use of these essay primes were designed with the assumption that the reflection prime, relative to the intuition prime, would increase the likelihood of Type 2 processes intervening on default intuitions resulting in putatively more rational judgments (Rand et al., 2012; Shenhav et al., 2012). That remains the starting assumption here. In that case we should see decrease in wishful thinking in the reflection condition.

However—with the exception of the unpublished results mentioned previously—this sort of prime has not been used in such an obviously motivated decision-making context[12] and it is possible that it will behave differently, especially given our earlier consideration of motivated reasoning within the dual-process framework. One possibility is that the reflection prime will promote the use of cognitive processes involved in rationalizing without promoting increased awareness and review of decisions. Another possibility is that the reflection prime will prime the participant's perception of having been objective which may decrease the drive or perceived need to actually be so. Both of these possibilities would predict that our reflection prime would increase bias.

A final alternative is that the thinking style prime may alter how participants weight the evidence when updating their task state belief, just like in Experiment 5. In that case we should expect to see an interaction of prime and outcome, but not necessarily an effect on bias.

Method

Participants. We recruited 423 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). We paid participants $1.50 for their participation. All other recruitment details were the same as Experiment 2.

Procedure. The procedure was the same as Experiment 2 except for the differences noted below.

Essay Prime. Participants were randomly assigned to one of three essay prime conditions: intuition, take time, or reflection. The essay prime required participants to describe a time that—

[12] Though of course belief in god—the target judgment of Shenhav, Rand, and Greene (2012)—and how much to contribute in a public goods game—the target judgment of Rand, Greene, and Nowak (2012)—have presumably some motivational component. Indeed it *might* be argued that the effect of the reflection prime in Rand et al. (2012) was to facilitate motivated reasoning and thereby increase selfish decisions, much like the opportunity to flip a coin in Batson (2007).

depending upon condition—their intuition/first instinct, taking their time, or reflecting lead to a good outcome for them.

We were worried that the amount of instruction our task required would make it difficult to effectively prime participants. If we had them write the essay prior to the instructions, we risked the instructions interfering with the effect of the prime. If we had them write the essay after the instructions and immediately prior to taking the task, the curious insertion of an unrelated tasks might have confused participants, or rendered them suspicious about the purpose of the essay. Our solution was to frame the essay as a memory task. We initially had them write the essay at the very start of their participation, with a warning that they would have to recall the topic of their essay later and would get a bonus of 10c if they did so accurately. Then, immediately prior to the Casino Game we had them recall their essay. All participants received the memory bonus.

We initially asked them to write their essay using the following prompt (here tailored for the intuition condition):

> However, before we start, we would like you to write about an event from your life. We want to use this as a memory measure. We will ask you to recall this event later during the task.
>
> **Please write a paragraph (approximately 8-10 sentences) describing a time your intuition/first instinct led you in the right direction and resulted in a good outcome.**

Immediately prior to beginning the Casino Game, we asked them to recall their essay with the following prompt:

> Before you proceed to the main task, we would like you to recall the event you wrote about earlier, and describe that event again in the space below.

If the event that you write about now matches the event you wrote about earlier, we will pay you an extra bonus of 10 cents, on top of any other bonus you earn today. We are not concerned with the details, just so long as it is clear that you are describing the same event as before.

As a reminder, the event described a time when your intuition/first instinct led you in the right direction and resulted in a good outcome.

Results

4 participants suffered technical problems leaving 419 participants for analysis (M_{age} = 35.1, 57% female). No participants were otherwise excluded.

Two-Back Outcome Model Measure. We used a mixed effects logistic regression model to test the effect of our essay primes on mode choice. The model was similar in construction to Experiment 4a but since we used a between-participant manipulation in this study specifying a random-effects slope for condition was redundant. As such, the full model contained the full interaction of condition—with intuition set as the reference level—with both previous outcome variables as fixed effects, and the by-participant random intercepts as random effects.

The results showed no effect of either take time (β = -0.01, z = -0.08, p = 0.936) or reflection (β = 0.07, z = 0.79, p = 0.429), but reflection did show an interaction with both the outcome variables (see Table 18 for full model details).

Visual inspection of the results (see Figure 20) suggested that—much like Experiments 4a and 4b—this reflection interaction effect may be primarily driven by an increased likelihood to choose positive mode when the game was more likely to be in negative mode, which is to say on those trials immediately preceded by a negative outcome. To explore this we split our data by whether the previous trial was positive or negative and ran our mixed effect model separately on both sets of data (since we split the data based on previous outcome, the previous outcome regressor was dropped from the models as redundant). The results showed that there was a significant effect of reflection on the negative trials ($\beta = 0.38$, $z = 2.63$, $p = 0.009$) but no significant effect on the positive trials. See Table 18 for full model details.

Reaction Time. We analyzed reaction time (see Figure 21) using the same model as in Experiment 4a with the difference that since our manipulation was now a between-participants condition we did not include a by-participants random slope estimation for the effect of condition in the random effects structure. Interestingly we saw no main effect of choice in our intuition condition ($\beta = -0.01$, $t = -0.63$, $p = 0.532$) but we did see an interaction of choice and the reflection condition ($\beta = -0.05$, $t = -2.91$, $p = 0.004$), implying that the effect of choice we have seen throughout most of our experiments was only operative in the reflection condition. Perhaps surprisingly, reflection was our quickest condition, but this difference was not significant. Neither reflection nor intuition were significantly quicker or slower than intuition (see Appendix G for full model details).

Demographics. We tested for the effect of demographic factors as per Experiment 4a and Experiment 5. No significant main effects of any demographic factor were found (all $p > 0.05$).

Table 18

Experiment 6a: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode

Fixed Effects	*Estimate*		
	All trials	*Negative outcome trials only*	*Positive outcome trials only*
Intercept (bias)	0.24***	-1.18***	1.7***
Outcome t-1	1.24***		
Outcome t-2	1.02***	1.11***	1.12***
TakeTime	-0.01	0.06	-0.05
Reflection	0.07	0.38**	-0.2
Ot-1 x Ot-2	0.02		
Ot-1 x TakeTime	-0.04		
Ot-1 x Reflection	-0.21***		
Ot-2 x TakeTime	-0.03	-0.04	-0.02
Ot-2 x Reflection	-0.13***	-0.18***	-0.13**
Ot-1 x Ot-2 x TT.	-0.01		
Ot-1 x Ot-2 x Refl.	-0.003		

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. * $p < .05$, ** $p < .01$, *** $p < .001$

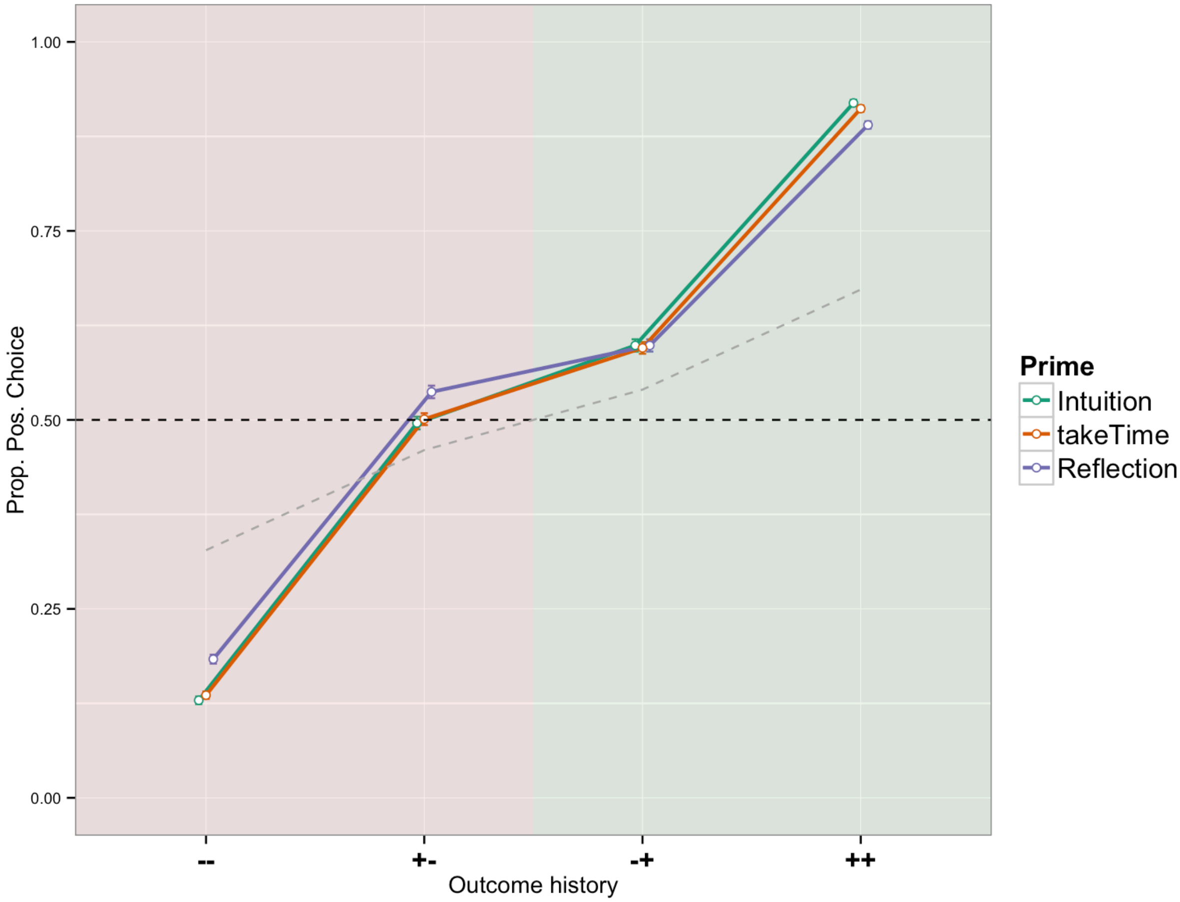


Figure 20. Choice behavior for Experiment 6a. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line. The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

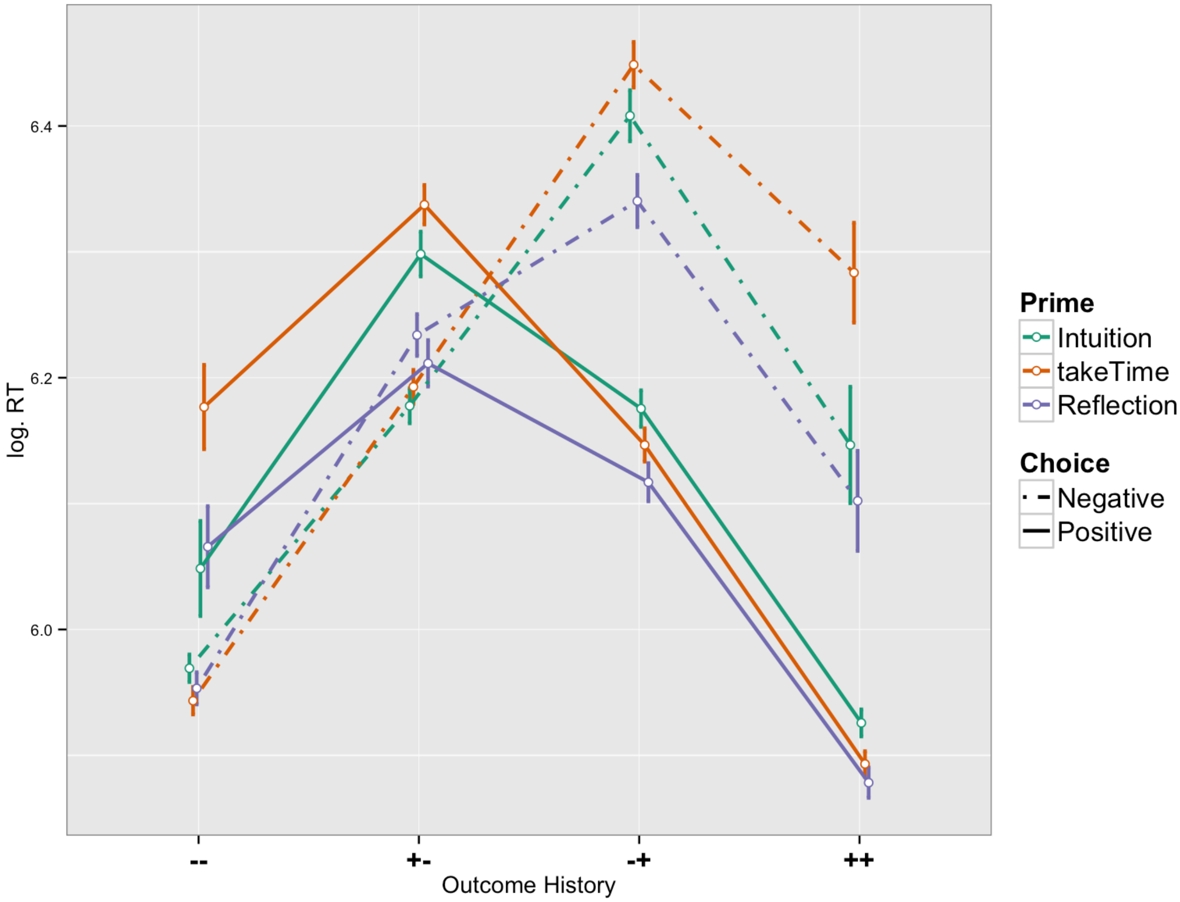


Figure 21. Reaction times for Experiment 6a. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Discussion

There were three main results in Experiment 6a. First, we see an increase in bias for the reflection condition localized to the negative trials. Second, we did not see a main effect of condition on reaction time. Third, the typical reaction time pattern we have seen across previous experiments, where positive judgments are made quicker than negative judgments, was only seen in the reflection condition.

While it is hard to explain why the effect of choice that we have seen so often in previous experiments was not present in the intuition condition—or the take time condition for that matter—the fact that it was present in the reflection condition is at least consistent with the increased bias in that condition. So it would seem that our reflection essay prime had the unexpected effect of increasing wishful thinking which may be reflected in its effect on reaction time.

However, we did not see the general increase in reaction time that Type 2 processes are associated with. The implication might be that if our reflection essay prime is indeed increasing wishful thinking, it is not obviously doing so through a manipulation of Type 2 processing. As mentioned in the introduction to the experiment, there may be alternative explanations that could account for this. However, given the numerous unexpected aspects of our results—the increase in bias in the reflection condition, the localization of that increase, and the lack of a difference in average reaction time across condition—we wanted to replicate the results before reading too much into them. This was the aim of Experiment 6b.

For Experiment 6b we dropped the take time condition. The prime didn't discernably alter participants' reaction time and exhibited an almost identical pattern of judgment as the intuition condition. This might be taken to imply that reflection was our "active" manipulation, i.e. the one that shifted participants' behavior from the default, rather than intuition, which would accord with the results of the pervious experiments in this chapter. In any case, since the take time condition was not an initially high priority, and since it failed to distinguish itself from intuition, it was seen as redundant and dropped from the replication.

Experiment 6b

Method

Participants. We recruited 843 participants from the Amazon Mechanical Turk online labor marketplace (www.mturk.com). All other recruitment details were the same as Experiment 2.

Procedure. The procedure was identical to Experiment 6a, except that we no longer had a take time condition, and one other inadvertent difference. As reported, Experiments 3, 4a, 4b, and 5 all differed from Experiment 2 in regard to two timing parameters: the time that the cogwheel rotated on-screen before the trial outcome was displayed, and the time that the participant's choice of mode was highlighted on screen before the beginning of the next trial. In Experiment 2, these parameters were both set as 600ms. In the four other studies mentioned, they were set at 300 ms and 400 ms respectively. As described earlier, those changes were made in an effort to make the game run more smoothly in response to participant feedback from Experiment 2. Experiment 6a was launched prior to the full analysis of that feedback and as such, retained the timing parameters of Experiment 2. However Experiment 6b was launched using the same core computer code as the other studies and as such, included the updated timing parameters. The updated code was used because it allowed for paying bonuses with much greater efficiency. The change in timing parameters was an oversight.

Results

3 participants suffered technical problems and a further 220 failed to complete the task, leaving 620 participants for analysis (M_{age} = 33.2, 57% female). It is unclear why so many participants dropped out. Our initial aim was to recruit 600, and the vast majority of the dropouts did so after a screen or two of the experiment. As such, we suspect that high number of participants requested attracted a rush of more casual Mechanical Turk workers, who signed in to our task to see what it was like, before deciding it wasn't for them. No participants were otherwise excluded.

All analyses were identical to Experiment 6a with the exception that there was no take time condition. However, in this iteration we expected a condition effect specifically in the negative trials, i.e. when the choice was preceded by a negative outcome. We did not see such an effect ($\beta = 0.1$, $z = 1.13$, $p = 0.259$), nor did we see either a condition effect across all trials ($\beta = -0.03$, $z = -0.49$, $p = 0.624$), or an effect localized in the positive trials ($\beta = -0.11$, $z = -1.3$, $p = 0.193$). However, we did see a replication of the interaction of reflection with outcome ($\beta = -0.09$, $z = -5$, $p < 0.001$) that we saw in Experiment 6a, implying that previous outcome had less impact upon choice in the reflection condition than in the intuition condition. The full results are shown in Table 19.

As in Experiment 6a, we see no main effect of reflection on reaction time ($\beta = -0.04$, $t = -0.82$, $p = 0.41$). However, in contrast to Experiment 6a, we did see a main effect of choice in the intuition condition ($\beta = -0.05$, $t = -6.21$, $p < 0.001$). This effect was not significantly different in the reflection condition ($\beta = 0.002$, $t = 0.24$, $p = 0.812$), implying that the positive judgments were faster than negative judgments in both conditions.

Finally, no significant main effects of any demographic factor were found (all $p > 0.05$).

Table 19

Experiment 6b: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode

Fixed Effects	*Estimate*		
	All trials	*Negative outcome trials only*	*Positive outcome trials only*
Intercept (bias)	0.22***	-1.06***	1.56***
Outcome t-1	1.17***		
Outcome t-2	1.08***	1.18***	1.16***
Reflection	-0.03	0.1	-0.12
Ot-1 x Ot-2	-0.01		
Ot-1 x Reflection	-0.09***		
Ot-2 x Reflection	-0.08***	-0.12***	-0.06*
Ot-1 x Ot-2 x Refl.	0.02		

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. * $p < .05$, ** $p < .01$, *** $p < .001$

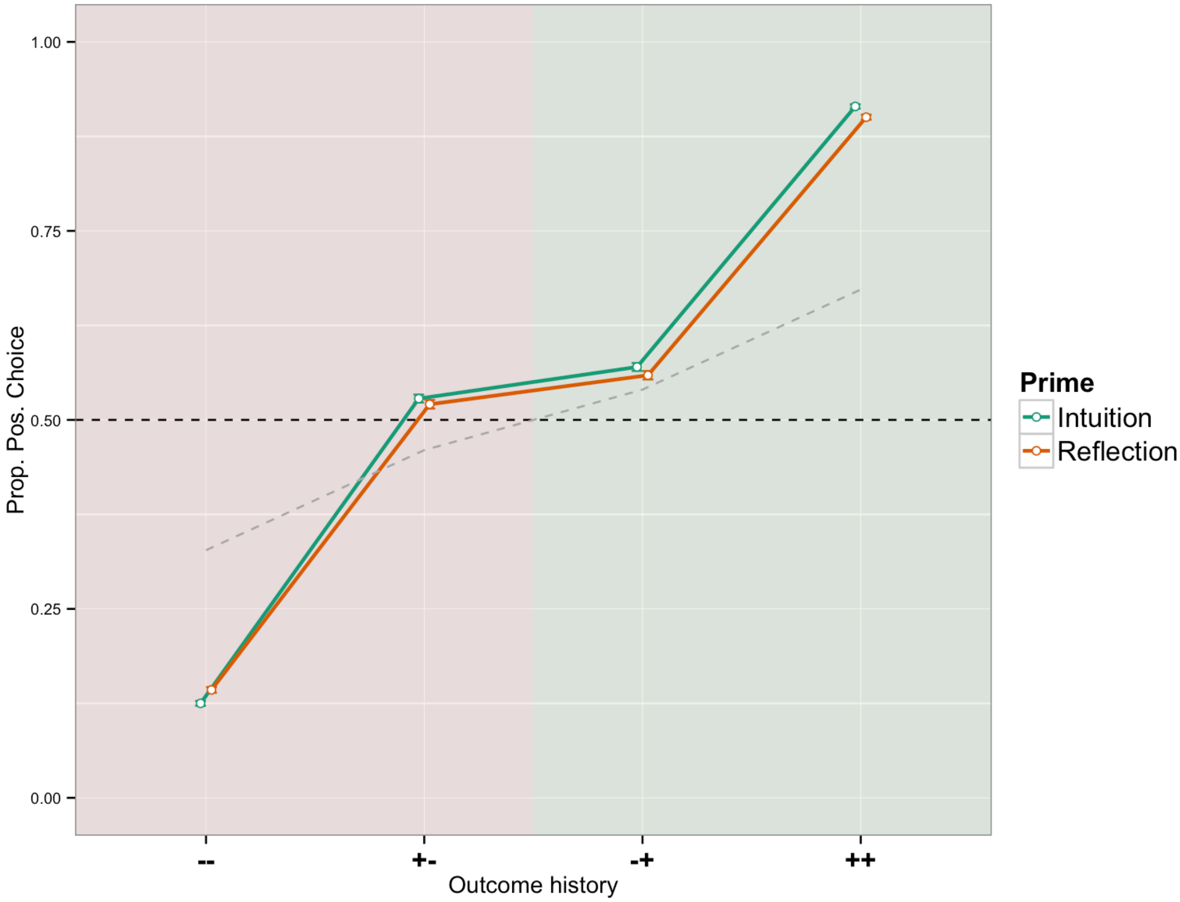


Figure 22. Choice behavior for Experiment 6b. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008) The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line. The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

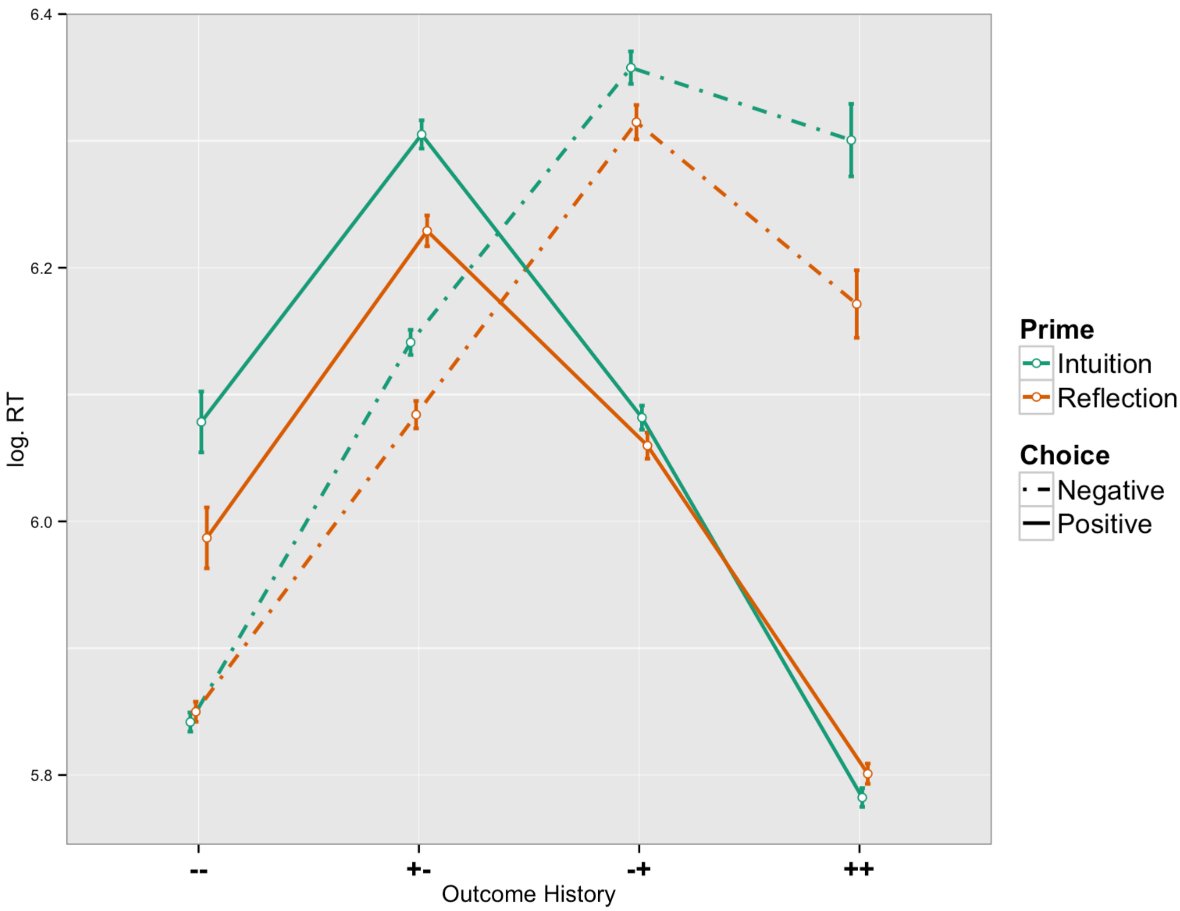


Figure 23. Reaction times for Experiment 6b. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the log transform of reaction time. Error bars were calcaulted with the within-participant correction outlined in Morey (2008). The points in each series have been laterally shifted relative to one another to avoid potential superimposition.

Discussion

The main, and crucial, difference between our results here and Experiment 6a is that we no longer see any increase in bias in the reflection condition. Otherwise we see a very similar pattern of results. Our prime manipulation did affect judgments across both experiments such that the effect of previous outcomes was softened in the reflection condition. That is, previously rewarding outcomes were less likely to cause the participant to choose Positive Mode, and previously punishing outcomes were less likely to cause the participant to choose Negative Mode in the reflection condition than in the intuition condition. This needs to be distinguished from the results of Experiment 5 where we also saw an interaction of condition and outcome. In that case participants were *more* affected by the most recent outcome when under load, but *less* affected by the outcome from two trials ago meaning that, under load, participants relied more upon recent outcomes at the expense of their reliance upon less recent outcomes. Here participants were less sensitive to *both* recent and less recent outcomes.

So it would seem then that our prime did not affect wishful thinking but rather affected the way participants weighted the evidence. Plausibly this could have been due to an increase in Type 2 processing. Increased reflection, for instance, could result in greater skepticism about new evidence. However, our reaction time results speak against this. Again, we saw no main effect of condition on reaction time. Any amount of explicit reflection would presumably take some amount of time. As such, we contend that our prime didn't so much increase reflection during the game so much as cause a retuning of some sort of implicit learning parameters.

We did see one reaction time difference between conditions which could be considered to support this view. As per usual, there was a significant interaction of choice and outcome on reaction time, i.e. participants are quicker to go with the evidence than they are to go against it.

This effect was significantly attenuated in the reflection condition (see Appendix G). That is, participants were relatively slower to agree and relatively quicker to disagree.

We merged both data sets to confirm these results across both experiments, performing the same reaction time and judgment analyses on the merged data set as we did on the two individual data sets. Perhaps unsurprisingly, given the larger number of participants in Experiment 6b, the results of the merged analysis exhibited the same pattern as those here. See Appendix I for full details.

Why didn't our essay prime affect wishful thinking? As discussed in Experiment 5, it may be necessary to give participants something meatier to chew upon. On any particular decision in our task, there isn't much to reflect upon as the evidence only varies along a single dimension. A setup where the evidence varies along multiple dimensions might provide increased opportunity for processes such as biased search, attribute reweighting, and selective rule following to take effect, and thereby increase the chances of finding an effect of thinking style upon those processes.

On the other hand, the essay primes themselves may simply not be powerful enough to affect Type 2 processes over the course of a repeated decision-making task. Type 2 processes are effortful and may be difficult to maintain for the whole duration. Indeed, the two previous success cases for the manipulation were single decision tasks (Rand et al., 2012; Shenhav et al., 2012).

In any case, any subsequent investigation would have to contend with the effect of the primes on evidence weighting that we found here. In order to tease any effect on wishful thinking out from underneath the evidence weighting effect, we would need to compare our current task to a control version of the task with two neutral modes rather than a rewarding and punishing mode. Any effect common to the two conditions would be attributable to evidence weighting. Any residual effect of the primes in the wishful thinking condition could be safely called an effect on wishful thinking.

This remains a direction for future investigation. As it stands, though we have seen an effect of thinking style upon evidence weighting in our task, this did not seem to affect wishful thinking.

Concluding remarks

Experiments 4a and 4b showed that delaying decision time decreased the bias. And even though we successfully disrupted working memory in Experiment 5, the bias was unaffected. On the whole, the results of chapter 3 show that automatic processes alone are sufficient for wishful thinking. Though controlled, Type 2 processing inhibits the bias when induced to play a role, it does not typically contribute to the bias, either antagonistically or complementarily, absent such an inducement. The results thus describe the biased updating that leads to wishful thinking as an automatic process rumbling in the cognitive background beyond our awareness and control, but which is liable to be corrected when our attention is turned towards it.

It seems then that there is no role for Type 2 processing in generating or amplifying the bias. Yet we may not want to be too hasty in ruling out the potential for such a role. Though we have trumpeted our paradigm as a paragon of simplicity, we may have been too assiduous in ridding the task of complexity, and inadvertently circumscribed only part of the problem space; one that is more congenial to Type 1 than Type 2 processes.

Yet this is less a flaw of the task than a potential demonstration of the value of its approach. Having mapped out one corner of the problem space we are now free to explore the rest. Numerous paths lie open to us. We could devise a more complex set of associations between states and outcomes so that there are numerous pieces of evidence—as opposed to a mere positive or negative outcome—that a participant has to consider when trying to figure out the task state (see Yang & Shadlen, 2007). We could use evidence more amenable to either rule-based or associative learning. We could pre-train participants on the evidence and create experts. All such ideas constitute viable extensions of our task and demonstrate its utility without compromising its original aims. In short, we have made tangible progress in decomposing wishful thinking in Chapter 3, and have oriented ourselves for the road ahead.

Chapter 4

Optimism about motivated bias

There is a strong, intuitive appeal to the view of motivated bias as the result of a ruminative process, in which one's transgressions are rationalized as peccadillos, setbacks as opportunities, defects as virtues, and failures of self-control as licensed rewards. Upon this view we consciously direct our attention towards the favorable and away from the unfavorable aspects of the situation until we have safely spun the desirable belief in a cocoon of plausibility. This is an occasional, effortful process, employed for matters of personal import, and though it is irrational, it is understandable, once you weigh the comfort of the resultant belief against the vagueness of any potential downstream costs.

However, this picture cannot be reconciled with the results of the eight experiments in this dissertation. In our experiments, a potential wishful thinker had little complexity in which to maneuver. The causal workings of our task were straightforward and clear to our participants. The evidence the participant had to work with was simply related to the desired belief, and the content of the desired belief was defined by little other than its desirability. The cost of inaccuracy was also clear, and there was no extra-experimental benefit to be accrued from wishful thinking. One's self-esteem, expectations for the future, or social standing were not at stake. That we elicited wishful thinking at all in such a paradigm is already difficult to explain upon the ruminative account.

More importantly, the wishful thinking bias we elicited survived the disruption of working memory, implying that the controlled, effortful processes required by a ruminative account are not behind it. What's more, when we did encourage the engagement of controlled cognitive processes—by enforcing a delayed response—we found that the bias attenuated, demonstrating that the role for such processes is to intervene and correct the bias rather than to generate it. Far from an occasional, effortful rationalization that thrives on evidential complexity and uncertain costs, the wishful thinking bias we engendered is a simple, biased, belief updating process that operates automatically and beneath our awareness. That such a process exists is what we have learned from this dissertation (see also Lieberman et al., 2001).

But our hope is that the merit of these studies lies not just in what we have learned here, but also in their heralding of a new approach to motivated bias that will ultimately enable us to more fully articulate the psychological processes behind it, probe their neural basis, and understand their evolutionary payoff, their role in mental health, and their relation to other pervasive and fundamental beliefs.

It is in the precision and flexibility that our approach affords that we invest such hope. Our task is simple, with few and readily identifiable moving parts. There are two hidden states, one of which is associated with a greater frequency of rewarding outcomes, and the other with a greater frequency of punishing outcomes. The task transitions between the states with a certain frequency. There is a reward for the participant if they correctly diagnose the current state, which they can attempt through a reverse inference from recent outcomes. These different components to the task are trivial to vary parametrically, increase in complexity, and present in different formats. This enables us to investigate and define the circumstances under which different processes and capacities constrain and contribute to biased updating, including those processes implicated in motivated reasoning—biased hypothesis generation, truncated evidence search, biased evidence weighting—as well as more general processes and capacities such as self-control, reward sensitivity, episodic simulation, and working memory capacity. Indeed, we have already considered how increasing the complexity of the outcomes associated with each state may lead to a greater requirement for the use of working memory when trying to use those outcomes to diagnose the game state.

As well as allowing for finer assays of the processes behind motivated bias, our task also allows for finer articulation of those processes. Its quantitative nature and repeated trials make it amenable to formal modeling. As a simple illustration, we can consider the adaptation of the ideal observer model of a reversal learning paradigm from Hampton et al. (2006) described in Appendix E. This model, though applicable to our task, lacks any provision for biased updating. In deciding how to introduce such a provision we must become precise about where in the belief updating process we expect the bias to occur. Should we alter the prior likelihood of favored hypotheses as indicated by Dunning et al. (1995) and Kunda (1987)? Should we alter the precision of the likelihood estimates as indicated by Lench, Smallman, Darbor, and Bench (2014)? Should we alter how new evidence is weighted, again as suggested by Kunda (1987)? Devising formal models of biased belief updating forces us to precisely articulate how we think such updating occurs, enables direct comparison between different models of biased updating, potentially allows for a more sensitive measure of individual differences in biased updating, and helps us integrate our investigation with more general models of learning and decision-making.

Our task is also congenial to more subtle behavioral measures that may aid in teasing apart and identifying these processes. Pupillometry is a temporally precise and continuously varying indicator of load on working memory (Laeng, Sirois, & Gredeback, 2012), and eye-tracking is a spatially precise measure of attention (Isaacowitz, 2005; Isaacowitz, Wadlinger, Goren, & Wilson, 2006) which are straightforwardly employed in our paradigm (e.g. by spatially separating the presentation of rewards and punishments).

Indeed, the use of such implicit measures may help us contribute to a debate over the role of self-signaling in self-deception (Mele, 1997; Mijovic-Prelec & Prelec, 2009). This idea build upon self-perception theory (Bem, 1967) by suggesting that we select actions that are typically performed by someone in a desirable environment so that we will then go on to interpret those actions as evidence that we are in a desirable environment. This is an ingenuous but ornate cognitive model and by comparing measures of biased belief that are witnessed by the agent to those that are not, we may be able to assess how much this contributes to biased updating.

Both the articulation of biased belief updating in computational models, and the capability to parametrically modify our task gives us wonderful leverage to investigate its neural basis. By fitting such models to participant behavior we generate specific trial by trial computational signals that can subsequently be used as predictors of neural response (Daw, 2011). We also suspect the task is sufficiently adaptable to facilitate its use with non-human animals. As well as allowing for the application of an even greater range of neuroscientific techniques, the simple question whether animals will actually exhibit biased belief updating is both fascinating and woefully underexplored. Investigating such a question will allow us to explore the phylogenetic distribution of the phenomenon and provide sorely needed empirical clues as to its evolutionary function (Egan et al., 2007; Santos & Rosati, 2015).

Our task may also be worth bringing to the issue of motivated bias and mental health. As described in Chapter 1, there is no shortage of evidence that wishful thinking and associated biases are severely compromised in depression (Strunk et al., 2006; Taylor & Brown, 1988), but it remains an open question how central this compromise is to the characterization of the illness (Sharot, Riccardi, Raio, & Phelps, 2007; Strunk et al., 2006). Is depression at least partly dependent upon the failure of the processes underlying motivated bias, or is the negative outlook associated with depression collateral fallout from the failure of other processes? We can test whether other components of depression, such as anhedonia or negative affect can influence wishful thinking in our task, e.g. by manipulating mood (J. S. Lerner & Keltner, 2001; Phelps, Lempert, & Sokol-Hessner, 2014). We can also attempt to pinpoint the specific components of motivated bias that are compromised in depression, e.g. is depression associated with irrationally pessimistic updating, or pessimistically biased priors (Garrett et al., 2014; Korn et al., 2014)? In this regard, our paradigm aligns with other quantitative approaches to mental health (Dayan, 2014; Mukherjee & Kable, 2014; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008).

We could even attempt to mitigate the severity of the deficit by using our task as a cognitive training tool. Such is the approach taken by attentional bias modification treatments (ABM) of social and generalized anxiety (Amir et al., 2009; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Schmidt, Richey, Buckner, & Timpano, 2009). In the ABM paradigm participants perform a dot probe task modified so that the probe is more associated with stimuli that the participant finds aversive, such as negative facial expressions in the case of social anxiety. Over time, the participant implicitly learns to bias their attention towards these stimuli, which can help alleviate the condition. This suggests that adding a similar hidden contingency between making positive mode judgments in our task and a helpful or desirable component of concurrent supplemental task (such as a dot probe task) might implicitly train participants to think the task is in positive mode. Indeed, recent work has shown that working memory can be retrained to bias itself away from retaining negative information (Robinaugh, Crane, Enock, & McNally, 2015), a result that aligns well with recent theories of working memory control (Hazy et al., 2007).

We have argued that our approach is valuable in emphasizing control even at the potential expense of external validity, and found that wishful thinking occurs even in a semantically sparse and well-controlled environment. This allows us to claim that wishful thinking is due to a general biased belief updating process, rather than a content specific process, which in turn opens up the question of how this general process relates to specific widely held beliefs. Some beliefs, despite their prevalence, seem underdetermined by evidence, such as political belief systems (Jost et al., 2007; Jost, Glaser, Kruglanski, & Sulloway, 2003), the belief that the world is just (Dalbert, 2009; M. J. Lerner & Miller, 1978), and the belief that it is governed by a higher power (Kay, Gaucher, Napier, Callan, & Laurin, 2008). Interestingly, the belief in a just world has similar signatures in terms of life outcomes and motivation, to more obviously biased beliefs such as optimism (Dalbert, 2009). Yet it is still not clear that a general bias in belief updating is behind the generation and maintenance of these beliefs, or whether they are the result of—for example—an innate disposition towards them, or represent acts of faith made in full awareness of potential counter-evidence.

Motivated bias is not a problem that sits cordoned off from the rest of psychology. As we hope this dissertation has shown, it benefits from the application of concepts and techniques from more general theories and investigations of learning, thinking, and decision-making. But it also serves as a challenge to and testing ground for such theories, and its investigation ultimately illuminates the architecture of the mind and the forces determining our evolution. In all of the above, we do not mean to advocate specifically for our paradigm as a silver bullet in the hunt for solutions to the conundrums of motivated bias. Rather we wish to advocate more generally for the use of controlled, quantifiable, models of those conundrums—of which we feel our paradigm is a clear exemplar—as a fresh approach to their investigation that complements the work done to date. Motivated bias is a deep problem, but we are optimistic about its resolution.

References

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The Self in Social Judgment*, *1*, 85–106.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*(4), 441–485. http://doi.org/10.1037/0096-3445.108.4.441

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., & Chen, X. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*(5), 961–973. http://doi.org/10.1037/a0016685

Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, *103*(4), 718–735. http://doi.org/10.1037/a0029395

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*, *40*(7), 817–826. http://doi.org/10.1016/S0028-3932(01)00178-6

Aronson, E. (1997). Back to the Future: Retrospective Review of Leon Festinger's "A Theory of Cognitive Dissonance." *The American Journal of Psychology*, *110*(1), 127–137. http://doi.org/10.2307/1423706

Aronson, E., & Carlsmith, J. M. (1963). Effect of the severity of threat on the devaluation of forbidden behavior. *The Journal of Abnormal and Social Psychology*, *66*(6), 584–588. http://doi.org/10.1037/h0039901

Aue, T., Nusbaum, H. C., & Cacioppo, J. T. (2011). Neural correlates of wishful thinking. *Social Cognitive and Affective Neuroscience*. http://doi.org/10.1093/scan/nsr081

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Babad, E. (1997). Wishful Thinking Among Voters: Motivational and Cognitive Influences. *International Journal of Public Opinion Research*, *9*(2), 105–125. http://doi.org/10.1093/ijpor/9.2.105

Baker, L. A., & Emery, R. E. (1993). When every relationship is above average: Perceptions and expectations of divorce at the time of marriage. *Law and Human Behavior*, *17*(4), 439–450. http://doi.org/10.1007/BF01044377

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., & Dai, B. (2014). lme4: Linear mixed-effects models using Eigen and S4 (Version 1.1-7). Retrieved from http://cran.us.r-project.org/web/packages/lme4/index.html

Batson, C. D. (2007). Moral masquerades: Experimental exploration of the nature of moral motivation. *Phenomenology and the Cognitive Sciences*, *7*(1), 51–66. http://doi.org/10.1007/s11097-007-9058-y

Beer, J. S., & Hughes, B. L. (2010). Neural systems of social comparison and the "above-average" effect. *NeuroImage*, *49*(3), 2671–2679. http://doi.org/10.1016/j.neuroimage.2009.10.075

Bem, D. J. (1967). Self-perception: an alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*(3), 183–200. http://doi.org/10.1037/h0024835

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, *67*(3), 366–381. http://doi.org/10.1037/0022-3514.67.3.366

Carson, R. C., Hollon, S. D., & Shelton, R. C. (2010). Depressive realism and clinical depression. *Behaviour Research and Therapy*, *48*(4), 257–265. http://doi.org/10.1016/j.brat.2009.11.011

Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review*, *30*(7), 879–889. http://doi.org/10.1016/j.cpr.2010.01.006

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. Guilford Press.

Chance, Z., Norton, M. I., Gino, F., & Ariely, D. (2011). Colloquium Paper: Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, *108*(Supplement_3), 15655–15659. http://doi.org/10.1073/pnas.1010658108

Chowdhury, R., Sharot, T., Wolfe, T., Düzel, E., & Dolan, R. J. (2014). Optimistic update bias increases in older age. *Psychological Medicine*, *44*(09), 2003–2012. http://doi.org/10.1017/S0033291713002602

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making under Risk. *The Journal of Neuroscience*, *29*(40), 12574–12583. http://doi.org/10.1523/JNEUROSCI.2614-09.2009

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *The Journal of Neuroscience*, *22*(11), 4563–4567.

Dalbert, C. (2009). Belief in a just world. *Handbook of Individual Differences in Social Behavior*, 288–297.

Dana, J., Weber, R., & Kuang, J. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. http://doi.org/10.1007/s00199-006-0153-z

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision Making, Affect, and Learning: Attention and Performance XXIII* (Vol. 23, pp. 3 – 38). Oxford University Press.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated Reasoning and Performance on the was on Selection Task. *Personality and Social Psychology Bulletin*, *28*(10), 1379–1387. http://doi.org/10.1177/014616702236869

Dayan, P. (2014). Rationalizable Irrationalities of Choice. *Topics in Cognitive Science*, *6*(2), 204–228. http://doi.org/10.1111/tops.12082

Dickinson, A. (1980). *Contemporary animal learning theory* (Vol. 1). CUP Archive.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*(4), 568.

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Chapter 10 Motivated Moral Reasoning. In *Psychology of Learning and Motivation* (Vol. 50, pp. 307–338). Elsevier.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology*, *75*(1), 53–69. http://doi.org/10.1037/0022-3514.75.1.53

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology*, *69*(1), 58.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082–1090. http://doi.org/10.1037/0022-3514.57.6.1082

Egan, L. C., Santos, L. R., & Bloom, P. (2007). The Origins of Cognitive Dissonance Evidence From Children and Monkeys. *Psychological Science*, *18*(11), 978–983. http://doi.org/10.1111/j.1467-9280.2007.02012.x

Eil, D., & Rao, J. M. (2011). The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–138. http://doi.org/10.1257/mic.3.2.114

Eldar, E., & Niv, Y. (2015). Interaction between emotional state and learning underlies mood instability. *Nature Communications*, *6*. http://doi.org/10.1038/ncomms7149

Epley, N., & Whitchurch, E. (2008). Mirror, Mirror on the Wall: Enhancement in Self-Recognition. *Personality and Social Psychology Bulletin*, *34*(9), 1159–1170. http://doi.org/10.1177/0146167208318601

Evans, J. S. B. T. (1984). Heuristic and analytic processes in reasoning*. *British Journal of Psychology*, *75*(4), 451–468. http://doi.org/10.1111/j.2044-8295.1984.tb01915.x

Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278. http://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. http://doi.org/10.1177/1745691612460685

Faraway, J. J. (2005). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. CRC press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Garrett, N., & Sharot, T. (2014). How robust is the optimistic update bias for estimating self-risk and population base rates? *PloS One*, *9*(6), e98848.

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression. *Frontiers in Human Neuroscience*, *8*. http://doi.org/10.3389/fnhum.2014.00639

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*(2), 107–119. http://doi.org/10.1037/0003-066X.46.2.107

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, *60*(4), 509–517. http://doi.org/10.1037/0022-3514.60.4.509

Gramzow, R. H., Willard, G., & Mendes, W. B. (2008). Big tales and cool heads: Academic exaggeration is related to cardiac vagal reactivity. *Emotion*, *8*(1), 138–144. http://doi.org/10.1037/1528-3542.8.1.138

Greenwald, A. G. (2012). There Is Nothing So Theoretical as a Good Method. *Perspectives on Psychological Science*, *7*(2), 99–108. http://doi.org/10.1177/1745691611434210

Guenther, C. L., & Alicke, M. D. (2008). Self-enhancement and belief perseverance. *Journal of Experimental Social Psychology*, *44*(3), 706–712. http://doi.org/10.1016/j.jesp.2007.04.010

Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS–US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*(1), 31–42. http://doi.org/10.1037/0097-7403.5.1.31

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The Role of the Ventromedial Prefrontal Cortex in Abstract State-Based Inference during Decision Making in Humans. *The Journal of Neuroscience*, *26*(32), 8360–8367. http://doi.org/10.1523/JNEUROSCI.1010-06.2006

Harding, E. J., Paul, E. S., & Mendl, M. (2004). Animal behaviour: Cognitive bias and affective state. *Nature*, *427*(6972), 312–312. http://doi.org/10.1038/427312a

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.

Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2007). Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1601–1613. http://doi.org/10.1098/rstb.2007.2055

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related Reversal Learning after Surgical Excisions in Orbito-frontal or

Dorsolateral Prefrontal Cortex in Humans. *Journal of Cognitive Neuroscience*, *16*(3), 463–478. http://doi.org/10.1162/089892904322926791

Isaacowitz, D. M. (2005). The Gaze of the Optimist. *Personality and Social Psychology Bulletin*, *31*(3), 407–415. http://doi.org/10.1177/0146167204271599

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Selective preference in visual fixation away from negative images in old age? An eye-tracking study. *Psychology and Aging*, *21*(1), 40–48. http://doi.org/10.1037/0882-7974.21.1.40

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. http://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. http://doi.org/10.1038/nature10384

Johnson, D. D. P., & Tierney, D. (2011). The Rubicon Theory of War: How the Path to Conflict Reaches the Point of No Return. *International Security*, *36*(1), 7–40. http://doi.org/10.1162/ISEC_a_00043

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. http://doi.org/10.1037/0033-2909.129.3.339

Jost, J. T., Napier, J. L., Thorisdottir, H., Gosling, S. D., Palfai, T. P., & Ostafin, B. (2007). Are Needs to Manage Uncertainty and Threat Associated With Political Conservatism or Ideological Extremity? *Personality and Social Psychology Bulletin*, *33*(7), 989–1007. http://doi.org/10.1177/0146167207301028

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2013). Motivated numeracy and enlightened self-government. *Yale Law School, Public Law Working Paper*, (307). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2319992

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. http://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341–350. http://doi.org/10.1037/0003-066X.39.4.341

Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology*, *95*(1), 18–35. http://doi.org/10.1037/0022-3514.95.1.18

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(03), 579–592. http://doi.org/10.1017/S0033291713001074

Krizan, Z., & Windschitl, P. D. (2009). Wishful Thinking about the Future: Does Desire Impact Optimism? *Social and Personality Psychology Compass*, *3*(3), 227–243. http://doi.org/10.1111/j.1751-9004.2009.00169.x

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*(4), 636.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. http://doi.org/10.1037/0033-2909.108.3.480

Kurzban, R., & Aktipis, C. A. (2007). Modularity and the Social Mind: Are Psychologists Too Self-ish? *Personality and Social Psychology Review*, *11*(2), 131–149. http://doi.org/10.1177/1088868306294906

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2014). lmerTest: Tests in Linear Mixed Effects Models (Version 2.0-20). Retrieved from http://cran.us.r-project.org/web/packages/lmerTest/index.html

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. http://doi.org/10.1177/1745691611427305

Lench, H. C., Smallman, R., Darbor, K. E., & Bench, S. W. (2014). Motivated perception of probabilistic information. *Cognition*, *133*(2), 429–442. http://doi.org/10.1016/j.cognition.2014.08.001

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159. http://doi.org/10.1037/0022-3514.81.1.146

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, *85*(5), 1030–1051. http://doi.org/10.1037/0033-2909.85.5.1030

Lieberman, M. D. (2007). Social Cognitive Neuroscience: A Review of Core Processes. *Annual Review of Psychology*, *58*(1), 259–289. http://doi.org/10.1146/annurev.psych.58.110405.085654

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do Amnesics Exhibit Cognitive Dissonance Reduction? The Role of Explicit Memory and Attention in

Attitude Change. *Psychological Science*, *12*(2), 135–140. http://doi.org/10.1111/1467-9280.00323

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*(4), 395–410. http://doi.org/10.3758/BF03200866

Lovallo, D., & Kahneman, D. (2003). Delusions of Success. *Harvard Business Review*, *81*(7), 56–63.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology*, *111*(1), 107–123. http://doi.org/10.1037/0021-843X.111.1.107

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com.ezp-prod1.hul.harvard.edu/sol3/papers.cfm?abstract_id=979648

McKay, R. T., & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, *32*(06), 493–510. http://doi.org/10.1017/S0140525X09990975

Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*(01), 91–102. http://doi.org/null

Mijovic-Prelec, D., & Prelec, D. (2009). Self-deception as self-signalling: a model and experimental evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1538), 227–240. http://doi.org/10.1098/rstb.2009.0218

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502–517. http://doi.org/10.1037/0033-295X.115.2.502

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends in Cognitive Sciences*, *14*(10), 435–440. http://doi.org/10.1016/j.tics.2010.07.004

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, *4*(2), 61–64.

Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S.-J., & Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences*, *110*(41), 16396–16401. http://doi.org/10.1073/pnas.1305631110

Mueller, S. T., & Zhang, J. (2006). A Non-Parametric ROC-Based Measure of Sensitivity. In *Third workshop on ROC analysis in Machine Learning, Pittsburgh, USA*. Retrieved from http://obereed.net/docs/MuellerZhangICML2006.pdf

Mukherjee, D., & Kable, J. W. (2014). Value-Based Decision Making in Mental Illness A Meta-Analysis. *Clinical Psychological Science*, 2167702614531580. http://doi.org/10.1177/2167702614531580

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural Prediction Errors Reveal a Risk-Sensitive Reinforcement-Learning Process in the Human Brain. *Journal of Neuroscience*, *32*(2), 551–562. http://doi.org/10.1523/JNEUROSCI.5498-10.2012

Norton, M. I., Vandello, J. A., & Darley, J. M. (2004). Casuistry and Social Category Bias. *Journal of Personality and Social Psychology*, *87*(6), 817–831. http://doi.org/10.1037/0022-3514.87.6.817

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and Reasoning in Moral Judgment. *Cognitive Science*, *36*(1), 163–177. http://doi.org/10.1111/j.1551-6709.2011.01210.x

Peterson, J. B., DeYoung, C. G., Driver-Linn, E., Séguin, J. R., Higgins, D. M., Arseneault, L., & Tremblay, R. E. (2003). Self-deception and failure to modulate responses despite accruing evidence of error. *Journal of Research in Personality*, *37*(3), 205–223. http://doi.org/10.1016/S0092-6566(02)00569-X

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, *37*(1), 263–287. http://doi.org/10.1146/annurev-neuro-071013-014119

Pinker, S. (2011). Representations and decision rules in the theory of self-deception. *Behavioral and Brain Sciences*, *34*(01), 35–37.

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76–87. http://doi.org/10.1016/j.jpsychires.2008.03.001

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71–99. http://doi.org/10.1016/j.jfineco.2006.09.003

Quattrone, G. A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and on the voter's illusion. *Journal of Personality and Social Psychology*, *46*(2), 237–248. http://doi.org/10.1037/0022-3514.46.2.237

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489*(7416), 427–430. http://doi.org/10.1038/nature11467

R Development Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from http://www.r-project.org/

Robinaugh, D. J., Crane, M. E., Enock, P. M., & McNally, R. J. (2015). Training the removal of negative information from working memory: A preliminary investigation of a working memory bias modification task. *Cognition and Emotion*, *0*(0), 1–12. http://doi.org/10.1080/02699931.2015.1014312

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology*, *59*(2), 229–241. http://doi.org/10.1037/0022-3514.59.2.229

Santos, L. R., & Rosati, A. G. (2015). The Evolutionary Roots of Human Decision Making. *Annual Review of Psychology*, *66*(1), 321–347. http://doi.org/10.1146/annurev-psych-010814-015310

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063.

Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology*, *118*(1), 5–14. http://doi.org/10.1037/a0013643

Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, *84*(1), 60–79. http://doi.org/10.1037/0022-3514.84.1.60

Sharot, T., Guitart-Masip, M., Korn, C. W., Chowdhury, R., & Dolan, R. J. (2012). How Dopamine Enhances an Optimism Bias in Humans. *Current Biology*, *22*(16), 1477–1481. http://doi.org/10.1016/j.cub.2012.05.053

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*. http://doi.org/10.1073/pnas.1205828109

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. http://doi.org/10.1038/nn.2949

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature*, *450*(7166), 102–105. http://doi.org/10.1038/nature06280

Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, *141*(3), 423–428. http://doi.org/10.1037/a0025391

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology*, *38*, 183.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. http://doi.org/10.1037/0033-295X.84.2.127

Skatova, A., Chan, P. A., & Daw, N. D. (2013). Extraversion differentiates between model-based and model-free strategies in a reinforcement learning task. *Frontiers in Human Neuroscience*, *7*. http://doi.org/10.3389/fnhum.2013.00525

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22. http://doi.org/10.1037/0033-2909.119.1.3

Sloman, S. A., Fernbach, P. M., & Hagmayer, Y. (2010). Self-deception requires vagueness. *Cognition*, *115*(2), 268–281. http://doi.org/10.1016/j.cognition.2009.12.017

Stanovich, K. E., & Toplak, M. E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind & Society*, *11*(1), 3–13. http://doi.org/10.1007/s11299-011-0093-6

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*(2), 161–188. http://doi.org/10.1037/0096-3445.127.2.161

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*(05), 701–717.

Strunk, D. R., Lopez, H., & DeRubeis, R. J. (2006). Depressive symptoms are associated with unrealistic negative predictions of future life events. *Behaviour Research and Therapy*, *44*(6), 861–882. http://doi.org/10.1016/j.brat.2005.07.001

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, *47*(2), 143–148. http://doi.org/10.1016/0001-6918(81)90005-6

Szu-Ting Fu, T., Koutstaal, W., Poon, L., & Cleare, A. J. (2012). Confidence judgment in depression and dysphoria: The depressive realism vs. negativity hypotheses. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(2), 699–704. http://doi.org/10.1016/j.jbtep.2011.09.014

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193–210. http://doi.org/10.1037/0033-2909.103.2.193

Taylor, S. E., Kemeny, M. E., Reed, G. M., Bower, J. E., & Gruenewald, T. L. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, *55*(1), 99–109. http://doi.org/10.1037/0003-066X.55.1.99

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003a). Are self-enhancing cognitions associated with healthy or unhealthy biological profiles? *Journal of*

Personality and Social Psychology, *85*(4), 605–615. http://doi.org/10.1037/0022-3514.85.4.605

Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003b). Portrait of the self-enhancer: Well adjusted and well liked or maladjusted and friendless? *Journal of Personality and Social Psychology*, *84*(1), 165–176. http://doi.org/10.1037/0022-3514.84.1.165

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*(3), 107–140. http://doi.org/10.1016/j.cogpsych.2011.06.001

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, *20*(4), 410–433. http://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Trivers, R. (2000). The Elements of a Scientific Theory of Self-Deception. *Annals of the New York Academy of Sciences*, *907*(1), 114–131. http://doi.org/10.1111/j.1749-6632.2000.tb06619.x

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed Criteria Redefining Merit to Justify Discrimination. *Psychological Science*, *16*(6), 474–480. http://doi.org/10.1111/j.0956-7976.2005.01559.x

Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology*, *44*(5), 1334–1338. http://doi.org/10.1016/j.jesp.2008.03.010

Von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*(01), 1–16. http://doi.org/10.1017/S0140525X10001354

Von Hippel, W., & Trivers, R. (2013). Self-deception to deceive others. In J. P. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social thinking and interpersonal behavior* (pp. 65 – 80). New York: Psychology Press.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806–820. http://doi.org/10.1037/0022-3514.39.5.806

Williams, E. F., & Gilovich, T. (2008). Do people really believe they are above average? *Journal of Experimental Social Psychology*, *44*(4), 1121–1128. http://doi.org/10.1016/j.jesp.2008.01.002

Windschitl, P. D., Scherer, A. M., Smith, A. R., & Rose, J. P. (2013). Why so confident? The influence of outcome desirability on selective exposure and likelihood judgment. *Organizational Behavior and Human Decision Processes*, *120*(1), 73–86. http://doi.org/10.1016/j.obhdp.2012.10.002

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv Preprint arXiv:1308.5499*. Retrieved from http://arxiv.org/pdf/1308.5499

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, *447*(7148), 1075–1080. http://doi.org/10.1038/nature05852

Appendix A

Instruction wording for all experiments

Experiment 1

Wording is specifically for when the risk preference task came first.

Screen 1

Welcome to The Casino Game! This task is like playing the slot machines at a casino, only better. In a casino the outcome is entirely due to luck and the odds are stacked against you. In this task, though there is some luck involved, it is within your power to greatly increase your chance of earning a large bonus.

We are going to ask you to play two games and take a survey at the end. In all, the task should take about 30 minutes to complete. In order to help you make the right choices, we will explain how each of our slot machines work before allowing you to play each game.

Please read the instructions carefully. It is important to us that you understand everything. The more you know about how the machines work, the more money you will earn today.

Screen 2

In each game you will be playing with a machine that has two levers. Each lever will either give you tokens, or take tokens away from you. On each lever, you will see the amount of tokens at stake if you choose that lever. For every 100 tokens you collect, you will earn 1 cent. To choose a lever, you must press the left or right arrow key.

Screen 3

Here is a picture of each Lever. [picture of each lever appeared at this point]. As you can see, one Lever will either give you 150 tokens or take 150 tokens away from you, and the other will either give you 120 tokens or take 120 tokens away from you. We call the 150 tokens Lever the 'High-Range' Lever, and the 120 tokens Lever the 'Low-Range Lever'. You can win more tokens on the High-Range Lever, but you can also lose more tokens on the High-Range Lever.

Screen 4

In order to ensure you have understood the instructions, we will occasionally ask you some simple comprehension questions. The first question is below.

Q: If I choose the 'High-Range' Lever the two possible outcomes are:

I will either receive 150 tokens or lose 150 tokens

I will either receive 150 tokens or lose 120 tokens

I will either receive 120 tokens or lose 150 tokens

I will either receive 120 tokens or lose 120 tokens

Screen 5

In this first game, you will also see a message telling you how likely it is that you will win or lose tokens on each turn. You will see this message between the two levers. A picture of what that looks like is below.

As you can see the message tells you how likely you are to WIN tokens. So a low percentage means that you are likely to LOSE tokens.

If you think about it, you will see that it makes sense to choose the High-Range Lever when you think you are likely to win, so that you maximize your winnings, and to choose the Low-Range Lever when you think you are likely to lose, so that you minimize your losses.

However, most of the time it is never guaranteed that you will win or lose, so most choices are a risk. Choose whatever you feel comfortable with.

Screen 6

The second game is different than the first. In the last game, how likely it was that the machine would give you tokens was different on nearly every go. This time, the machine will only ever be in one of two different 'modes': Positive Mode and Negative Mode.

In Positive Mode, the machine gives you tokens most of the time (~66% of the time), but it will occasionally take tokens away from you (~33% of the time). In Negative Mode, the machine takes tokens away from you most of the time (again ~66% of the time), but it will occasionally give you tokens instead (~33% of the time).

Obviously, it is better to be in Positive Mode!

Screen 7

Please answer the following question:

Q: When the machine is in Positive Mode it:

Gives you tokens all the time

Gives you tokens most of the time

Takes tokens back most of the time

Takes tokens back all the time

Screen 8

Please answer the following question.

Q: A message with a low percentage means:

I have a high chance of winning tokens and a low chance of losing tokens on this turn

I have a low chance of winning tokens and a high chance of losing tokens on this turn

Screen 9

If you think about it, you will see that the best option is different in each mode. In Positive Mode, since you are mostly gaining tokens, you should choose the lever that gives you the most tokens, i.e the High-Range Lever (which will mostly give you 150 tokens). But when the machine is in Negative Mode, since you are mostly losing tokens, you should choose the lever that takes the least amount of tokens away from you, i.e. the Low-Range Lever (which will mostly only take away 120 tokens).

If you successfully follow this strategy, you will greatly increase your chances of earning a large bonus payment.

Screen 10

Please answer the following question:

Q: When the machine is in Positive Mode, I should:

Choose the High-Range Lever

Choose the Low-Range Lever

Screen 11

When the machine is in Negative Mode, I should:

Choose the High-Range Lever

Choose the Low-Range Lever

Screen 12

Before you play this game, we think it would be helpful to try a practice session.

In this practice session, sometimes the machine will be in the Positive Mode, and sometimes it will be in the Negative Mode. As you will see, the switch between modes doesn't happen on every turn. It only happens every once in a while. It will tell you on screen which mode the machine is in, so you will know when it has changed.

You will have 30 'trials' on the machine. Please note that any tokens you earn in this practice session will not count towards your final tally.

Screen 13

As you have seen, the machine switches modes every once in a while. It is important to realize that these switches occur entirely at random. That is to say, they are not programmed ahead of time and nothing you do has any influence on whether the machine switches or not.

* Given that the switches are random, sometimes there may be a long time between switches, and sometimes they might occur in quick succession. You may have noticed this in the practice session.

Screen 14

Please answer the following question:

Q: The machine switches mode:

Only when I make the same choice 5 times in a row

According to a pre-programmed schedule

At random, every once in a while

Screen 15

There is one big difference between the practice task you just completed and the game you are about to play. In the game, we will not tell you which mode the machine is in! There will be no message on the screen! What should you do?

Don't worry. Though this makes it a bit more challenging, you can still tell what mode the machine is in by paying attention to what the machine does. If it is mostly giving you tokens, then it must be in Positive Mode. If it is mostly taking tokens back from you, it must be in Negative Mode. So, although you won't see a message telling you what mode the machine is in, you can still figure it out by tracking whether the machine is mostly giving you tokens or mostly taking them away.

Screen 16

Please answer the following question:

Q: To figure out what mode the machine is in, I should:

Stare intently at the screen

Pay attention to what the machine is doing. If it is mostly giving me tokens, then it must be in Positive Mode. If it is mostly taking tokens away, it must be in Negative Mode

Guess

Screen 17

Thank you for your patience with the instructions. We have listed a reminder of the main points at the bottom of the screen in case it is helpful.

When you click to the next screen, you will be taken to the second game. You start with 10000 tokens. You will have 150 'trials' on the machine to win as many tokens as possible. The more tokens you get, the greater your bonus.

Good luck!

A reminder of the main points:

Important Point #1 In Positive Mode the machine mostly gives you tokens. In Negative Mode, the machine mostly takes tokens away.

Important Point #2 In Positive Mode, you should choose the lever that gives you the most tokens, i.e. the High-Range Lever. When the machine is in Negative Mode, you should choose the lever that takes the least amount of tokens away, i.e. the Low-Range Lever

Important Point #3 The machines switches modes every once in a while. It decides to do so entirely at random.

Important Point #4 We will not tell you which mode the machine is in. You can still tell what mode the machine is in by paying attention to what the machine does. If it is mostly giving you tokens, then it must be in Positive Mode. If it is mostly taking tokens away, it must be in Negative Mode.

Experiment 2

Screen 1

Welcome to The Casino Game!

This task is like playing the slot machines at a casino, only better. In a casino the outcome is entirely due to luck and the odds are stacked against you. In this task, as well as asking you to play on a slot machine, we will ask you to figure out what the machine is doing, and we will give you an extra bonus for doing so. So at least some of your bonus will be due to your own skill.

In order to help you do well at this task, we will explain how the slot machine works before allowing you to play the game.

Please read the instructions carefully. It is important to us that you understand everything. The more you know about how the machine works, the more money you will earn today.

Screen 2

Instructions: The Basics

You will be playing with a simple slot machine that will either give you tokens, or take tokens away from you. On the machine, you will see the amount of tokens at stake for every pull of the lever. In this case the machine will either give you 120 tokens or take 120 tokens away from you. For every 100 tokens you collect, you will earn 1 cent. To activate the machine, you must press the down arrow key.

Screen 3

Instructions: A Quick Practice Session

Let's try a very brief practice session to get familiar with the slot machine. You will have 5 'trials' on the machine. Remember to press the down arrow key to activate the machine. Please note that any tokens you earn in this practice session will not count towards your final tally.

Screen 4

Instructions: Two Modes

The most important thing you need to know is that the machine can be in either of two different 'modes': Positive Mode and Negative Mode. In Positive Mode, the machine gives you tokens most of the time (~66% of the time), but it will occasionally take tokens away from you (~33% of the time) **[Note: participants were not informed of the exact contingencies for Experiments 3, 4a, 4b, 5, 6a, and 6b]**. In Negative Mode, the machine takes tokens away from you most of the time (again ~66% of the time), but it will occasionally give you tokens instead (~33% of the time).

This means that the more time the machine is in Positive Mode, the more tokens you should expect to get, and the greater your bonus at the end. Obviously, it is better to be in Positive Mode!

Screen 5

In order to ensure you have understood the instructions, we will occasionally ask you some simple comprehension questions. The first question is below.

Q: When the machine is in Positive Mode it:

Gives you tokens all the time

Gives you tokens most of the time

Takes tokens back most of the time

Takes tokens back all of the time

Screen 6

Instructions: The Judgment Task

Just like in a casino, how well you do on the slot machine is a matter of luck. But we are also giving you the chance to use your own judgment to earn some extra tokens. After every pull of the lever, once you have seen the outcome, we will ask you to make your best guess as to what 'Mode' the machine is in: Positive Mode or Negative Mode. You will get a bonus of 30 tokens for every correct answer. You choose 'Positive Mode' by pressing the left arrow, and 'Negative Mode' by pressing the right arrow.

Screen 7

Instructions: Practice Session 2

Let's try another practice session so you can get familiar with the task. In this practice session, sometimes the machine will be in the Positive Mode, and sometimes it will be in the Negative Mode. As you will see, the switch between modes doesn't happen on every turn. It only happens every once in a while. Just for this practice session, it will tell you on screen which mode the machine is in, so you will know when it has changed. This message will not be there during the actual game. You will have 20 'trials' on the machine so that you can get a good sense of how the machine behaves. Again, please note that any tokens you earn in this practice session will not count towards your final tally.

Screen 8

Instructions: Switching Modes

Well done. As you have seen, the machine switches modes every once in a while. It is important to realize that these switches occur entirely at random. That is to say, they are not programmed ahead of time and nothing you do has any influence on whether the machine switches or not. For instance, if you choose 'Positive Mode', it will NOT make it any more or less likely that the machine will be in 'Positive Mode' for the next turn. And if you choose 'Negative Mode', it will NOT make it any more or less likely that the machine will be in 'Negative Mode' for the next turn.

* Given that the switches are random, sometimes there may be a long time between switches, and sometimes they might occur in quick succession. You may have noticed this in the practice session.

Screen 9

Please answer the following question:

Q: The machine switches mode:

Only when I make the same choice five times in a row

According to a pre-programmed schedule

At random, every once in a while

Whenever I choose Positive Mode

Screen 10

Instructions: How to do well

As we mentioned, unlike in the practice session, there will be no message on screen telling you what mode the machine is in. So how can you figure it out? If you think about it, you will notice that you can tell what mode the machine is in by paying attention to what the machine does. If it is mostly giving you tokens, then it must be in Positive Mode. If it is mostly taking tokens back from you, it must be in Negative Mode. So, although you won't see a message telling you what mode the machine is in, you can still figure it out by tracking whether the machine is mostly giving you tokens or mostly taking them away.

Screen 11

Please answer the following question:

Q: To figure out what mode the machine is in, I should:

Stare intently at the screen

Pay attention to what the machine is doing. If it is mostly giving me tokens, then it must be in Positive Mode. If it is mostly taking tokens away, it must be in Negative Mode.

Guess

Screen 12

Instructions: A Final Important Point

Thank you for your patience with the instructions. There is just one more final point that we want to emphasize before you start the task. As you may have noticed during the practice session, the machine doesn't always give you tokens in Positive Mode, and it doesn't always take tokens away from you in Negative Mode. This means that even if the machine has just given you tokens, that doesn't necessarily mean that it is in Positive Mode, and if it has just taken tokens away from you, that doesn't necessarily mean that it is in Negative Mode. So, when trying to decide what mode the machine is in, it is important that you think about whether or not the machine has been giving you tokens over the last number of turns, rather than just thinking about the most recent outcome.

Screen 13

Please answer the following question:

Q: When deciding what mode the machine is in, I should:

Think about whether or not the machine has been giving me tokens over the last number of turns.

Think only about what happened on the most recent turn.

Screen 14

When you click to the next screen, you will be taken to the main task. We have listed a reminder of the main points from the instructions below in case it is helpful.

You start with 10000 tokens. You will have 150 ‘trials’ on the machine to win as many tokens as possible. The more tokens you get, the greater your bonus. Remember, even though you won't see it until the end, you can earn up to 4500 extra tokens depending on your judgments!

Good luck!

Experiment 3: Memory Task Instructions

Screen 1

Instructions: The Memory Task

Thank you for your patience with the instructions so far.

As we mentioned earlier, to make things a little more difficult for you we are also going to give you a memory task on top of the judgment task!

Every so often we will pause the game and ask you to remember what happened a certain number of turns ago. Did you get a "+120 tokens" or a "-120 tokens" on that turn?

For instance, we might ask you what happened 2 turns ago. In that case you would need to cast your mind back beyond the most recent outcome, to the outcome from the turn before that, and tell us whether the machine gave you +120 tokens or -120 tokens on that turn.

For every outcome you remember correctly you will receive 100 extra bonus tokens!

Screen 2

Instructions: Practice Session 3

Let's try a final practice session so you can get familiar with the memory task. Just like with the judgment task, you choose your answer for the memory task ('+120 tokens' or '-120 tokens') using the arrow keys. Like before, since it is a practice session, we will tell you on screen which mode the machine is in. This message will not be there during the actual game. Finally, just a reminder that any tokens you earn in this practice session will not count towards your final tally.

Screen 3

Thank you for completing the training session.

Just to be certain things are clear, please answer the following question:

Q: Imagine that on your last three turns, you received +120 tokens, then -120 tokens, then +120 tokens. Now we ask you to recall the outcome from two turns ago. The correct answer is:

+120 tokens

-120 tokens

Experiment 5: Load Task instructions

The following assumes the load task came first.

Screen 1

Instructions: The Memory Task

Thank you for your patience with the instructions so far.

As we mentioned earlier, to make things a little more difficult for you we are also going to give you a memory task on top of the judgment task!

Essentially, we are going to ask you to hold a letter in mind while you are playing the game. Every so often, we will ask you what that letter is. For every letter you remember correctly you will receive 100 extra bonus tokens.

Click below to learn about this in more detail.

Screen 2

The Memory Task: Three Steps

The memory task works in three steps.

Step 1: First we will present you with the letter ‘A’. You will see it appear on the screen where the cog wheel usually appears. You start by keeping that letter in mind.

When you see it appear, you must press the letter ‘A’ on the keyboard in order to continue playing the game.

Note that we will always start with the letter ‘A’.

Screen 3

Step 2: Every time you pull the lever on the slot machine, you should change the letter you are keeping in memory to the next letter in the alphabet.

For instance, if ‘A’ was the letter you were currently trying to remember, then you should update it to ‘B’. If ‘C’ was the letter you were trying to remember, you should change it to ‘D’ and so on.

Screen 4

Step 3: This will continue for a number of turns. Eventually we will ask you what letter you are thinking of. You will see '"which letter?" appear where the cog wheel usually appears. When that happens, you should press a letter key on the keyboard to indicate your answer.

If you have correctly kept track of the letter on every turn, you will get 100 tokens. If you are one letter away from the correct answer, we will still give you 50 tokens.

Once you have given your answer, it will reset to the letter 'A' again.

Screen 5

Instructions: Practice Session 3

Let’s try a final practice session so you can get familiar with the memory task.

Remember to press the ‘A’ key when you see the letter ‘A’. Also remember to give your answer when you see “which letter?” by pressing the appropriate key. Like before, since it is a practice session, we will tell you on screen which mode the machine is in. This message will not be there during the actual game.

Finally, just a reminder that any tokens you earn in this practice session will not count towards your final tally.

Screen 6

Thank you for completing the training session.

Just to be certain things are clear, please answer the following question:

Q: I should update the letter to the next letter in the alphabet whenever:

I receive a reward from the machine

I 'pull the lever'

you tell me to

Appendix B

Demographic and task engagement items

How well do you feel you understood the instructions? *Not at all (1) --- Perfectly (7)*

How much did you enjoy the games? I found it… *Very boring (1) --- Lots of fun (7)*

How hard did you try to earn the largest bonus payment you could? *I didn't try at all (1) --- I did my best (7)*

Did you employ any particular strategies in either of the game? If so, please tell us about them here. *[open ended response]*

What did you think this study was trying to test? *[open ended response]*

Do you have any advice for us about how to improve the games so that future players feel more engaged? *[open ended response]*

Feel free to note anything else that you would like to bring to our attention here. It will be much appreciated. *[open ended response]*

What is your age? *[open ended response]*

What is your gender? *Female; Male*

What is your ethnicity? *[open ended response]*

Please indicate the highest education level of you and your parents or guardian(s).

* If you do not / did not live with either or both of your parents, please complete information for your legal guardian(s) * *No formal education; Less than 7th grade; Junior High / Middle School (9th grade); High School / GED; Some college; 2 year college degree; 4 year college degree; Some graduate or professional school; Masters degree; Professional degree (JD/MD); Doctoral degree*

What would you say was the relative immediate income of your family during your childhood? *Low (1) --- Average (5) --- High (9)*

What is your annual household income? *$15,000 or less; $15001 - $25000; $25001 - $35000; $35001 - $50000; $50001 - $65000; $65001 - $85000; $80001 - $100000; over $100000*

Where did you spend most of your childhood? *[list of countries]*

What is your native language? *English; Arabic; Bengali; Chinese; French; German; Hindi; Japanese; Korean; Malay / Indonesian; Portuguese; Russian; Spanish; Urdu; Other*

When it comes to social issues, how liberal or conservative are you? *Extremely Liberal (1) --- Moderate, Middle of Road (4) --- Extremely Conservative (7)*

When it comes to fiscal issues, how liberal or conservative are you? *Extremely Liberal (1) --- Moderate, Middle of Road (4) --- Extremely Conservative (7)*

Items Specific to Experiments 3, 4b, 5, and 6b

Some of our participants report using particular strategies or rules to guide their decisions during the game. For instance, they might decide that they will only switch their response if they get two of the opposite outcome in a row. Others don't follow rules like these. Instead they just make their best guess at every turn. Did you use any strategies or rules when playing the game? *Yes; No*

If you did follow a strategy or rule, please tell us what it was here. *[open ended response]*

What did you think this study was trying to test? *[open ended response]*

Do you have any advice for us about how to improve the games so that future players feel more engaged? *[open ended response]*

Feel free to note anything else that you would like to bring to our attention here. It will be much appreciated. *[open ended response]*

What is your Mechanical Turk ID? This will enable us to pay bonuses more efficiently. *[open ended response]*

Item Specific to Experiment 3

It is very important to us to know how you approached the memory task. Please answer the next question honestly. Your answer will not affect your bonus at all. When playing the game, did you try keep track of the outcomes of each turn? *I tried my best to keep them in mind; I didn't deliberately keep track of the outcomes, I just tried to recall them whenever I was asked; The memory task was too hard, I just made random guesses; I wrote / typed the outcomes down; Other:*

Items specific to Experiment 5

Was this game too difficult? *Yes; No*

It is very important to us to know whether you were trying to keep the letter in mind while playing the game. Please answer the next question honestly. Your answer will not affect your bonus at all. When playing the game, how did you try keep track of the letter you were given to remember? *I tried my best to keep it in mind; I wrote / typed it down; It was too hard, I didn't even try to remember it; Other:*

Finally, in previous versions of this game, it seems that 1-2% of our participants have trouble with the game loading. We have no clue why this is! If the game didn't load for you, could you please let us know which browser version you use (e.g. "Chrome ver 38")? This might help us eliminate the problem. Also note that if you had any technical difficulties and were unable to earn a bonus as a result, we will pay you the average bonus received by the other participants.

Appendix C

Effect of demographic items on choice per experiment

Fixed Effects	*Exp 1*	*Exp 2*	*Exp 3*	*Exp 4a*	*Exp 4b*	*Exp 5*	*Exp 6a*	*Exp 6b*
Intercept (bias)	**0.15****	**0.23*****	**0.23*****	**0.26*****	**0.25*****	**0.21*****	**0.2*****	**0.19*****
Outcome t-1	**0.93*****	**1*****	**1.18*****	**1.05*****	**0.93*****	**1.05*****	**1.07*****	**0.99*****
Age	0.07	0.002	-0.03	0.002	0.02	-0.01	-0.04	0.01
Gender	**0.1***	**0.07***	-0.06	0.004	0.05	0.06	-0.04	-0.004
Social Conservatism	-0.3	0.03	0.03	0.02	0.01	<0.001	0.05	0.05
Fiscal Conservatism	0.07	-0.05	-0.003	-0.07	-0.01	0.008	-0.02	-0.04
Education	-0.04	-0.01	-0.02	**-0.09***	0.05	0.06	0.04	0.03
Income	0.05	0.01	0.02	0.02	-0.001	-0.01	0.07	-0.03
Delay				**-0.04***	**-0.06*****			
Load						0.02		
Reflection							0.07	-0.03
Take Time							0.002	

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. Only main effects are listed for reasons of space. Significant effects highlighted in boldface. * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix D

Effect of engagement items on choice per experiment

	Exp1	*Exp2*	*Exp3*	*Exp4a*	*Exp4b*	*Exp5*	*Exp6a*	*Exp6b*
Intercept (bias)	**0.15****	**0.27*****	**0.28*****	**0.33*****	**0.29*****	**0.23*****	**0.23*****	**0.22*****
Outcome t-1	**0.96*****	**1.17*****	**1.37*****	**1.27*****	**1.06*****	**1.24*****	**1.16*****	**1.13*****
Outcome t-2	**0.63*****	**1.04*****	**0.98*****	**1.02*****	**0.94*****	**0.96*****	**0.97*****	**1.04*****
Comprehension	-0.06	-0.04	-0.04	0.001	**-0.09***	0.02	0.03	-0.02
Enjoyment	0.05	0.05	0.01	0.08	0.06	0.06	0.04	0.01
Try	-0.03	0.01	0.08	-0.07	0.003	0.05	<0.001	**0.06***
Comp Check Fail	0.08	-0.05	-0.02	0.002	-0.02	-0.01	0.03	0.03
Strategizer	-0.01	**-0.07***	-0.03	-0.03	-0.05	-0.001	-0.06	**-0.05***
Guessed Purpose	0.03	-0.01	-0.02	0.01	**0.1****		0.001	-0.003
Ot-1 x Ot-2	**-0.03***		**0.03***					
Delay				-0.07	-0.06			
Delay x Ot-1					**0.1*****			
Delay x Ot-2					**0.06****			
Load						0.02		
Reflection							0.1	-0.02
Take Time							0.002	

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. All estimates listed. Some models were pared back to enable convergence. Significant effects highlighted in boldface. * p < .05, ** p < .01, *** p < .001

Appendix E

Ideal Observer Model

We require our ideal observer model to generate the best trial by trial estimates of the current task state given explicit knowledge of the task parameters (state transition probabilities; likelihood of different outcomes per state) and the outcome history. To do this, we used a modified version of the Hidden Markov Model from (Hampton et al., 2006). For each trial, the probability of being in Positive Mode is calculated by:

$$P(ST_T = Pos.|O_T) = \frac{P(O_T|S_T = Pos.)P(S_{T-1} = Pos.)}{\sum_{1..k} P(O_T|S_T = k)\, P(S_{T-1} = k)} \quad (1)$$

where the left hand term refers to the probability that the agent is in Positive Mode given the outcome they just encountered, $P(O_T|S_T = Pos.)$ is the probability of the current outcome if the task was in Positive Mode; $P(S_{T-1} = Pos.)$ refers to the prior probability of being in Positive Mode (as estimated from the previous trial), and the denominator refers to the probability of the current outcome across all states, i.e. the base rate of the outcome.

Since the task can switch between modes after each trial, this probability is then tempered by the probability that the task has switched:

$$P(S_{T+1} = Pos.) = \big(P(S_T = Pos.)\times(1-\gamma)\big) + (P(S_T = Neg.)\times\gamma) \quad (2)$$

where γ is the probability that the task has switched states.

We used this ideal observer model to generate estimates of the likelihood of being in Positive Mode for each outcome history bin given a starting estimate of 0.5, and running the model over the sequence of outcomes that defined that bin.

Appendix F

Effect of trial, outcome, and condition on mode judgment for experiments in Chapter 3.

	Modeled without random slopes					*Modeled without outcome regressor*				
Fixed Effects	*Exp4a*	*Exp4b*	*Exp5*	*Exp6a*	*Exp6b*	*Exp4a*	*Exp4b*	*Exp5*	*Exp6a*	*Exp6b*
Intrcpt (bias)	**0.25***	**0.26***	**0.2***	**0.15**	**0.2***	**0.18***	**0.22***	**0.17***	**0.12**	**0.18***
Ot-1	**1.02***	**0.9***	**1***	**1.02***	**0.97***					
Trial	0.002	**0.03***	**0.04***	0.03	0.01	-0.01	0.03	0.05	0.05	0.01
TrialQ	0.01	-0.01	0.01	0.05*	-0.02	0.03	0.001	0.01	0.06	-0.03
Del.	-0.05	**-0.14***				-0.05	**-0.14***			
Ld.			-0.003					-0.01		
Refl.				0.07	-0.06				0.09	-0.05
Take Time				0.001					0.01	
Ot-1 x Tr.	0.01	-0.02	-0.03	-0.01	**-0.02***					
Ot-1 x TrQ	0.03	**0.03***	**0.04***	**0.06**	0.01					
Del x Tr	**0.06***	-0.01				**0.06**	-0.01			
Del x TrQ	0.01	**0.08***				0.001	**0.07***			
Del x Ot-1	**0.08***	**0.11***								
Ld. x Tr			-0.04					-0.05		
Ld. x TrQ			0.02					0.02		
Ld. x Ot-1			**0.12**							
Refl. x Tr				-0.01	**0.04***				-0.03	0.04
Refl. x TrQ				0.004	0.03				-0.02	0.03
Refl. x Ot-1				**-0.13***	**-0.07**					
TT. x Tr				-0.04					-0.05	
TT. x TrQ				-0.002					-0.02	
TT. x Ot-1				-0.001						

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode. Significant effects highlighted in boldface. * p < .05, ** p < .01, *** p < .001

Appendix G

Effects of choice, outcome, and condition on reaction time for experiments in Chapter 3

Fixed Effects	*Exp 4a*	*Exp 4b*	*Exp 5*	*Exp 6a*	*Exp 6b*
Intercept	**6.67*****	**6.58*****	**6.18*****	**6.18*****	**6.12*****
Choice	**-0.04*****	**-0.03*****	**-0.03****	-0.01	**-0.05*****
Outcome t-1	**0.01*****	0.004	-0.002	0.01	**0.02*****
Outcome t-2	**0.01****	-0.004	-0.001	**-0.02***	**0.02*****
Delay	**0.37*****	**0.61*****			
Load			**0.34*****		
Reflection				-0.04	-0.04
Take Time				0.01	
Choice x Ot-1	**-0.1*****	**-0.08*****	**-0.16*****	**-0.12*****	**-0.15*****
Choice x Ot-2	**-0.02*****	**-0.02*****	**-0.05*****	**-0.04*****	**-0.05*****
Ot-1 x Ot-2	**-0.06*****	**-0.07*****	**-0.11*****	**-0.09*****	**-0.1*****
Delay x Choice	**0.03*****	**0.02*****			
Delay x Ot-1	-0.003	0.002			
Delay x Ot-2	**-0.01***	0.002			
Load x Choice			0.02		
Load x Ot-1			0.002		
Load x Ot-2			-0.006		
Reflection x Choice				**-0.05****	0.003
Reflection x Ot-1				0.001	0.01
Reflection x Ot-2				0.02	-0.01
Take Time x Choice				-0.02	
Take Time x Ot-1				-0.001	
Take Time x Ot-2				0.01	

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as -1 for Negative Mode and 1 for Positive Mode. Only main effects and first level interactions are listed for reasons of space. Significant effects highlighted in boldface. * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix H

3-back plots of behavior and reaction time for each experiment

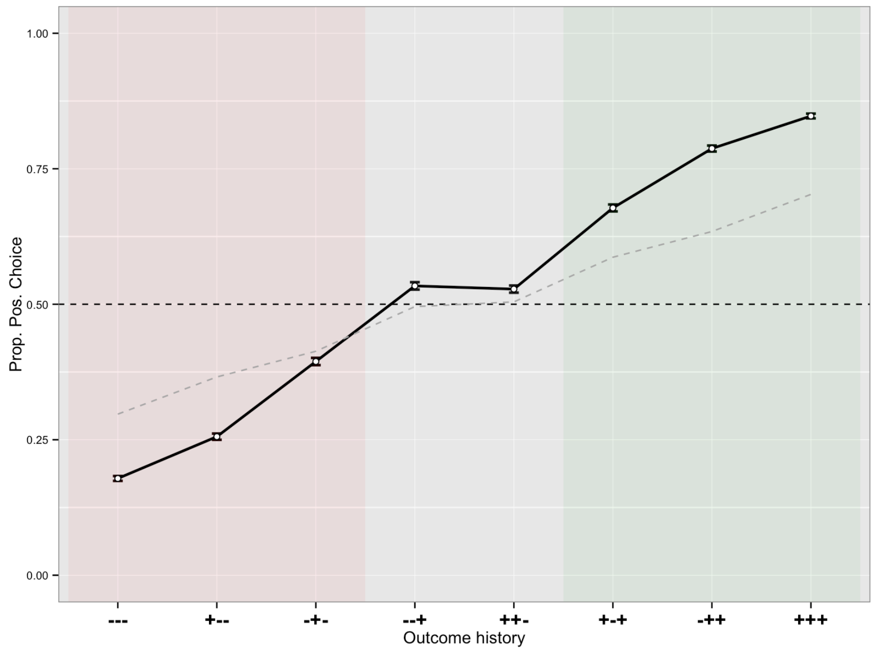


Figure H1. Experiment 1 choice

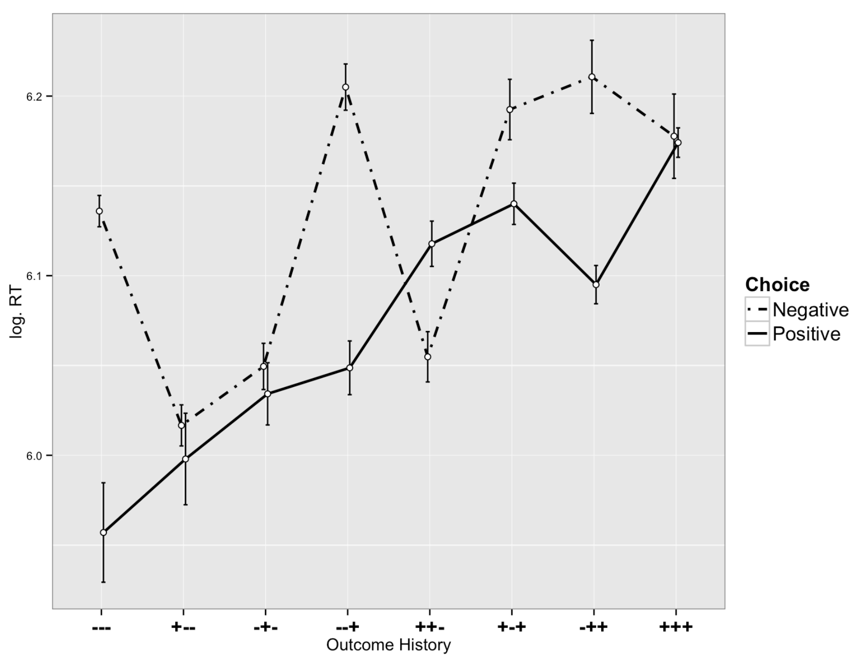


Figure H2. Experiment 1 reaction time

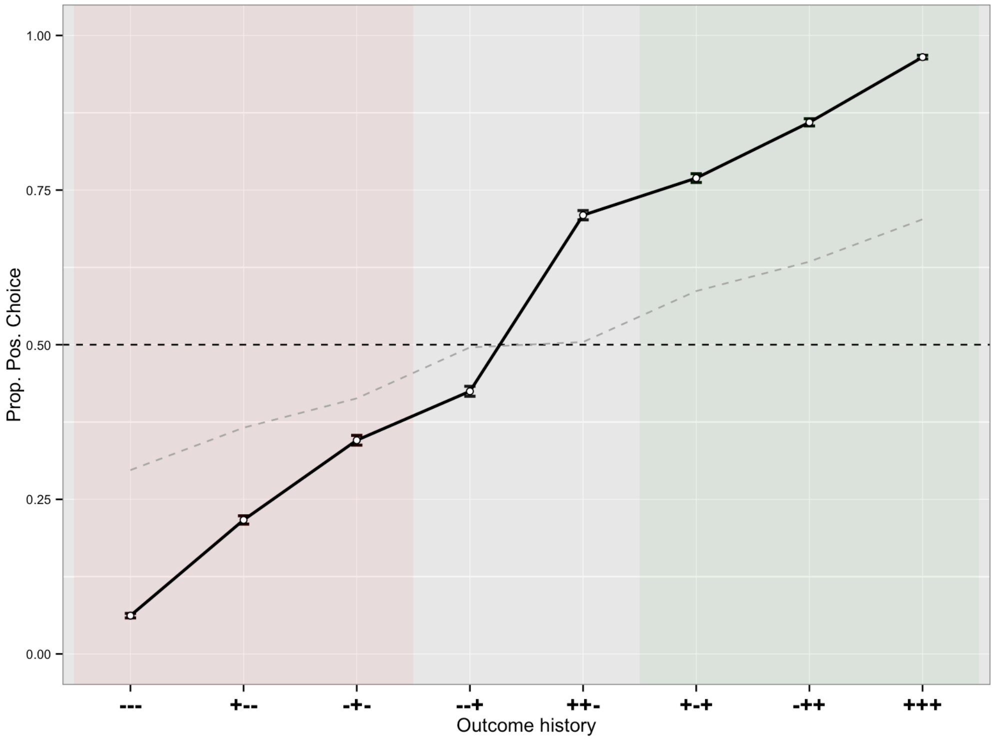


Figure H3. Experiment 2 choice

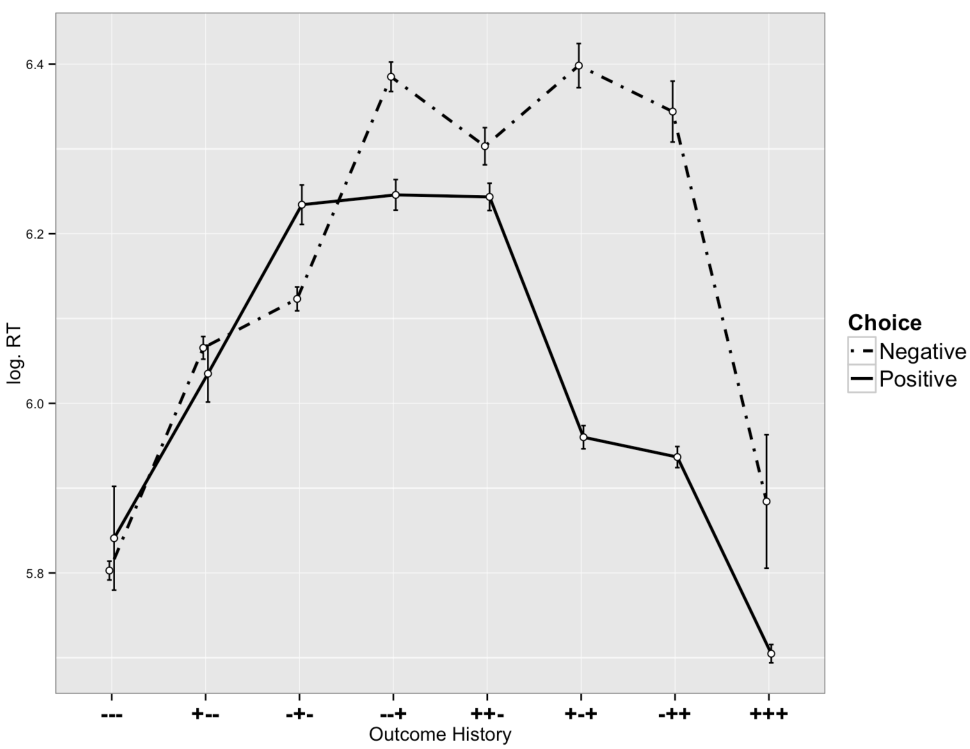


Figure H4. Experiment 2 reaction time

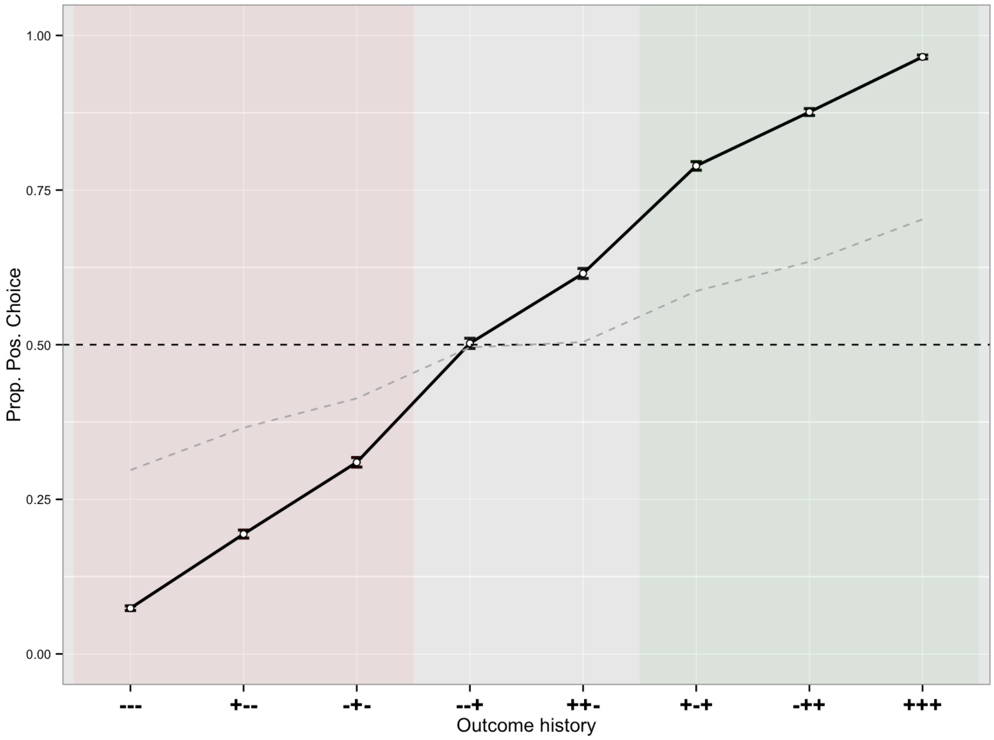


Figure H5. Experiment 3 choice

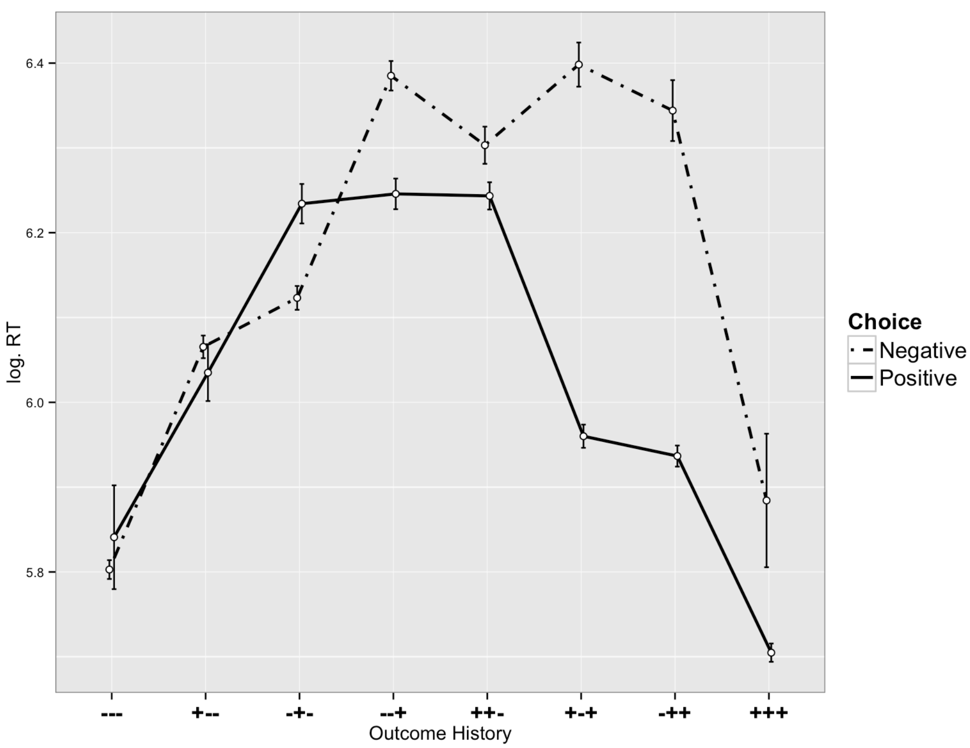


Figure H6. Experiment 3 reaction time

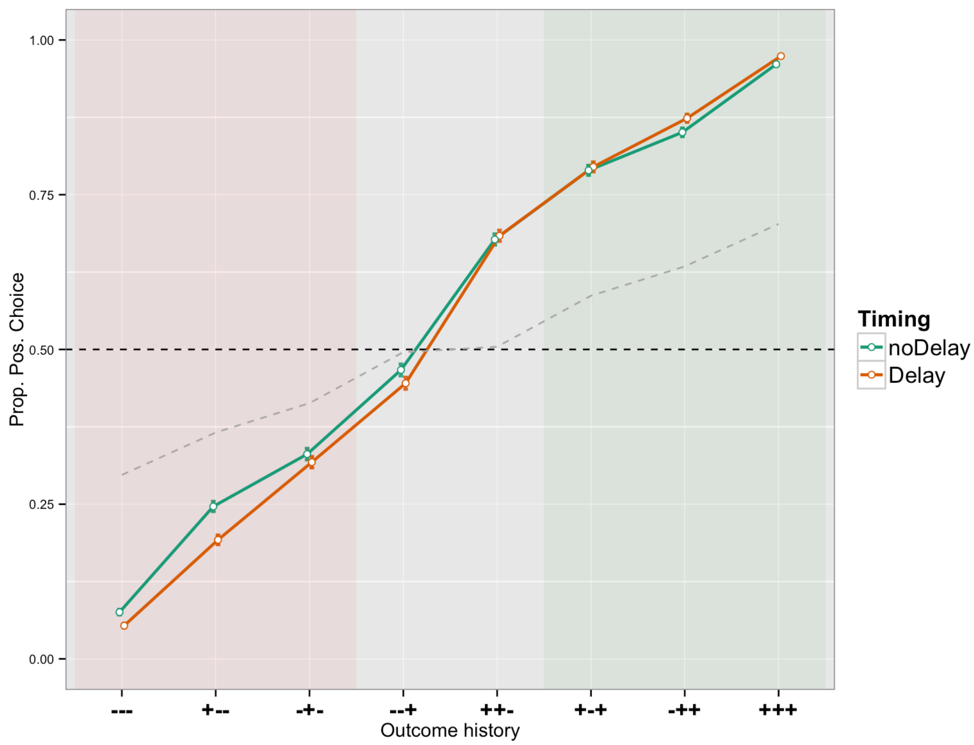


Figure H7. Experiment 4a choice

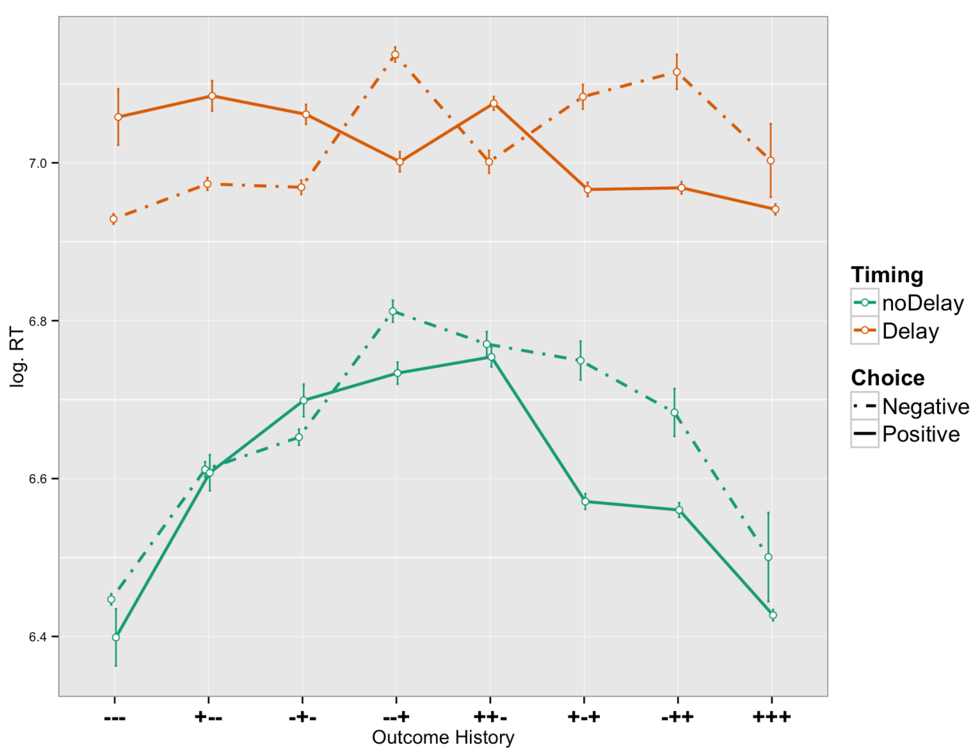


Figure H8. Experiment 4a reaction time

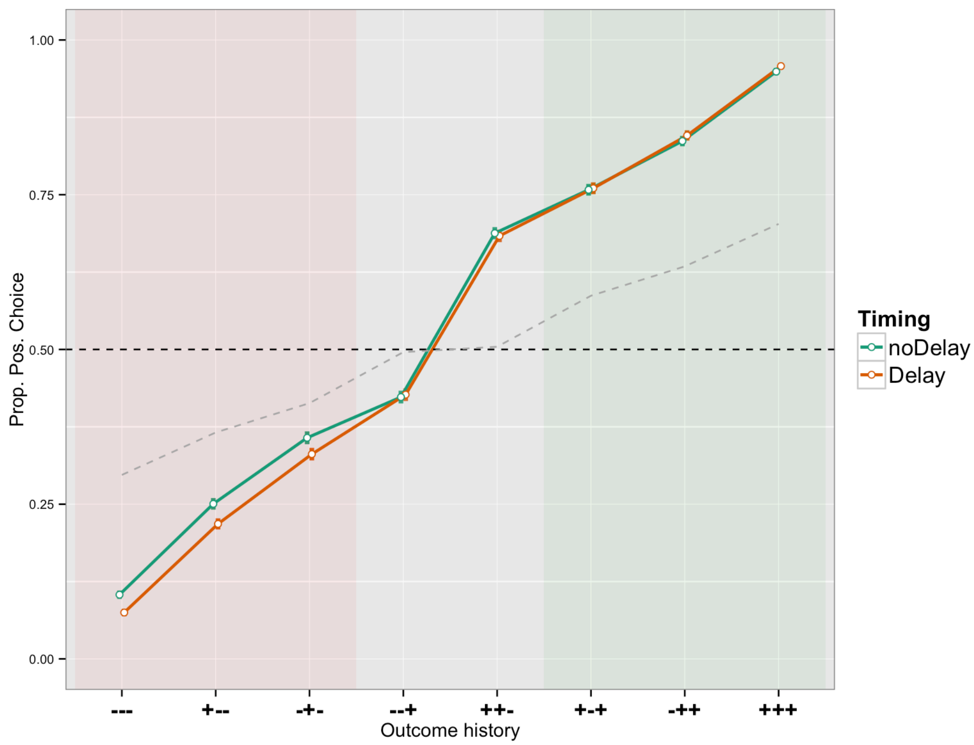


Figure H9. Experiment 4b choice

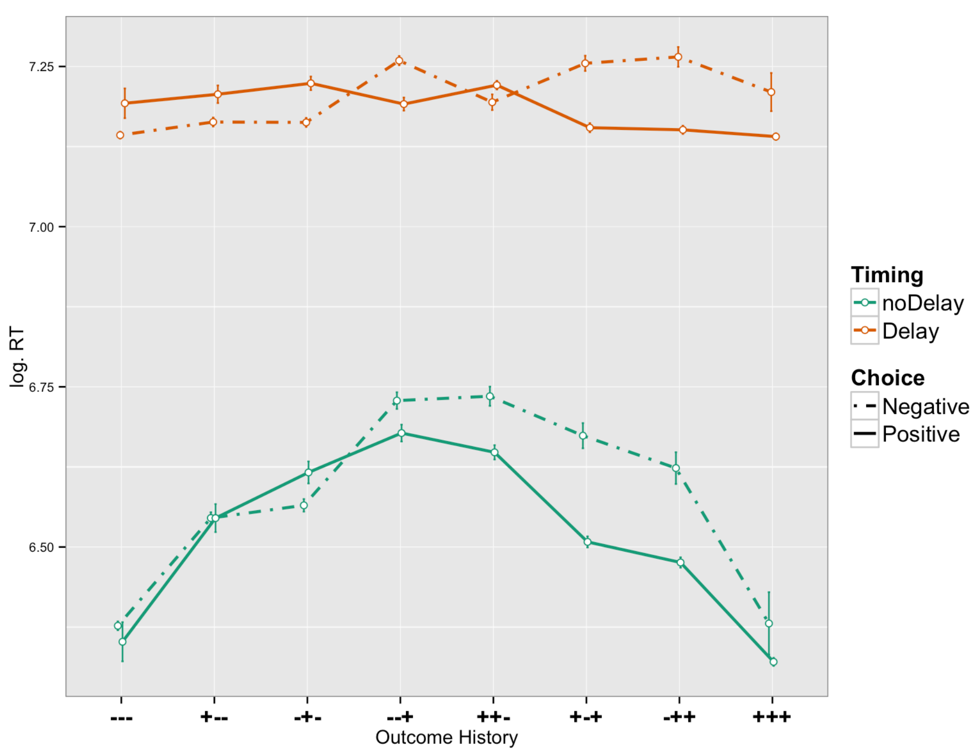
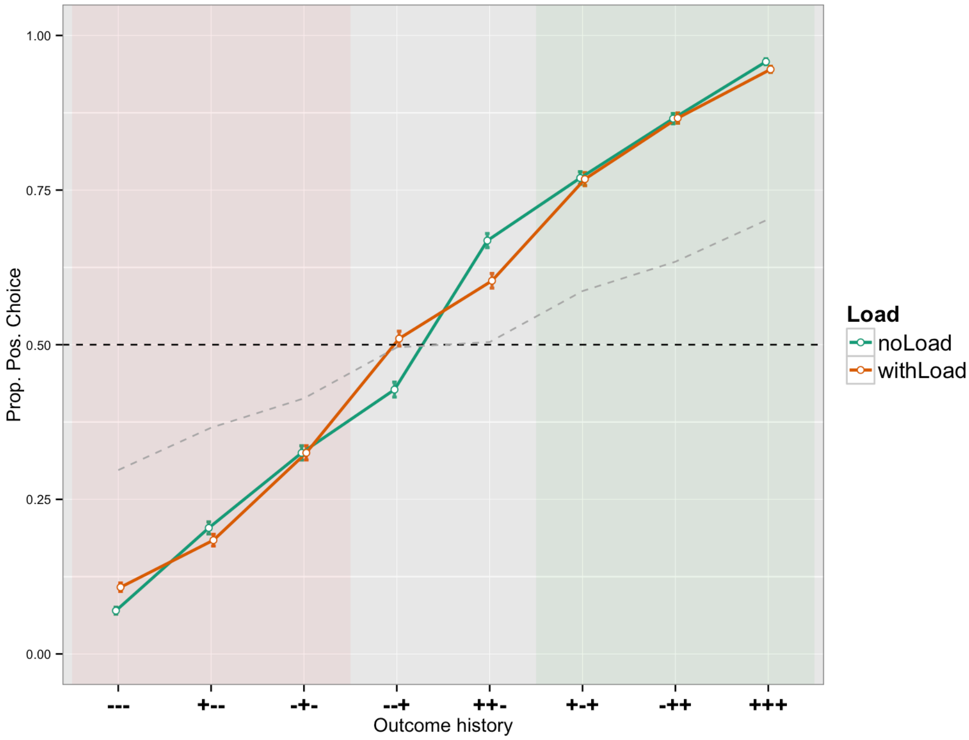


Figure H10. Experiment 4b reaction time



Figure_apx 11. Experiment 5 choice

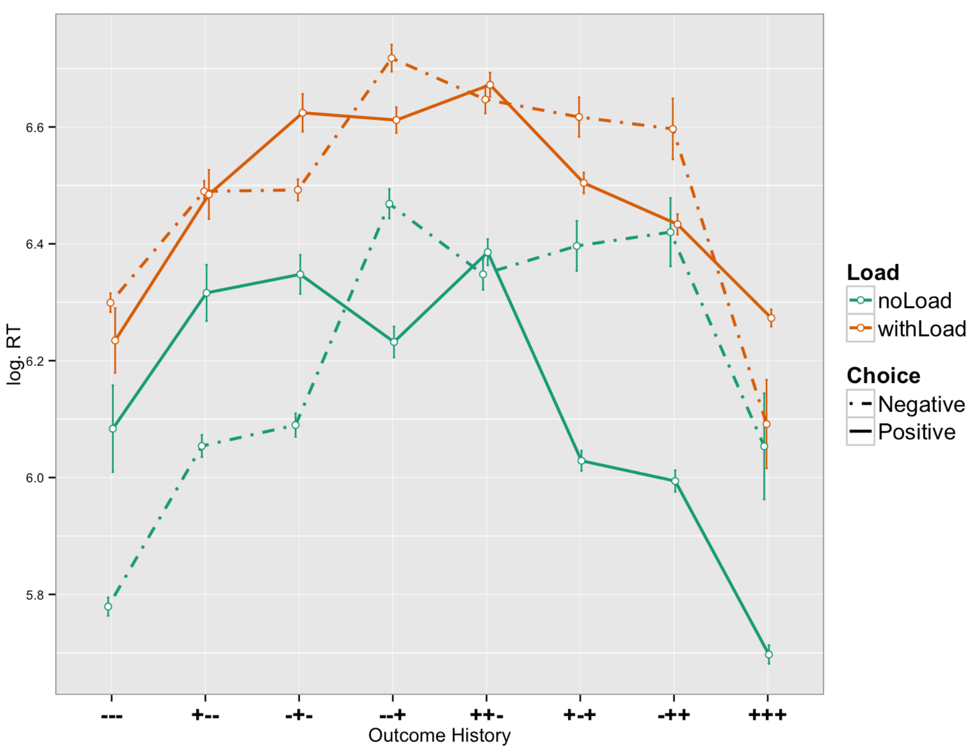


Figure H12. Experiment 5 reaction time

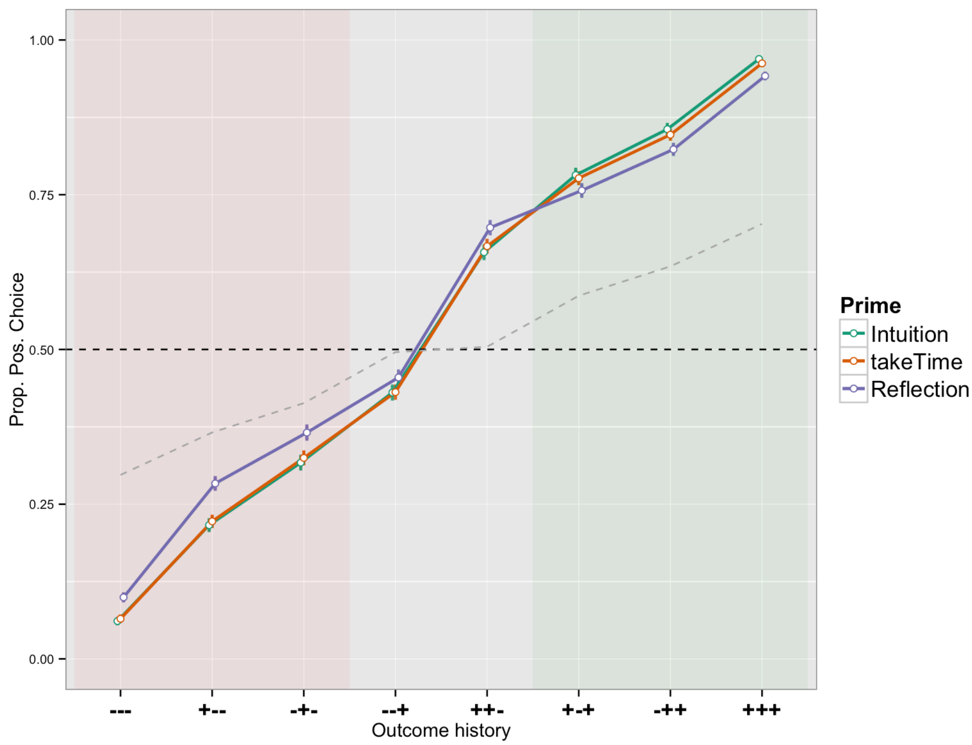


Figure H13. Experiment 6a choice

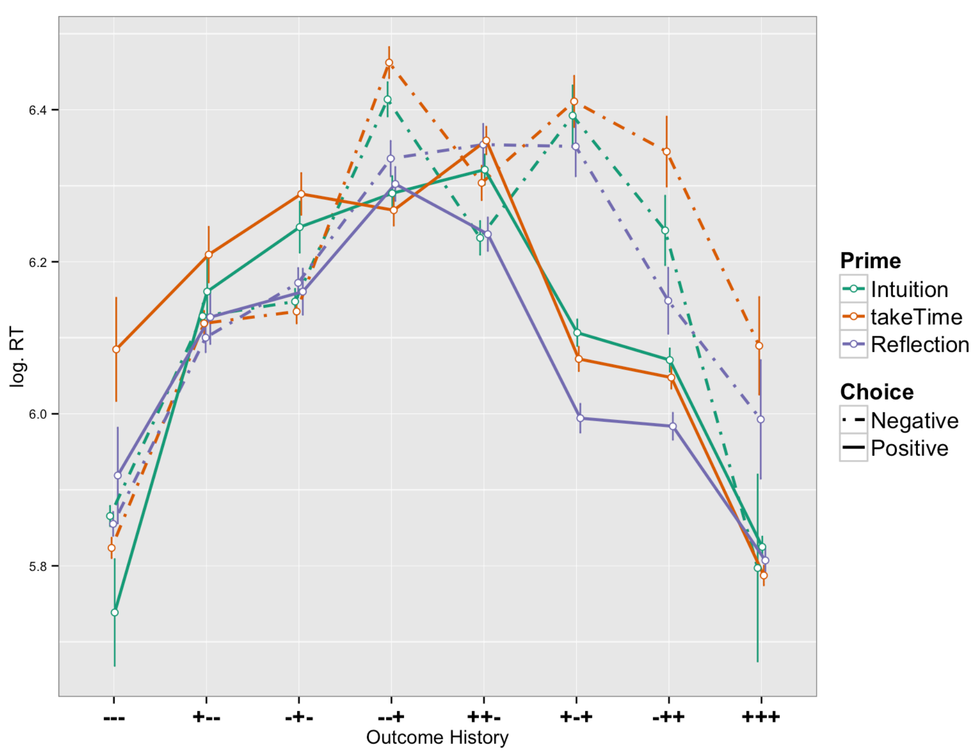


Figure H14. Experiment 6a reaction time

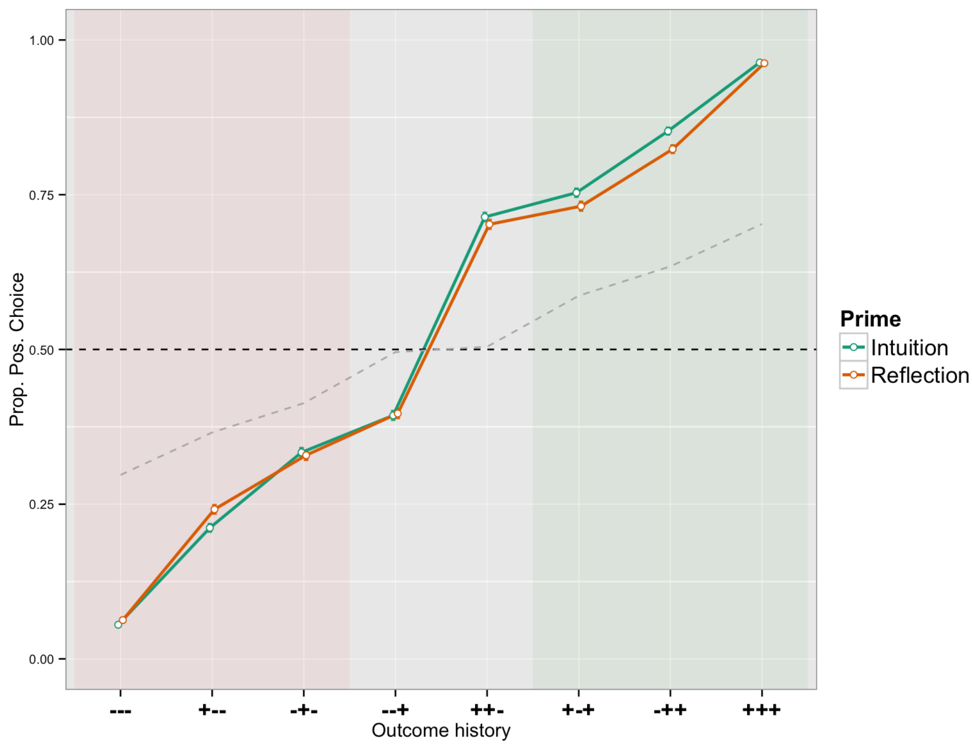


Figure H15. Experiment 6b choice

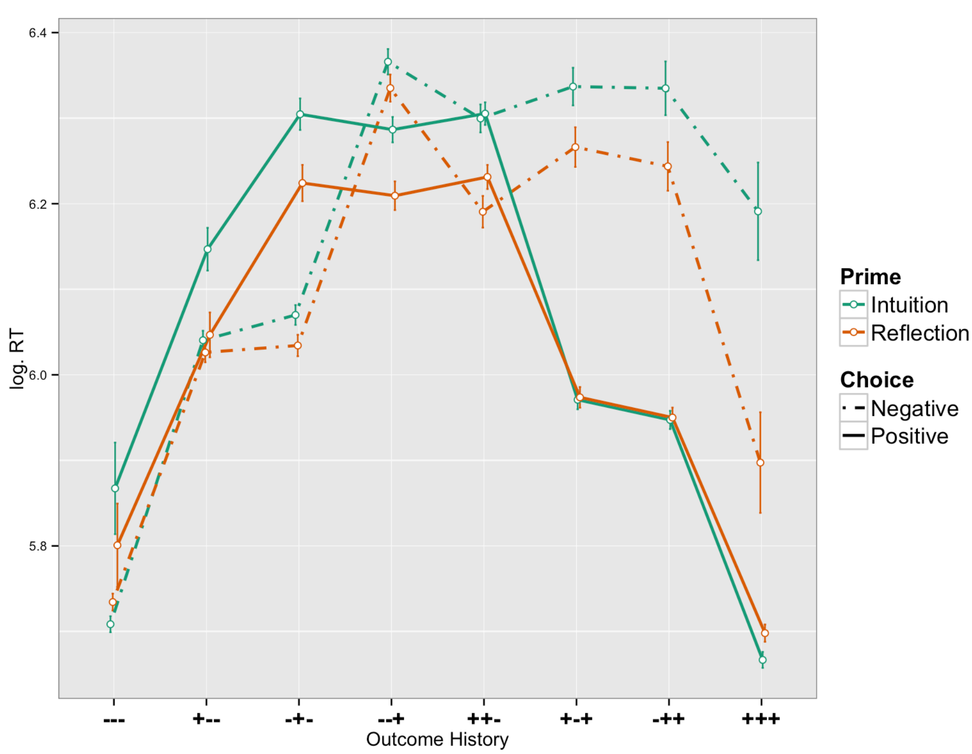


Figure H16. Experiment 6b reaction time

Appendix I

Analysis of Merged Data from Experiments 6a and 6b

Table I1

Experiment 6a and 6b: Effect of previous trial outcomes and condition on likelihood to choose Positive Mode

Fixed Effects	*Estimate*	*Std. Error*	*z value*	*p*
Intercept (bias)	0.23	0.03	7.33	< 0.001
Outcome t-1	1.19	0.01	111.1	< 0.001
Outcome t-2	1.07	0.01	99.57	< 0.001
Reflection	0.01	0.04	0.12	0.908
Ot-1 x Ot-2	<0.001	0.01	0.02	0.983
Ot-1 x Reflection	-0.13	0.02	-8.41	< 0.001
Ot-2 x Reflection	-0.1	0.02	-6.55	< 0.001
Ot-1 x Ot-2 x Refl.	0.01	0.01	0.93	0.354

Note. Outcome t-1 represents most recent outcome. Outcome regressors coded as 1 for reward and -1 for punishment. Choice coded as 0 for Negative Mode and 1 for Positive Mode.

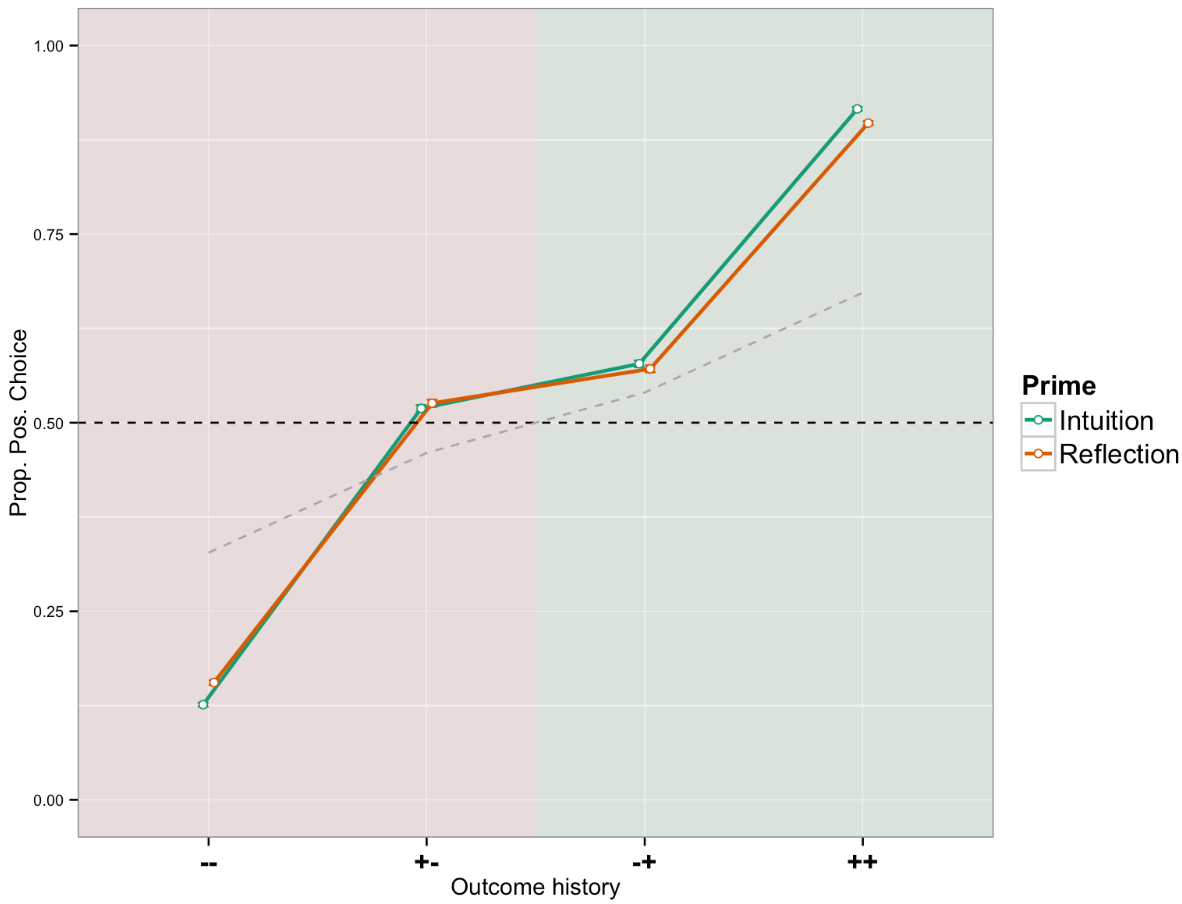


Figure I 1. Choice behavior for the merged data of Experiments 6a and 6b. Outcome history prior to choice is plotted on the x-axis. Each bin reads as two outcomes ago followed by most recent outcome. The y-axis represents the proportion of Positive Mode judgments for each bin, with error bars calculated using the adjustment in Morey (2008). The light grey dotted line represents the likelihood of being in Positive Mode (corresponding to y-axis values) as calculated by an ideal observer model (see Appendix E for details). Background color represents when a rational observer would choose Positive Mode (green) or Negative Mode (red). Bias can be evaluated as an asymmetry of choice about the equi-proportional line.