TI 2019-042/I
Tinbergen Institute Discussion Paper

Anticipatory Anxiety and Wishful Thinking

Jan Engelmann[1]
Maël Lebreton[2]
Peter Schwardmann[3]
Joël van der Weele[1]
Li-Ang Chang[4]

[1] University of Amsterdam
[2] University of Geneva
[3] LMU Munich
[4] CREED - University of Amsterdam

Anticipatory Anxiety and Wishful Thinking*

Jan Engelmann[1], Maël Lebreton[2], Peter Schwardmann[3], Joël van der Weele[1], and Li-Ang Chang[1]

[1]CREED - University of Amsterdam

[2]University of Geneva

[3]LMU Munich

June 17, 2019

Abstract

It is widely hypothesized that anxiety and worry about an uncertain future lead to the adoption of comforting beliefs or "wishful thinking". However, there is little direct causal evidence for this effect. In our experiment, participants perform a visual pattern recognition task where some patterns may result in the delivery of an electric shock, a proven way of inducing anxiety. Participants engage in significant wishful thinking, as they are less likely to correctly identify patterns that they know may lead to a shock. Greater ambiguity of the pattern facilitates wishful thinking. Raising incentives for accuracy does not significantly decrease it.

JEL classification: C91, D83

Keywords: confidence, beliefs, anticipatory utility, anxiety, motivated cognition

*The authors wish to thank participants in Frankfurt, Bonn and Maastricht for comments. Nik Mautner Markhof provided excellent research assistance.

1 Introduction

Many commonly held beliefs seem to be inspired by their comforting properties rather than their realism. For instance, billions of adherents of the major religions believe in an afterlife, despite a lack of evidence for its existence. Moreover, religiosity is higher in populations that face more unpredictable shocks like earthquakes (Sinding Bentzen, 2019), and declines with the provision of alternative forms of insurance (Auriol et al., 2017). Populist politicians that promise easy fixes find more support in areas with weak economic prospects and declining growth rates (Mughan et al., 2003; Obschonka et al., 2018). People at risk of serious diseases avoid medical testing in order to remain optimistic about their health status (Lerman et al., 1998; Oster et al., 2013; Ganguly and Tasoff, 2016). While these facts are suggestive, they do not establish a direct causal effect of anticipatory anxiety on the adoption of comforting beliefs, or "wishful thinking".

To fill this gap, we study the effect of experimentally induced anticipatory anxiety on belief formation. Participants in our experiment are incentivized to identify the tilt (left vs. right) of a pattern flashed on a screen. The pattern's actual tilt is randomly determined by the computer, with both tilts being equally likely. We address the key challenge of experimentally inducing anxiety by wiring participants to an electric stimulation device that delivers an electric shock after some patterns. In our main treatment, we vary whether electric shocks are associated with right-tilted or left-tilted patterns. Thus, unlike in most field settings, it is clear to participants that the shock, the cause of their anxiety, is outside of their control. We identify wishful thinking by comparing participants' beliefs in the pattern recognition on patterns associated with a shock with their beliefs on patterns associated with no shock. In further treatment dimensions we vary the ambiguity of the pattern and the incentives for accurate reporting.

We find strong evidence for wishful thinking. Subjects are significantly less accurate and confident in identifying patterns that may lead to an electric shock. Because our setting allows us to run many trials per subject and administer within-subject treatments, we are able to classify participants according to their propensity to engage in wishful thinking. We uncover substantial heterogeneity in wishful thinking, with a majority of subjects displaying the phenomenon. We also find that greater ambiguity of the pattern leads to significantly more wishful thinking. Higher incentives for accuracy do not have a significant effect.

Our preregistered experiment contributes to the literature in three ways. First and foremost, we demonstrate the causal effect of anxiety on self-deception and overoptimistic beliefs.[1] A few studies investigate wishful thinking over desirable outcomes and find somewhat mixed results. Mijović-Prelec et al. (2010) compare a treatment with a high monetary prize and low incentives for accuracy with a treatment that features a low monetary prize and high incentives for accuracy. Beliefs are more favourably biased in the former. In Coutts (2019) and Mayraz (2011), subjects engage in estimation tasks where they have an exogenously given monetary stake in some of the outcomes. Both studies document optimism, although Coutts (2019) finds it only for one out of three tasks. Barron (2016) finds no evidence for wishful thinking in the updating behavior of experimental subjects. Relative to the use of monetary prizes in these studies, electric shocks have the advantage of being precisely-timed and salient consumption events and a proven method of inducing anticipatory utility.[2]

Second, the positive effect of greater ambiguity of the pattern on wishful thinking demonstrates the importance of the cognitive limits of self-deception (see also Sloman et al., 2010; Caplin and Leahy, 2019). Our results support the idea that the avoidance of precise information, such as diagnostic medical tests, is an important facilitator of wishful thinking. Because evidence constrains people's ability to self-deceive, they may be wary of receiving it.[3] The important role of ambiguity in our experiment speaks to a small literature on motivated visual perception that investigates whether people "see what they want to see" in ambiguous visual evidence (Dunning and Balcetis, 2013). To verify that people truly believe their reported motivated perceptions, this literature relies on implicit questionnaire items as well as eye tracking measures and reaction times (Balcetis and Dunning, 2006). We provide complementary evidence from incentivized elicitations that make truthful reporting a money maximizing strategy.

Third, we test a core prediction of theories that model wishful thinking as a trade-off between its anticipatory utility benefits and the (material) costs of adopting wrong beliefs (Brunnermeier and Parker, 2005; Bénabou and Tirole, 2011). We do not find evidence that

[1] Falk and Zimmermann (2017) provide complementary experimental evidence on preferences over information about the occurrence of an electric shock, e.g. whether participants have a preference over early or late resolution of uncertainty.

[2] Cook and Barnes (1964) and Loewenstein (1987) show evidence that if possible, people dislike to delay electric shocks in order to minimize anxiety. For more evidence on the relation between shocks and anxiety, see Berns et al. (2006); Schmitz and Grillon (2012); Engelmann et al. (2015, 2019).

[3] This point also helps address the criticism by Spiegler (2008) that models of wishful thinking (e.g. Brunnermeier and Parker 2005) cannot account for people's tendency to avoid anxiety-inducing information. Instead, our results suggest that information preferences can be captured by introducing into these models cognitive costs of self-deception that are increasing in the precision the evidence encountered by the agent.

higher incentives reduce wishful thinking, and therefore no support for these theories. However, the evidence from our pattern recognition task leaves little scope for "motivated reasoning" and, instead, reflects an instantaneous form of self-deception (see also Kappes and Sharot 2019). It is plausible that material incentives play a more important role in the more deliberate process of "motivated reasoning", just as they have been shown to matter in the remembering of ego-relevant feedback (Zimmermann, 2019).[4]

In the next section, we describe our experimental design. Section 3 introduces a simple theoretical model that helps us derive our hypotheses. Section 4 features the results of our experiment. In section 5, we discuss robustness and different interpretations of our results, before the last section concludes.

2 Experimental design

Procedures. Our computerized experiment took place in the CREED experimental laboratory at the University of Amsterdam. It was registered on aspredicted.org, the preregistration can be downloaded here). Sixty subjects participated in individual sessions. Upon coming to the lab, subjects read the instructions (available in Appendix B), signed a consent form and answered several control questions to determine their understanding of the task and the belief elicitation mechanism. The experimenter pointed out any wrong answers and discussed the correct answer until the participant indicated (s)he understood them.

Next, the strength of the electric shock was calibrated. The wrist of the participant's non-dominant hand was connected to a Digitimer DS5 isolated bipolar current stimulator, which itself was connected to MATLAB through National Instruments USB x-series. The participant induced herself with a series of shocks, which she rated on a pain scale of 0 (not painful at all) to 10 (extremely painful). The calibration was complete when the subject rated the pain as 7 on the scale three consecutive times. The maximum possible shock strength was set to 5V 25mA and the duration of the shock was set to 50ms (Engelmann et al., 2015, 2019).

The task and confidence elicitation procedures are outlined in detail in Figure 1 and were inspired by Lebreton et al. (2018). Briefly, the main task consisted of three parts ("sessions") and lasted approximately 45 minutes. Each part was made up of 4 blocks consisting of

[4] Our findings are consistent with findings of studies looking at wishful thinking in the gain domain, where increasing incentives for accuracy has no effect on the belief bias in Mayraz (2011) and increases it in Coutts (2019).

18 trials each. At the start of each block, the participant was informed of the treatment for the next 18 trials via a "block cue". Treatment conditions included that shocks would occur after left-, or right-tilted visual patterns, whether payment for correct performance was 1 or 20 Euros and how difficult discrimination was (see the "Treatments" below for details). On a given trial, participants were asked to distinguish whether a pattern (Gabor patch) was tilted to the left or to the right. Subjects first indicated their answer ("left" or "right") and then submitted their confidence in the answer on a probabilistic scale from 50 to 100. Their confidence rating was incentivized with a Becker-deGroot-Marshak or "matching probabilities" mechanism. This mechanism makes it incentive compatible to state the true beliefs, regardless of the risk preferences of the participant.[5]

The resolution of the BDM procedure was not shown until after completion of the experiment. Instead, after the confidence rating, a screen appeared for 2-8 seconds with the text "Please wait... resolving shock..." followed by a message saying either "No Shock" or "Shock!", in which the participant received an electric shock dependent on the trial type. At the end of each block the overall accuracy of decisions in that block was displayed.

Treatments. We conduct three treatment variations, all of which were administered within-subject. Our first and main treatment varied whether a shock was administered after right-tilted or left-tilted patterns. We made it very clear to subjects that the occurrence of the shock depended only on the actual pattern, and not on their report. We distinguish between trials in which shock and correct answer were *Aligned* and trials in which they were *Non-Aligned*. Wishful thinking will be measured as the difference between average accuracy or confidence *Non-Aligned* and *Aligned*.

In our second treatment we varied the difficulty of the pattern recognition task within each block, as measured by the degree of the tilt from the vertical line. The task had three difficulty levels that were calibrated to result in accuracy levels of 60%, 70% and 80%. We will refer to task difficulty as the "ambiguity" of the evidence. Initially, these levels were calibrated on the basis of a pilot, and were the same for all subjects. After each part (4 blocks of 18 trials), difficulty levels were recalibrated for each subject, using a logistical

[5] Subjects indicate their point of indifference $x \in \{50, 55, ..., 95, 100\}$ between Lottery 1, earning a prize M with probability x, and Lottery 2, earning a prize M if their answer was correct, where M varied between experimental conditions. After the choice, the computer randomly draws a number $z \in \{50, 55, ..., 95, 100\}$. The participant receives Lottery 1 if $x \geq z$, and Lottery 2 otherwise. Schlag et al. (2015) provide details about the origins and incentive compatibility of this mechanism, see Trautmann and van de Kuilen (2014) and Hollard et al. (2016) for evidence.

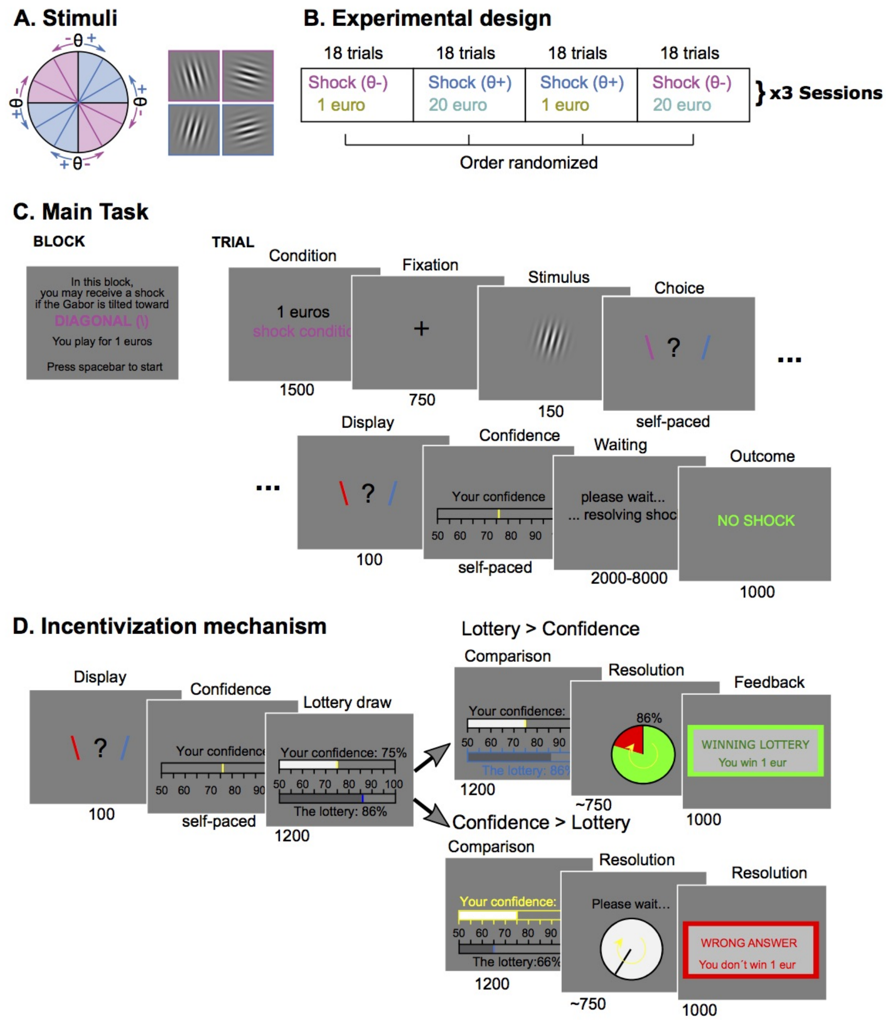


Figure 1: **Experimental Design.** **A.** Participants are tasked with determining the tilt of a grating (Gabor patch). The circle on the left pictures the entire stimulus space. Left-tilted gratings are colored in pink and denoted by $\theta-$, whereas right-tilted gratings are colored in blue and denoted by $\theta+$. The four insets on the right display examples of gratings that were used in the task. **B.** Each participant completed three sessions, each divided in four blocks of 18 trials. The four blocks correspond to four conditions of a 2x2 factorial design (Shock x Incentive). The Shock treatment varied whether a shock was administered if the grating was right-tilted or whether it was administered if the grating was left-tilted. The Incentive treatment varied whether the potential prize in the belief elicitation was 1 or 20 euros. **C.** Successive screens displayed in one trial are shown from left to right with durations in milliseconds. Left: At the beginning of each block, participants first saw a screen informing them about the Shock and Incentive conditions. Right: before each trial, subjects were reminded of the Shock and Incentive (1500ms) and after briefly seeing a fixation cross (750ms), the grating was flashed (150ms). This was followed by a response screen that depicted the two possible tilts on the left and right side of the screen. Participants could choose either tilt by pressing the left or right arrow on the keyboard (self-paced). The side of the screen on which each tilt was displayed was balanced to average out any lateralization effects. Then, the chosen class briefly turned red (100ms) and participants were asked to indicate their confidence in their choice on a scale from 50% (completely uncertain) to 100% (certainty). They did this by moving a cursor on a horizontal scale using the left and right arrows (5% steps), and could validate their answer with the spacebar (self-paced). Next, participants faced an anticipation screen (2000-8000ms), asking them to wait for the Shock resolution. Finally, the electric shock was administered or not (1000ms). No trial-by-trial feedback was given about the correctness of the guess, but the average performance was communicated at the end of each block. **D.** Before the experiment, the BDM mechanism used to elicit beliefs was explained to participants and they were then given an opportunity to gain experience with the mechanism. After the experiment, two trials were randomly selected for payment. The Figure depicts the set of screens in the practice run.

performance function. This procedure rules out effects of learning or fatigue.

Our final treatment varied in the potential prize for the BDM mechanism, which was either 1 euro or 20 euros. We call these the *Low Incentive* and *High Incentive* conditions respectively. These two cross-cutting treatments resulted in four conditions that were varied randomly across blocks of 18 trials each.

Payments. Subjects were paid at the end of the experiment. Participants' earnings consisted of an 10 euro show-up payment, plus the earnings from the accuracy payments of one randomly drawn trial from both the low and high incentive condition. Thus, payments varied between 10 and 31 euros.

3 Model

In this section, we present a stylized model of wishful thinking in our laboratory context in order to derive our main hypotheses. The model is in the spirit of Brunnermeier and Parker (2005) in that beliefs treated as objects of choice, but it includes a cognitive cost of self-deception as in Bénabou and Tirole (2002).

The correct answer, or state of the world, $\theta \in \{l, r\}$, is either that the pattern is left-tilted (l) or that it is right-tilted (r). A participant's initial probabilistic belief that $\theta = r$ is given by p. Subject to some cost, she may then self-deceive into a new belief $\hat{p}$. Assuming that the participant states her chosen belief $\hat{p}$, the Becker-DeGroot-Marshak mechanism implies the following expected material payoffs from potentially winning a prize M

$$\pi(p, \hat{p}) = \frac{1}{2}\left(1 + 2p\hat{p} - \hat{p}^2\right) M$$

The probability of winning the prize is maximized at $\hat{p} = p$. Therefore, if material payoffs were the only object in a participant's utility function, then she would not self-deceive. The material cost of self-deception is given by $\pi(p, p) - \pi(p, \hat{p})$.

A participant's anxiety of the electric shock is based only on her chosen beliefs $\hat{p}$ and is given by

$$\sigma(\alpha\hat{p} + (1 - \alpha)(1 - \hat{p}))$$

It depends on $\alpha \in \{0, 1\}$, which takes a value of 1 if and only if the shock is administered for right-tilted patterns. The variable σ captures the importance of anticipatory utility concerns, or a participant's innate anxiety.

Suppose next that self-deception is not frictionless, but instead subject to a quadratic cognitive cost $\lambda(s)(p-\hat{p})^2$. The cognitive cost function is increasing in the distance between a participant's initial belief and her chosen belief. λ captures the magnitude of the cognitive cost and we make the assumption that λ is increasing in s, the strength of the signal a participant encounters. The more ambiguous a pattern, the lower is s. Then, a subject's expected utility is given by

$$U = \frac{1}{2}\left(1 + 2p\hat{p} - \hat{p}^2\right) M - \sigma(\alpha\hat{p} + (1-\alpha)(1-\hat{p})) - \lambda(s)(p-\hat{p})^2$$

Maximizing the above expression with respect to $\hat{p}$ yields a participant's optimal belief

$$\hat{p}^* = p - \frac{\sigma(2\alpha - 1)}{2(M + \lambda(s))}$$

From the optimal beliefs we can derive hypotheses about our experimental treatments. We consider the case in which the correct answer is $\theta = r$, so that $\hat{p}$ is the belief in the correct answer. The case of $\theta = l$ is symmetric. Then, the *Aligned* condition corresponds to $\alpha = 1$ and the *Non-Aligned* condition corresponds to $\alpha = 0$. The amount of wishful thinking is given by

$$W := \hat{p}^*(\alpha = 0) - \hat{p}^*(\alpha = 1) = \frac{\sigma}{M + \lambda(s)} \tag{1}$$

From (1), and under the assumption that σ and λ are positive, we can derive the following main hypotheses.

Hypothesis 1 (Wishful thinking) *There is positive wishful thinking.*

We can also formulate the following two secondary hypotheses. The effect of ambiguity on wishful thinking follows directly from our assumption that $\lambda'(s) > 0$.

Hypothesis 2 (Ambiguity) *Wishful thinking decreases when the pattern is easier to identify, i.e.* $\frac{dW}{ds} < 0$.

Hypothesis 3 (Incentives) *Wishful thinking declines in the size of incentives. More precisely,* *a. In the Aligned condition, belief in the correct answer is higher when M is larger, i.e.* $\frac{d\hat{p}^*(\alpha=1)}{dM} > 0$.

b. In the Non-Aligned condition, belief in the correct answer is lower when M is larger, i.e. $\frac{d\hat{p}^*(\alpha=0)}{dM} < 0$.

4 Results

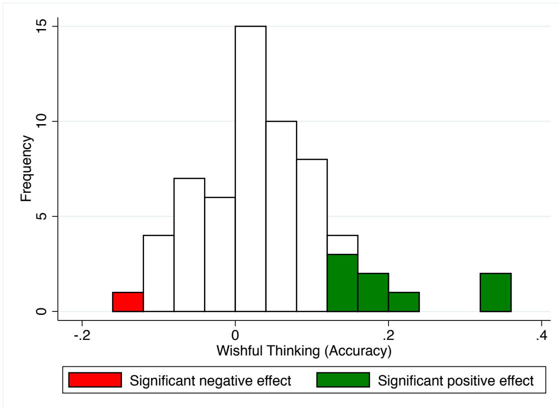
4.1 Wishful thinking

To test our main hypotheses, we consider two outcome variables: Accuracy and Belief. Both outcome measures are derived by taking a subject's average across all trials belonging to a given condition, yielding one observation per subject per condition. All results are robust if we use all trial data and regressions with subject fixed effects. "Accuracy" is the fraction of correct answers on the initial binary choice. The variable "Belief" measures the subjective belief in the correct answer based on our elicitations of confidence. Beliefs vary on a scale from 0 (meaning the subject indicated 100% confidence in the wrong answer) to 100 (meaning the subject indicated 100% confidence in the correct answer). We define wishful thinking as the average Accuracy/Belief on patterns associated with no shock minus the average Accuracy/Belief on patterns associated with a shock.

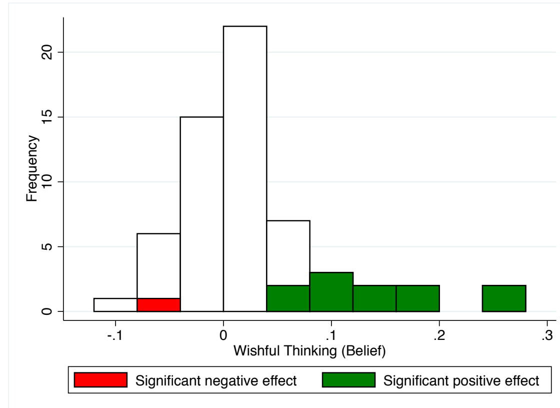
While Accuracy and Belief are related, we believe that they are of independent interest. Accuracy on the initial binary choice provides a first, "snap-shot" judgment of the right answer. Choices elicited in similar 2-alternative forced choice tasks have been extensively used in cognitive psychology, notably to reveal features of cognition that participants cannot report explicitly, are not aware or conscious of. Beliefs, on the other hand, integrate an explicit, subjective judgment of performance, in the form of confidence ratings. Psychological research shows that accuracy and beliefs may diverge, as the information underlying accuracy judgments may undergo continued processing and/or may not be fully accessible in the construction of confidence (Fleming and Daw, 2017).

Following our preregistered analysis,[6], we test for wishful thinking in Accuracy and Belief, pooling difficulty and incentive levels. We find statistically significant evidence for both accuracy ($\mu = 0.041$, $s.d. = 0.098$, $d = 0.62$, $t = 3.26$, $d.F. = 59$, $p = 0.0009$, one-sided), and

[6] See here `http://aspredicted.org/blind.php?x=mb5y37`.



(a) Accuracy



(b) Beliefs

Figure 2: Frequency of wishful thinking in accuracy and belief measures. Observations highlighted in green or red represent individuals for whom we can reject that wishful thinking is zero at the 95% confidence level.

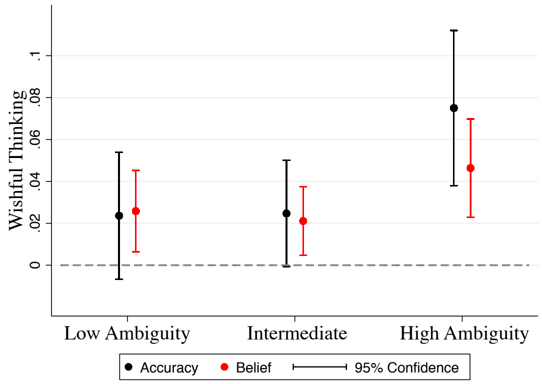
belief ($\mu = 0.031$, $s.d. = 0.076$, $d = 0.47$, $t = 3.18$, $d.F. = 59$, $p = 0.0011$, one-sided) and confirm Hypothesis 1. The effect sizes are sizable at about 40% of a standard deviation.[7]

We also analyse heterogeneity in wishful thinking. Figure 2 shows the histogram of the treatment effect, and demonstrates that about 63% of subjects shows positive wishful thinking for Accuracy, and 67% for Beliefs. The remainder shows (weakly) negative wishful thinking, or pessimism, which could be driven by disappointment aversion or a wish to brace oneself psychologically against shocks. Our within-subject design with 216 trials allows us to even make some statistical statements at the individual level. For the green data points in Figure 2, we reject the null hypothesis of no wishful thinking at the 5 percent level, using logit and OLS regressions for accuracy and belief respectively. Thus, 8 (11) out of 60 individuals exhibit significant wishful thinking in Accuracy (Beliefs). Note that by drawing on only within-subject variation, these are rather conservative tests, which are supplementary to as opposed to reflective of our the group-level analysis.

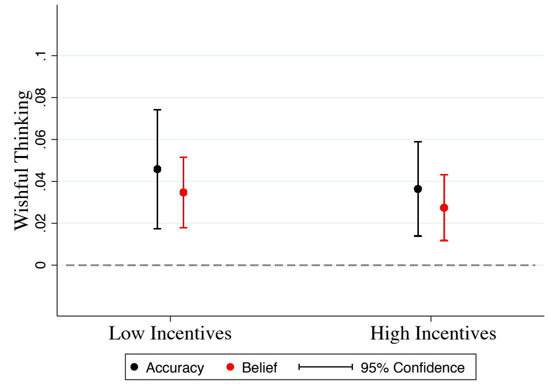
4.2 Ambiguity

To test Hypothesis 2, we contrast wishful thinking across the three levels of ambiguity. Figure 3, panel (a) shows the average level of wishful thinking across conditions for both Accuracy

[7] In the descriptive statistics, d stands for Cohen's d, defined as the difference in means of the two samples divided by the *pooled* standard deviation, which is smaller than the standard deviation of the difference.



(a) Ambiguity



(b) Incentives.

Figure 3: Wishful thinking across pattern ambiguity levels and incentive conditions. Left panel shows wishful thinking across ambiguity levels, as defined by the tilt from the vertical line. Bars indicate 95% confidence intervals.

and Belief. In addition to modest levels of wishful thinking at lower levels of ambiguity, the data show a pronounced jump at the highest level of ambiguity. More specifically, neither in terms of Accuracy nor in terms of Beliefs is wishful thinking significantly different at low and intermediate ambiguity. However, focussing on Accuracy, wishful thinking under high ambiguity is significantly higher than wishful thinking under low ambiguity (one-sided t-test, $p = 0.0113$) and wishful thinking under intermediate ambiguity (one-sided t-test, $p = 0.0099$). Similarly, focussing on Beliefs, wishful thinking under high ambiguity is higher than wishful thinking under low ambiguity (albeit with marginal statistical significance, one-sided t-test,$p = 0.0670$) or intermediate ambiguity (one-sided t-test, $p = 0.0293$).

4.3 Incentives

According to Hypothesis 3, higher incentives will reduce wishful thinking. As Panel (b) of Figure 3 shows, there are no meaningful changes in wishful thinking as a result of incentives. Differences between conditions are not statistically significant on a one-sided t-test for either Accuracy ($p = 0.333$) or Belief ($p = 0.318$).

More specifically, Hypothesis 3a) and 3b) predict that higher incentives increase belief in the correct answer in Aligned and decrease it in Non-Aligned. Table 1 shows OLS regressions of Accuracy and Belief on the shock alignment and its interaction with incentives. Column 1 and 4 replicate the main treatment effect, and also show that there is no interaction be-

	Accuracy			Belief in correct answer		
	All	Aligned	Non-Aligned	All	Aligned	Non-Aligned
	(1)	(2)	(3)	(4)	(5)	(6)
Aligned	-0.0441**			-0.0311***		
	(0.0171)			(0.0113)		
High Incentives	0.00174	0.0160	0.00174	0.00642	0.0115*	0.00642
	(0.0128)	(0.0114)	(0.0127)	(0.00792)	(0.00618)	(0.00791)
Aligned x High Inc.	0.0142			0.00510		
	(0.0170)			(0.00935)		
Constant	0.723***	0.678***	0.723***	0.653***	0.622***	0.653***
	(0.0110)	(0.0116)	(0.0110)	(0.00945)	(0.00898)	(0.00943)
Observations	240	120	120	240	120	120
R^2	0.056	0.010	0.000	0.044	0.007	0.002

Table 1: OLS regressions of Accuracy (columns 1-3) and Belief (columns 4-6) in the correct answer. Standard errors in parentheses, clustered by subject. $*p < 0.1, **p < 0.05, ***p < 0.01$.

tween the shock and incentives. Columns 2 and 5 show the data for the Aligned condition only, and show that in line with Hypotheses 3a), incentives increase performance, but not significantly so for accuracy. Columns 3 and 6 show the data for the Non-Aligned condition only. Contrary to hypothesis 3b) and in line with previous research (Pessoa and Engelmann, 2007; Engelmann, 2009), incentives increase accuracy and beliefs, but the effect is small and not statistically significant.

These results are consistent with higher incentives leading to a slight improvement in performance for both aligned and non-aligned patterns, perhaps through their effect on effort. In the Aligned condition, the incentive effect on effort and the incentive effect on wishful thinking go in the same direction, possibly accounting for the suggestive evidence of a positive effect in column 5. In the Non-Aligned condition, the two incentive effects go in opposite directions. Then, the small positive coefficients suggest that the incentive effect on effort drowns out the incentive effect on wishful thinking. We can conlude that if it exists at all, the effect of incentives on wishful thinking is rather small.[8]

[8] In an additional (unreported) analysis, we find no effect of an interaction between the level of ambiguity and the size of incentives.

5 Discussion

Here we discuss several interpretations of our results, starting with the interpretation of our main result as wishful thinking. First, one worry might be that despite several explicit understanding questions in the instructions, subjects somehow believed that their answers had a causal impact on the shock. This mistaken belief might lead subjects to switch their answers to the non-aligned pattern in order to avoid the shock. To address this point, we asked subjects in the closing questionnaire whether they agreed with the statement: "During the experiment, I believed the shocks were associated with my own answer, not with the true pattern." We find that 35 subjects (58%) "totally disagree" with this statement. Re-doing our analysis for only those participants, we find that our main effect is even stronger in this sample (Accuracy: $\mu = 0.051$, $s.d. = 0.104$, $d = 0.68$, $t = 2.87$, $d.F. = 34$, $p = 0.0035$, one-sided; Belief: $\mu = 0.041$, $s.d. = 0.085$, $d = 0.56$, $t = 2.88$, $d.F. = 34$, $p = 0.0034$, one-sided). Moreover, a regression analysis shows that the answer to the question does not predict wishful thinking. We conclude that confusion does not drive the result.[9]

A second concern might be that the more likely a subject is to receive a shock, the more noisy is her answer, reducing accuracy in the Aligned condition even in the absence of wishful thinking. This alternative, "noise-based" explanation predicts that subjects should be significantly more accurate in a condition where there is no threat of a shock at all. We can address this with data from an additional experiment ($N = 50$) that contained a No-Shock condition. In several blocks of this experiment, implemented in random order, subjects were informed that they would not receive a shock for any of the trials in this block. We find that Accuracy and Belief in the No-Shock condition are significantly lower than in the Non-Aligned condition. This is inconsistent with the noise explanation, but consistent with wishful thinking. In Appendix A we provide more details on this second experiment, and show that its main results are in line with the results of the current experiment.

Having ruled out credible confounds, our within-subject design allows us to ask whether a participant's propensity to engage in wishful thinking correlates with survey measures of anxiety and emotion regulation. However, first note that it is not clear how experienced

[9] In addition, our questionnaire featured an open question asking people "If you believed the shocks were associated with a particular event, what was the event?". 14 Subjects named some other event then the correct pattern. When excluding those 14 subjects, our main effect persists and remains highly significant (Accuracy: $\mu = 0.042$, $s.d. = 0.107$, $d = 0.58$, $t = 2.65$, $d.F. = 45$, $p = 0.0055$, one-sided; Belief: $\mu = 0.030$, $s.d. = 0.080$, $d = 0.44$ $t = 2.55$, $d.F. = 45$, $p = 0.0071$, one-sided).

	(1)	(2)	(3)	(4)	(5)
Beck Anxiety Inventory	-0.000831				-0.00170
	(0.00123)				(0.00135)
Anxious about shocks		-0.00457			0.00246
		(0.00654)			(0.00733)
Cognitive Reappraisal			-0.00283*		-0.00291*
			(0.00148)		(0.00156)
Expressive Suppression				0.00255	0.00247
				(0.00197)	(0.00215)
Constant	0.0375**	0.0345***	0.109**	-0.00905	0.0897
	(0.0162)	(0.0127)	(0.0428)	(0.0305)	(0.0603)
Observations	60	60	60	60	60
R^2	0.008	0.008	0.060	0.028	0.100

Table 2: OLS regressions of wishful thinking on items from the post-experiment questionnaire. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

anxiety should vary with wishful thinking. In terms of the model in Section 3, suppose that individuals only differed in their innate anxiety σ. Then, more anxious types would exhibit more wishful thinking. But, at the optimum, they would still experience more anxiety. Anxiety and wishful thinking would then be positively correlated. Next, suppose that individuals only differed in λ, i.e. their cognitive cost of self-deceiving. Then, those that engage in more wishful thinking will also experience less anxiety, leading to a negative correlation between wishful thinking and experienced anxiety. Thus, the model makes no clear prediction about the direction of the correlation between experienced anxiety and wishful thinking.

Our analysis is summarized in Table 2. In column 1, focusing on wishful thinking in beliefs, we find no correlation between wishful thinking and anxiety as measured by the Beck Anxiety Inventory, which indexes symptoms of clinical anxiety experienced by the individual in her everyday life (Beck et al., 1988). Column 2 shows that wishful thinking is not correlated with whether a participant stated that they were anxious about the shock in the experiment.

Next, we check for whether wishful thinking is correlated with answers to the questions of the Emotion Regulation Questionnaire (ERQ), a 10-item scale designed to measure respondents' tendency to regulate their emotions in two ways: (1) Cognitive Reappraisal and (2) Expressive Suppression (Gross and John, 2003). There is weak evidence that scoring high on the cognitive reappraisal items of the ERQ is associated with (slightly) less wishful thinking (column 3). Expressive suppression is not significantly correlated with wishful thinking (column 4). Controlling for all measures, Column 5 confirms the suggestive evidence for a correlation between wishful thinking and cognitive reappraisal. Cognitive reappraisal measures the tendency of participants to regulate their emotions by cognitively changing the meaning of emotional situations, relative to suppression. Our results suggest that cognitive reappraisal and wishful thinking are substitutes: people that commonly use cognitive reappraisal to control their emotions have less need for self-deception about the shock's occurrence. This, in turn, suggests that wishful thinking might be a separate category of emotions regulation strategies. However, we emphasize that this evidence is at most suggestive, as our result would not survive a correction for multiple comparisons across our four measures. Furthermore, only the result in column 3 is robust to using wishful thinking in terms of accuracy as the dependent variable.

Finally, we comment on the nature of motivated cognition documented in our experiment. Most of the relevant literature has focused on "motivated reasoning", a process by which people engage in the construction of self-serving arguments and biased information gathering in order to support their preferred conclusion (Kunda, 1990). Given the nature of our task, and the quick decision making, it seems implausible that such motivated reasoning is going on in our experiment. Instead, it appears that motivated cognition is instantaneous and occurs through some automatic or non-reflexive processes, adding to earlier findings of motivated perceptions and instantaneous updating (Balcetis and Dunning, 2006; Kappes and Sharot, 2019). The non-reflexive nature of wishful thinking may help explain why incentives have little effect in our setting. We conjecture that incentives are more important in situations where subjects engage more deliberatively in motivated reasoning and hence have more control over movements in their beliefs.

6 Conclusion

Philosophers and economists have long considered the importance of beliefs for people's well-being. For instance, Jevons (1879) argues that "the greatest force of feeling and motive arises from the anticipation of a long-continued future", while Bentham (1789) points to expectation as being among the most significant sources of pleasure and pain. Over the last decades, economists have introduced anticipatory feelings as a source of utility into their formal models (Loewenstein, 1987; Caplin and Leahy, 2001) and the notion of utility from anticipation has experienced somewhat of a "renaissance" (Loewenstein and Molnar, 2018).

Our experiment shows the importance of such anticipatory emotions, as the anticipatory anxiety induced by a non-noxious electric shock carries enough weight to inspire wishful thinking at a cost of lower accuracy payments. Our estimates show a large average effect of about 40% of a standard deviation that is most pronounced when the evidence is ambiguous. Wishful thinking exhibits large heterogeneity, with some subjects showing the opposite tendency. We find little evidence that a higher cost of wishful thinking reduces its prevalence.

Our results are applicable to decision making in a wide range of applications, as anticipatory anxiety has been invoked in decisions related to health, insurance, finance and politics.[10] They help explain why people seek solace and comfort in religious beliefs, why financial professionals ignore "red flags" about their asset portfolio, why many people most at risk for a particular disease are likely to avoid medical testing, and why voters who are concerned about their jobs and the future of their offspring are susceptible to reassuring political narratives. The crucial role of ambiguity gives a rationale for the avoidance of precise information such as that provided in medical tests (Lerman et al., 1998; Oster et al., 2013; Ganguly and Tasoff, 2016) and helps explain the persistence of beliefs in phenomena such as the afterlife, that by their nature do not admit clear evidence.

References

Akerlof, George A. and William T. Dickens, "The economic consequences of cognitive dissonance," *The American Economic Review*, 1982, *72* (3), 307–319.

Auriol, Emmanuelle, Julie Lassebie, Amma Panin, Eva Raiber, and Paul Seabright, "God insures those who pay? Formal insurance and religious offerings in Ghana," *Mimeo, Toulouse School of Economics*, 2017.

Balcetis, Emily and David Dunning, "See what you want to see: Motivational influences on visual perception," *Journal of Personality and Social Psychology*, 2006, *91* (4), 612–625.

[10] Examples are beliefs about health risks (Schwardmann, 2019), financial decisions and excessive trading (Brunnermeier and Parker, 2005; Eisenbach and Schmalz, 2015; Bridet and Schwardmann, 2017), time inconsistency (Caplin and Leahy, 2001; Kőszegi, 2010), occupational choice and labor market equilibrium (Akerlof and Dickens, 1982; Santos-Pinto et al., 2018), information acquisition (Yariv, 2002; Eliaz and Spiegler, 2006; Loewenstein, 2006), principal-agent communication (Kőszegi, 2006; Caplin and Leahy, 2004), self-image and taboos (Bénabou and Tirole, 2011), groupthink (Bénabou, 2013) and politics (Bénabou, 2008; Levy, 2014; Le Yaouanq, 2016).

Barron, Kai, "Belief updating: Does the'good-news, bad-news' asymmetry extend to purely financial domains?," Technical Report, WZB Discussion Paper 2016.

Beck, Aaron T, Norman Epstein, Gary Brown, and Robert A Steer, "An inventory for measuring clinical anxiety: psychometric properties.," *Journal of consulting and clinical psychology*, 1988, *56* (6), 893.

Bénabou, R and J Tirole, "Self-confidence and personal motivation," *The Quarterly Journal of Economics*, 2002, (August), 871–915.

Bénabou, Roland, "Ideology," *Journal of the European Economic Association*, 2008, *6* (May), 321–352.

_ , "Groupthink: Collective Delusions in Organizations and Markets," *The Review of Economic Studies*, 2013, *80* (2), 429–462.

_ **and Jean Tirole**, "Identity, Morals, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics*, 2011, *126* (2), 805–855.

Bentham, Jeremy, *An Introduction to the Principles of Morals and Legislation*, Oxford: Clarendon Press, 1789.

Berns, Gregory S, Jonathan Chappelow, Milos Cekic, Caroline F Zink, Giuseppe Pagnoni, and Megan E Martin-skurski, "Neurobiological Substrates of Dread," *Science*, 2006, *754* (May), 754–758.

Bridet, Luc and Peter Schwardmann, "Selling Dreams: Endogenous Optimism in Lending Markets," *Mimeo, Ludwig Maximilians Universität München*, 2017.

Brunnermeier, Markus K. and Jonathan A. Parker, "Optimal expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

Caplin, A and J Leahy, "Psychological expected utility theory and anticipatory feelings," *The Quarterly Journal of Economics*, 2001, (February 2001), 55–79.

Caplin, Andrew and John Leahy, "The supply of information by a concerned expert," *Economic Journal*, 2004, *114* (497), 487–505.

_ **and** _ , "Wishful Thinking," *NBER working paper 25707*, 2019.

Cook, John Oliver and Lehman W. Barnes, "Choice of delay of inevitable shock," *Journal of Abnormal and Social Psychology*, 1964, *68* (6), 669–672.

Coutts, Alexander, "Testing Models of Belief Bias: An Experiment," *Games and Economic Behavior*, 2019, *113*, 549–565.

Dunning, David and Emily Balcetis, "Wishful Seeing : How Preferences Shape Visual Perception," *Journal of Experimental Social Psychology*, 2013, *22* (1), 33–37.

Eisenbach, Thomas M. and Martin C. Schmalz, "Amxiety, overconfidence and excessive risk taking," *Staff Report no. 711, Federal Reserve Bank of New York*, 2015.

Eliaz, Kfir and Ran Spiegler, "Can anticipatory feelings explain anomalous choices of information sources?," *Games and Economic Behavior*, 2006, *56* (1), 87–104.

Engelmann, J. B., F. Meyer, E. Fehr, and C. C. Ruff, "Anticipatory Anxiety Disrupts Neural Valuation during Risky Choice," *Journal of Neuroscience*, 2015, *35* (7), 3085–3099.

Engelmann, Jan B., "Combined effects of attention and motivation on visual task performance: Transient and sustained motivational effects," *Frontiers in Human Neuroscience*, 2009, *3* (March), 1–17.

_ , Friederike Meyer, Christian C. Ruff, and Ernst Fehr, "The neural circuitry of affect-induced distortions of trust," *Science Advances*, 2019, *5* (3), eaau3413.

Falk, Armin and Florian Zimmermann, "Beliefs and utility: Experimental evidence on preferences for information," *Working Paper*, 2017.

Fleming, Stephen M. and Nathaniel D. Daw, "Self-evaluation of decision-making: A general bayesian framework for metacognitive computation," *Psychological Review*, 2017, *124* (1), 91–114.

Ganguly, Ananda and Joshua Tasoff, "Fantasy and Dread: The Demand for Information and the Consumption Utility of the Future," *Management Science*, 2016, (September 2017), mnsc.2016.2550.

Gross, James J and Oliver P John, "Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being.," *Journal of personality and social psychology*, 2003, *85* (2), 348.

Hollard, Guillaume, Sébastien Massoni, and Jean Christophe Vergnaud, "In search of good probability assessors: an experimental comparison of elicitation rules for confidence judgments," *Theory and Decision*, 2016, *80* (3), 363–387.

Jevons, William S., *The Theory of Political Economy*, MacmIllan and Company, 1879.

Kappes, Andreas and Tali Sharot, "The automatic nature of motivated belief updating," *Behavioural Public Policy*, 2019, *3* (1), 87–103.

Köszegi, Botond, "Emotional Agency," *The Quarterly Journal of Economics*, 2006, *121* (1), 121–155.

_ , "Utility from anticipation and personal equilibrium," *Economic Theory*, 2010, pp. 415–444.

Kunda, Ziva, "The case for motivated reasoning.," *Psychological bulletin*, 1990, *108* (3), 480.

Le Yaouanq, Yves, "A model of ideological thinking," *Mimeo, Ludwig Maximilians Universität München*, 2016, pp. 1–49.

Lebreton, Maël, Shari Langdon, Matthijs J. Slieker, Jip S. Nooitgedacht, Anna E. Goudriaan, Damiaan Denys, Ruth J. van Holst, and Judy Luigjes, "Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments," *Science Advances*, 2018, *4* (5), eaaq0668.

Lerman, Caryn, Chanita Hughes, Stephen J. Lemon, David Main, Carrie Snyder, Carolyn Durham, Steven Narod, and Henry T. Lynch, "What you don't know can hurt you: Adverse psychologic effects in members of BRCA1-linked and BRCA2-linked families who decline genetic testing," *Journal of Clinical Oncology*, 1998, *16* (5), 1650–1654.

Levy, Raphaël, "Soothing politics," *Journal of Public Economics*, 2014, *120*, 126–133.

Loewenstein, George, "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 1987, *97* (387), 666–684.

_ , "The Pleasures and Pains of Information," *Science*, 2006, *312* (May), 704–706.

— **and Andras Molnar**, "The renaissance of belief-based utility in economics," *Nature Human Behaviour*, 2018, *2* (3), 166–167.

Mayraz, Guy, "Wishful Thinking," *CEP Discussion Paper*, 2011, (1092).

Mijović-Prelec, Danica, Drazen Prelec, and Danica Mijovic-Prelec, "Self-deception as self-signalling : a model and experimental evidence," *Phil. Trans. of the Royal Society B*, jan 2010, *365* (1538), 227–240.

Mughan, A., C. Bean, and I. McAllister, "Economic globalization, job insecurity and the populist reaction," *Electoral Studies*, 2003, *22* (4), 617–633.

Obschonka, Martin, Michael Stuetzer, Peter J. Rentfrow, Neil Lee, Jeff Potter, and Samuel D. Gosling, "Fear, Populism, and the Geopolitical Landscape: The Sleeper Effect of Neurotic Personality Traits on Regional Voting Behavior in the 2016 Brexit and Trump Elections," *Social Psychological and Personality Science*, 2018, *9* (3), 285–298.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal expectations and limited medical testing: Evidence from huntington disease," *American Economic Review*, 2013, *103* (2), 804–830.

Pessoa, Luiz and Jan B. Engelmann, "Motivation sharpens exogenous spatial attention.," *Emotion*, 2007, *7* (3), 668–674.

Santos-Pinto, Luís, Michele Dell, and Luca David Opromolla, "A General Equilibrium Theory of Occupational Choice under Optimistic Expectations," *Mimo, University of Lausanne*, 2018.

Schlag, Karl H., James Tremewan, and Joël J. Van der Weele, "A Penny for Your Thoughts: A Survey of Methods for Eliciting Beliefs," *Experimental Economics*, 2015, *18* (3), 457–490.

Schmitz, Anja and Christian Grillon, "Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test)," *Nature Protocols*, 2012, *7* (3), 527–532.

Schwardmann, Peter, "Motivated health risk denial and ppreventative health care investments," *Journal of Health Economics*, 2019.

Sinding Bentzen, Jeanet, "Acts of God? Religiosity and Natural Disasters Across Subnational World Districts," *Economic Journal*, 2019, *In press.*

Sloman, Steven A., Philip M. Fernbach, and York Hagmayer, "Self-deception requires vagueness," *Cognition*, 2010, *115* (2), 268–281.

Spiegler, Ran, "On two points of view regarding revealed preference and behavioral economics," *The foundations of positive and normative economics: A handbook*, 2008, p. 95.

Trautmann, Stefan T. and Gijs van de Kuilen, "Belief Elicitation: A Horse Race among Truth Serums," *Economic Journal*, 2014, *125* (589), 2116–2135.

Yariv, Leeat, "I'll See It When I Believe It - A Simple Model of Cognitive Consistency," *Available at SSRN 300696*, 2002.

Zimmermann, Florian, "The dynamics of motivated beliefs," *American Economic Review*, 2019, *forthcoming.*

Appendices

A No-shock treatment

The second experiment ($N = 50$) was a precursor of our main experiment. The experiment had almost the same design as our main experiment and was preregistered on aspredicted.com (you can download the preregistration here). There were some small differences. First, the experiment also featured some blocks in which subjects did not face the threat of a shock. Second, while we used the same visual patterns for the task, subjects had to indicate whether they were vertically or horizontally oriented (rather than choosing the closest diagonal). Third, incentives were constant across the experiment. Finally, the experimental code exhibited a small bug which meant that the ambiguity levels were not equally calibrated across the *Aligned* and *Non-aligned* condition. While we are able to control for the ambiguity level (see below), the imperfect randomization ultimately caused us to run the main experiment, in order to derive the cleanest possible results. Having ruled out a noise-based explanation with the *No-Shock* condition, we simplified our procedures by dropping that condition in the main experiment and, instead, varied incentives in order to study their effect on wishful thinking.

For our purposes, two aspects of the second experiment are of interest. First, to investigate the "noise-based explanation" elaborated in Section 5, we look how the presence of a shock-threat affects Accuracy and Confidence. Table A.1 shows the result of an OLS regression of Accuracy and Belief, averaged by subject and condition, on treatment dummies. Because ambiguity was not well balanced between the Aligned and Non-aligned conditions, we control for ambiguity with a linear specification. We find that the Accuracy and Belief in the No Shock condition are substantially worse than in the Non-aligned condition. This shows that the presence of a shock does not in itself reduce performance, reinforcing our confidence that wishful thinking is driving our main result.

Second, we investigate whether our main treatment effect obtains also in this study. The experiment replicates our main results with very similar effect sizes (Accuracy: $\mu = 0.047$, $s.d. = 0.103$, $d = 0.55$, $t = 3.19$, $d.F. = 49$, $p = 0.0012$, one-sided; Belief: $\mu = 0.031$, $s.d. = 0.069$, $d = 0.38$, $t = 3.24$, $d.F. = 49$, $p = 0.0011$, one-sided). When we control for Ambiguity in A.1, the results remain unaltered, and show a highly significant effect of the Aligned condition which is 0.037 for Accuracy and 0.029 for Belief, closely mirroring the magnitudes in our main experiment.

	(1) Accuracy	(2) Belief
Aligned	-0.0373**	-0.0285***
	(0.0152)	(0.0103)
No Shock	-0.0339***	-0.0371***
	(0.0114)	(0.00736)
Ambiguity	-0.0349***	-0.0255***
	(0.00497)	(0.00411)
Constant	0.904***	0.810***
	(0.0132)	(0.0127)
Observations	600	600
R^2	0.098	0.087

Table A.1: OLS regressions of Accuracy and Belief on the experimental conditions. Standard errors in parentheses clustered at subject level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B Instructions

Below we provide our experimental instructions.

Instructions

Welcome to the experiment. Please read these instructions carefully and do not hesitate to contact the experimenter if you have any questions. You will receive a 10€ show up fee for coming to this experiment. You can earn additional money during the experiment, depending on your decisions. Your total earnings will be paid out at the end of the experiment.

The experiment will take about 75 minutes. During the experiment, you will be connected to a machine which can administer electric shocks to your wrist. This is an established, routine procedure and will do no harm to your health and safety. **Nevertheless, if you do not wish to participate in this experiment, please contact the experimenter now and you will be allowed to leave.**

The experiment consists of 4 parts. Shock Calibration, Practice, Experiment and Questionnaire.

Part 1: Shock Calibration

In this part we will calibrate the shock. You **cannot** earn any money during this part. You will receive a wristband on your left wrist, which is connected to the shock machine. You will then administer a series of electric stimuli to yourself, which you will be asked to rate on a pain scale level between 1 (not painful) and 10 (extremely painful). This will be repeated several times in order to calibrate your threshold for electric shocks. After the calibration is finished, please contact the experimenter.

Part 2: *Practice*

This part is intended to familiarize you with the procedures of the experiment and practice the task. During this part, (1) you **cannot** earn any money; and (2) you will **not** receive any electric stimuli. In this part you will perform a pattern recognition task 12 times. Each time, a pattern will flash on the screen, and you will be asked to indicate whether the pattern was tilted +45 degrees (/) or towards -45 degrees (\) using the arrow keys – see Figure 1 below for examples. After each pattern you will also be asked to indicate your confidence in the correctness of your own answer. You rate your confidence in percentages, where 50% means that you were not confident at all, and 100% means you were extremely confident on that trial. You will be able to do this through moving a slider on the screen using the arrow keys and then confirm your final answer by pressing the spacebar.

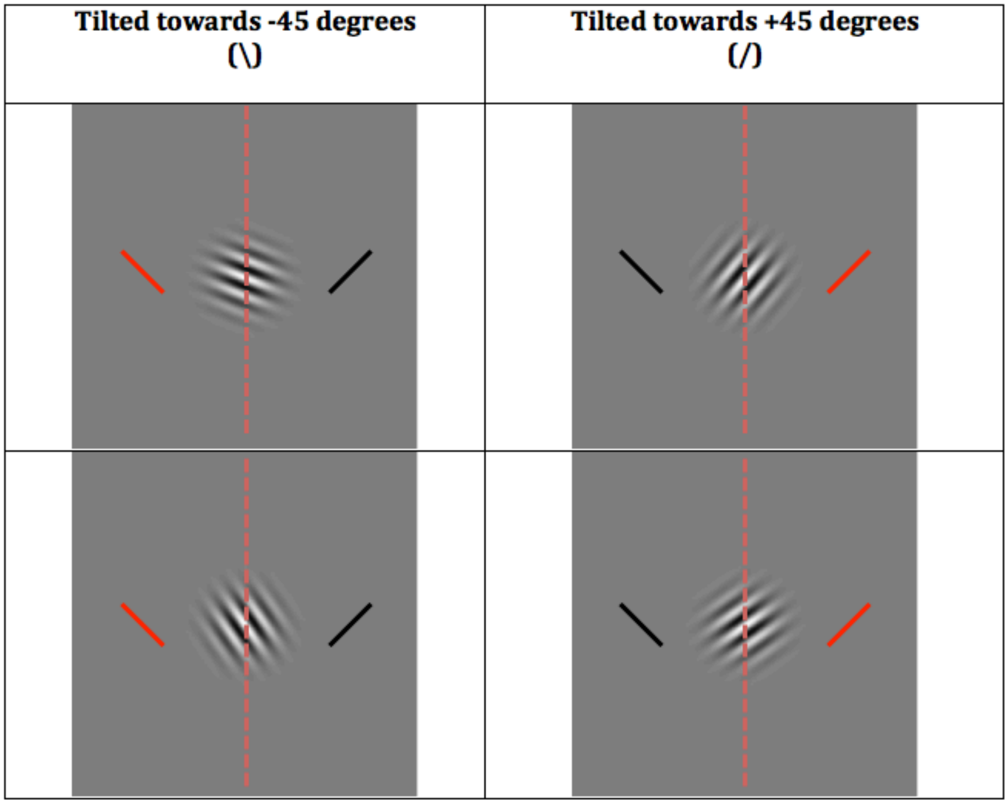


Figure 1: Examples of visual patterns. Correct answers are shown in red.

In the practice period, you will not be rewarded for your answers. In the real experiment, you will be able to win a monetary prize for accurate answers as follows. After confirming your confidence (C), the computer will draw a random lottery number (L) between 50 and 100. **If C > L**, you win a prize if your answer is correct and earn nothing otherwise**. If C < L**, the program will spin a wheel of fortune and you will win a prize with a probability of L%.

This procedure may seem complicated at first, but it is designed so that you maximize your chance of winning the prize by accurately and truthfully stating your confidence. Roughly speaking, the intuition is as follows: If you are confident your answer is correct, then you should state a high value of C, so that you are likely to get paid on the basis of your answer. If you are not very confident in your answer, then you should state a low value of C, in order to get paid on the basis of lottery L, independently of your answer.

The following two examples demonstrate this logic in more detail. First, suppose your true confidence is 75%, but instead of reporting this, you exaggerate and report C=100%. In this case, L is always lower than C, so you win the prize if your answer is correct. Since the true confidence in your answer is 75%, you can be 75% confident to win the prize. Now suppose you had instead stated your confidence truthfully as C=75% and the computer had drawn a number L>75%. In this case, you win the prize with a probability L, which is higher than 75%**.** Thus, you could have increased your chance of winning by reporting truthfully that C=75%.

As a second example, suppose again that your true confidence is 75%, but instead of reporting this, you report C=50%. In this case, L is always higher than C, so you will get paid with probability L. However, in the case that L<75%, you could have improved your chance of winning by reporting truthfully that C=75%, and getting paid on the basis of your answer.

The logic in these two examples holds for any level of confidence, and demonstrates that you maximize your chance of winning the prize if you state your true confidence.

Part 3: Experiment

This part consists of three identical sessions. Each session has 4 blocks, consisting of 18 pattern recognition tasks each. During this section, (1) you **can** earn money, and (2) you **may** receive electric shocks, as we explain now.

The task is the same as during the practice. A pattern will flash on the screen and you will be asked to indicate whether the patch was tilted more to the + 45 degree (/) or the -45 degree (\) diagonal, and to indicate your confidence (C) in the correctness of your answer. As explained above, your confidence statement may result in a prize. The size of the prize is randomly determined. For some tasks, the prize is 1€, for other tasks it is 20€. You will not observe whether you won the prize until the end of the experiment when we calculate your payments. As we explained above, you maximize your chance of winning the prize by accurately stating your true confidence.

At the start of each block a screen will appear to inform you that one of the following things is true in this block:

- You can win a **prize of 1€** for your confidence statement
- You can win a **prize of 20€** for your confidence statement

Moreover, we will tell you how the potential shock will be determined

- You may receive **an electric shock** only if the pattern tilts **towards +45 degrees (/)**
- You may receive **an electric shock** only if the pattern tilts **towards -45 degrees (\)**

Thus, the probability of receiving an electric shock only depends on the pattern, and is unrelated to your answers. For instance, if at the start of the block you are informed that you may receive an electric shock if the pattern is tilted towards +45 degrees (/), then you will receive a stimulus with a probability of 1/3 if the correct answer is (/), regardless of whether your own answer was (/) or (\).

After rating your confidence, a screen will appear stating "please wait, resolving shock" followed by either an electric shock or no electric shock.

After the completion of session 1 the visual patterns will have to be re-calibrated (as in part 1). Session 2 will commence immediately after the re-calibration. Session 2 is identical to session 1. And after the completion of session 2, the visual patterns will have to be re-calibrated again. Session 3 will commence after the session2.

Payment

After completion of session 3, **two** trials will be randomly selected for payment. We choose one trial among all the trials with a 1€ prize, and one trial among the trials with a 20€ prize. For both trials chosen for payment, you will be shown the pattern, your decision, your confidence rating, the lottery draw L and the payment outcome. The money earned in the experiment will be added to your 10€ show up fee. Thus, you can earn a minimum amount of 10€ and a maximum amount of 10+1+20=31€.

Once your final payment is displayed on the screen please contact the experimenter. We will ask you undergo a second shock calibration in order to ensure that you have calibrated it correctly at the start in order to ensure the validity of the experiment.

Part 4: Questionnaire

Before you receive payment, we will ask you to complete a questionnaire.

Confidentiality
All research data will remain completely anonymous and confidential. Any results used in scientific or other publications will be anonymous and cannot be linked to your identity.

Participation
Your participation in this study is voluntary. You may choose to withdraw from this study or decline to answer any questions at any time. If you decide to withdraw from the experiment, you will receive your show up fee.

Insurance
Participation in this study involves making simple choices where no prior knowledge is needed. The administration of electric stimuli / shock, which is a routine experimental procedure, will do no harm to your health or safety. Since this study poses no risks to your health or safety, the conditions of the regular liability insurance of the University of Amsterdam is applied.

Further information
If you have questions about this research beforehand or afterwards, please contact the responsible researcher Li-Ang Chang (e-mail: l.chang@uva.nl). In case of complaints about this study, you can contact Sophia de Jong, member of the ethical committee of the Ethics Committee Economics and Business (EBEC) of the University of Amsterdam (secbs-abs@uva.nl)

AGREEMENT

When you sign this document containing a written explanation of the experiment that you are participating in, you declare that you have read and understood the instructions and that all your questions have been answered by the experimenter. Moreover, with your signature you agree to participate in the procedures outlined in the instruction above.

[**Participant**]

"I have read and understood the information above and agree to participate in the current experiment and grant the experimenters permission to use my data. I reserve the right to withdraw from this agreement without giving any explanation, as well as to withdraw from participation in this experiment at any time."

Date:

……………………………........... ……………………………

Participant name *Signature*

[**Experimenter**]

"*I have explained the experiment to the participant. I will answer any further questions to my best knowledge.*"

Date:

……………………………........... ……………………………

Researcher name *Signature*

Quiz

1) Indicate whether the following patterns are tilted towards +45 degrees (/) or towards -45 degrees (\):

/ or \   / or \

/ or \   / or \

2) **Scenario:** You are informed that you may receive a shock if the pattern is tilted towards +45 degrees (/). The correct answer is **/** and your answer is ****, is it possible that you may receive an electric shock?
    a. Yes
    b. No

3) **Scenario:** You are informed that you may receive a shock if the pattern is tilted towards +45 degrees (/). The correct answer is **/** and your answer is **/**, is it possible that you may receive an electric shock?
    a. Yes
    b. No

4) **Scenario:** You are informed that you may receive a shock if the pattern is tilted towards -45 degrees (\). The correct answer is **** and your answer is **/**, is it possible that you may receive an electric shock?
    a. Yes
    b. No

5) **Scenario:** You are informed that you may receive a shock if the pattern is tilted towards -45 degrees (\). The correct answer is **/** and your answer is ****, is it possible that you may receive an electric shock?
    a. Yes
    b. No

6) **Scenario:** You truthfully report your confidence to be 75% and the computer draws a value L=90%. What is the probability of winning the prize if this trial is selected for payment?
    a. 75%
    b. 25%
    c. 90%
    d. 10%

7) **Scenario:** You truthfully report your confidence to be 90% and the computer draws a value L=86%. How will your payment be determined?
    a. I win the prize with 86% probability.
    b. I win the prize if my answer is correct.
    c. I win the prize with 10% probability.
    d. Impossible to tell.

8) **Scenario:** Your true confidence is 55%, but you report your confidence to be 75%. The computer draws a value L=60%. Would your chance of winning have been higher if you had reported 55% instead of 75%?
    a. Higher chance of winning by reporting 75%
    b. Higher chance of winning by reporting 55%
    c. It doesn't make a difference

9) **Scenario:** Your true confidence is 90%, but you report your confidence to be 75%. The computer draws a value L=80%. Would your chance of winning have been higher if you had reported 90% instead of 75%?
    a. Higher chance of winning by reporting 75%
    b. Higher chance of winning by reporting 90%
    c. It doesn't make a difference