informs
https://pubsonline.informs.org/journal/mnsc
MANAGEMENT SCIENCE
Articles in Advance, pp. 1–10
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Motivated Belief Updating and Rationalization of Information

Christoph Drobner,[a,b] Sebastian J. Goerg[a,c,*]

[a] TUMCS for Biotechnology and Sustainability, Technical University of Munich, 94315 Straubing, Germany; [b] Faculty of Economics, UniDistance Suisse, 3900 Brig, Switzerland; [c] TUM School of Management, Technical University of Munich, 80333 Munich, Germany
*Corresponding author
Contact: christoph.drobner@tum.de (CD); s.goerg@tum.de, https://orcid.org/0000-0002-1740-6870 (SJG)

Received: August 13, 2023
Revised: October 9, 2023
Accepted: January 3, 2024
Published Online in Articles in Advance: March 12, 2024

https://doi.org/10.1287/mnsc.2023.02537

Abstract. We study belief updating about relative performance in an ego-relevant task. Manipulating the perceived ego relevance of the task, we show that subjects substantially overweight positive information relative to negative information because they derive direct utility from holding positive beliefs. This finding provides a behavioral explanation why and how overconfidence can evolve in the presence of objective information. Moreover, we document that subjects who receive more negative information downplay the ego relevance of the task. These findings suggest that subjects use two alternative strategies to protect their ego when presented with objective information.

History: Accepted by Marie Claire Villeval, behavioral economics and decision analysis.
Funding: The authors gratefully acknowledge financial support from the ExperimenTUM.
Supplemental Material: The online appendices and data files are available at https://doi.org/10.1287/mnsc.2023.02537.

Keywords: motivated beliefs • optimistic belief updating • overconfidence • direct belief utility • Bayes' rule

1. Introduction

In standard decision theory, beliefs are unaffected by people's hopes and desires; instead, new information is processed in a Bayesian manner. This Bayesian model is difficult to reconcile with empirical evidence on overconfidence, which often leads to suboptimal decision making. Examples include excessive entry in competitive markets (Camerer and Lovallo 1999), distorted investment and merger decisions of managers and chief executive officers (CEOs) (Malmendier and Tate 2005, 2008), and polarization in politics (Ortoleva and Snowberg 2015). Furthermore, overconfident CEOs exhibit reduced responsiveness to performance feedback due to their optimistic financial outlook (Schumacher et al. 2020), whereas entrepreneurs' overconfident forecasts are associated with an increased risk of firm failures (Invernizzi et al. 2017). Although overconfident CEOs tend to explore novel technological avenues (Galasso and Simcoe 2011), there's evidence that overconfidence in product selection can lead to inferior outcomes compared with random choices (Feiler and Tong 2022). Thus, overconfidence is a common phenomenon, but a remaining puzzle is why and how overconfidence can evolve and persist in the presence of objective information.

We use a novel experimental design to provide causal evidence for the hypothesis that people overweight positive information relative to negative information because they derive direct utility from holding positive beliefs. Specifically, we study belief updating behavior in a *single event* (i.e., relative performance in an IQ test) and manipulate the perceived ego relevance of this event (i.e., how much people care about their relative performance in the IQ test). Our results show that subjects overweight positive information relative to negative information when the perceived ego relevance of the underlying event is increased. This finding provides a behavioral foundation for the persistence of overconfidence in ego-relevant settings despite the presence of objective information.

Previous experiments tested this *optimistic belief updating* hypothesis by comparing updating behavior between *different events*, which vary in their level of ego relevance (Eil and Rao 2011, Ertac 2011, Grossman and Owens 2012, Buser et al. 2018, Coutts 2019, Möbius et al. 2022, Coffman et al. 2023). For instance, Coutts (2019) compares updating behavior in beliefs about other's (ego-neutral) versus own (ego-relevant) IQ scores. Taken together, this experimental evidence has produced a variety of mixed results with evidence in favor of and against the optimistic belief updating hypothesis (Benjamin 2019, Barron 2021, Drobner 2022). One fundamental challenge of the methodology used in this literature is that *different events* vary in many dimensions, potentially confounding the causal relationship between ego relevance and belief updating. For instance, ego-relevant and ego-neutral events may differ in the size and ambiguity of prior beliefs, making it difficult to distinguish optimistic belief updating from prior-biased inference such as base-rate neglect (Barron 2021). One goal of this paper is to resolve this identification problem by introducing exogenous variation in ego relevance within a *single event* while holding other properties of the updating task fixed.

In our preregistered experiments,[1] subjects perform an IQ test, and we elicit their beliefs about the probability of scoring in the top half of the performance distribution. After the elicitation of initial beliefs, we provide subjects with different information about the importance of IQ tests. In the *High-Ego* treatment, subjects read an article containing scientific evidence arguing that IQ tests are a strong predictor for intelligence and future productivity. In the *Low-Ego* treatment, subjects read an article containing scientific evidence suggesting that IQ tests are not a valid measure for the complex phenomenon of intelligence. As a result, we argue that subjects in the *High-Ego* treatment perceive the IQ test as being more ego-relevant and consequently derive more direct belief utility than subjects in the *Low-Ego* treatment.[2] After the treatment manipulation, we provide subjects with two binary signals and elicit posterior beliefs about their relative performance in the IQ test. These signals are noisy but informative and we explicitly inform subjects that the true state of the world will not be resolved.

Overall, our results provide several important insights. First, we show that subjects update their beliefs more optimistically as direct belief utility increases. We provide several pieces of evidence in support of this finding: (i) we document more optimistic final beliefs in the *High-Ego* treatment compared with the *Low-Ego* treatment, (ii) we compare updating behavior to the Bayesian benchmark and show that subjects in the *High-Ego* treatment update their beliefs optimistically, whereas there is no such optimistic updating in the *Low-Ego* treatment, and (iii) we show evidence for motivated errors as the propensity of updates that go in the opposite direction of the Bayesian prediction increases for negative signals in the *High-Ego* treatment, whereas it is independent of the valence of signals in the *Low-Ego* treatment. Taken together, these results provide causal evidence for the optimistic belief updating hypothesis and confirm a broad range of theoretical models with direct belief utility (Bénabou and Tirole 2002, Caplin and Leahy 2019, Möbius et al. 2022).[3] Moreover, these results complement the finding of a contemporaneous project by Kozakiewicz (2022), who introduces exogenous variation in ego relevance by comparing updating behavior in response to either a realized signal or potential realizations of signals. In line with our results, Kozakiewicz (2022) documents a positive effect of direct belief utility on self-serving signal interpretations.

Second, we show that subjects alter their perceptions ex post about the ego relevance of the IQ test depending on the valence of signals received. Exploiting the noisy signal structure, we provide causal evidence that subjects consider the IQ test as being less ego-relevant, and they indicate exerting less effort in the IQ test as the number of negative signals increases. This finding complements evidence presented by Van der Weele and Siemens (2020) who find similar patterns in a self-signaling experiment, where subjects downplay the importance of doing well in a task if they receive negative performance feedback. Interestingly, we find that this ex post rationalization of information is predominantly driven by the minority of subjects with pessimistic updating patterns in the belief updating task.

2. Experimental Design

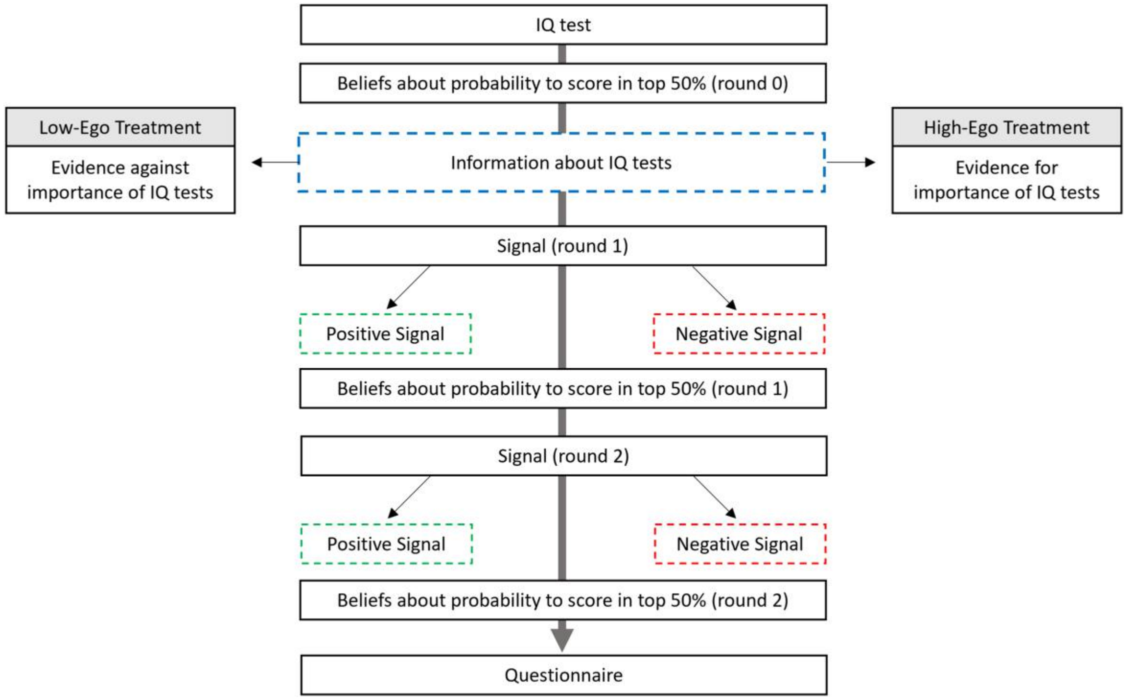
Figure 1 illustrates our experimental design. To estimate the causal effect of direct belief utility on belief updating, the experiment requires (i) a belief updating task and (ii) exogenous variation in subjects' perceived ego relevance of the underlying event. We capture these features by implementing the following experimental methodology. First, subjects performed an IQ-related test. Second, we elicited subjects' initial beliefs about their relative performance in the IQ test. Third, using a between-subjects design, we provided subjects with different information about the importance of IQ tests. Fourth, subjects received noisy but informative signals. Fifth, we elicited subjects' posterior beliefs. The last two stages were repeated such that subjects received two binary signals and reported their posterior beliefs twice.

One important aspect of the experimental design is that the treatment information was randomly assigned *after* the prior belief elicitation to rule out the possibility that other prior related errors such as base-rate neglect confound treatment differences in belief updating patterns. In addition, we explicitly informed subjects that the true state of the world remains uncertain during the course of the experiment. We implement this design feature as Drobner (2022) demonstrates that optimistic belief updating vanishes when subjects expect the immediate resolution of uncertainty. We now provide a detailed description of the different parts of the experiment.[4]

2.1. IQ Test

Subjects performed a quiz with puzzles from Civelli and Deck (2018) that are similar to the Raven Progressive Matrix test, which is commonly used as an IQ test. Subjects saw a set of 15 puzzles and had 30 seconds each to choose the correct answer from a set of four possible answers as illustrated in Figure 2. Subjects received a piece-rate payment that varied between €0.1 and €0.5 for each correct answer in the test. The size of the payments was randomly selected for each

Figure 1. (Color online) Experimental Design

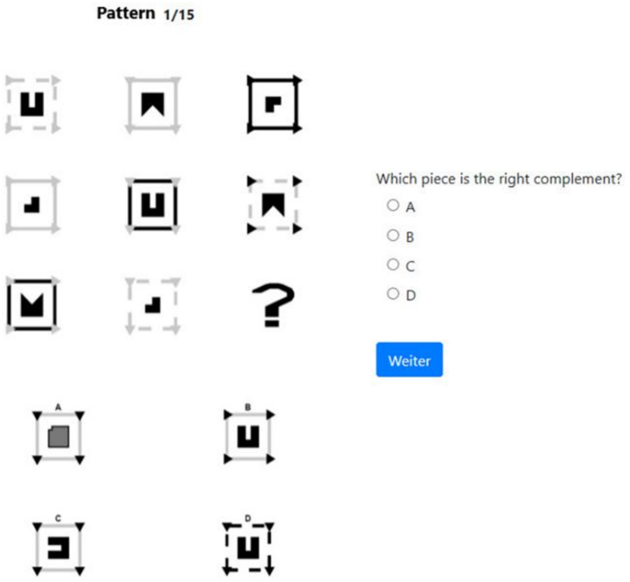


question to obfuscate the relationship between payments and IQ test performance.

2.2. Belief Elicitations

We elicited subjects' beliefs about the probability of scoring in the top half of the IQ test performance distribution in the session at three points at a time. In round 0, we elicited subjects' initial beliefs before receiving information. In round 1, we elicited subjects' beliefs after receiving the treatment information and the first binary signal about their relative performance. In round 2, we elicited subjects' beliefs after the receipt of the second binary signal about their relative performance. To incentivize truthful reporting, we implemented a variation of the Becker-DeGroot-Marschak (BDM) mechanism (Karni 2009). We asked subjects to state the probability x which makes them indifferent between winning a monetary prize of €2 with probability x and winning the same monetary prize if they indeed performed in the top half of the performance distribution within the session. This mechanism ensures that truthful reporting maximizes expected utility from payments regardless of subjects' risk preferences (Trautmann and van de Kuilen 2015).

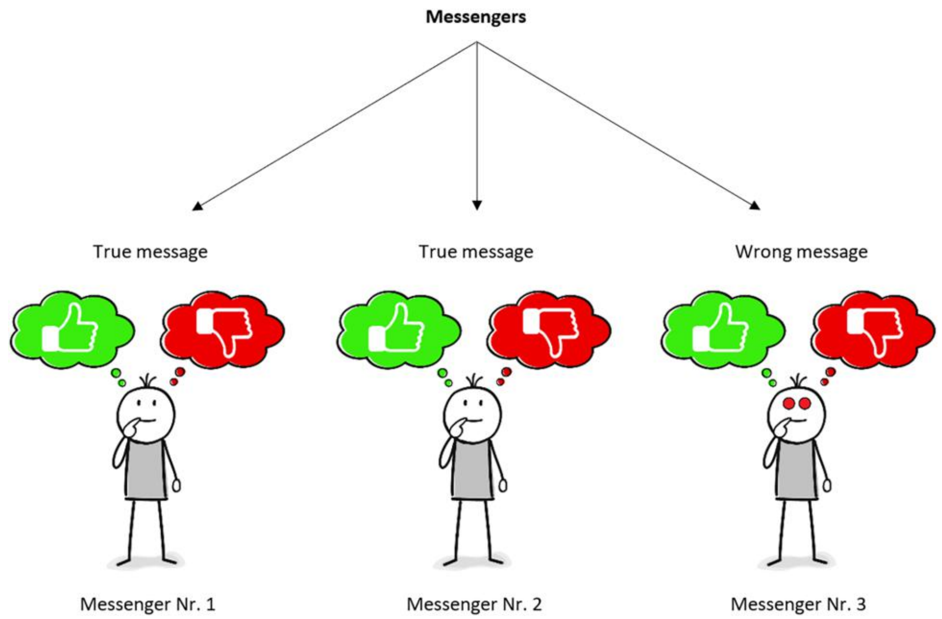
Figure 2. (Color online) IQ Test Question



2.3. Information About IQ Tests

In a between-subjects design, we asked subjects to read an article that contains simplified and shortened information summarizing scientific papers with evidence about the importance of IQ tests. Subjects in the *High-Ego* treatment received an article with scientific evidence in favor of IQ tests as predictors for success and well-being. Specifically, the article highlighted strong correlations between IQ and ego-relevant future life outcomes such as income and health. Subjects in the *Low-Ego* treatment received an article with scientific evidence against the validity of IQ tests as a measure for intelligence.

Figure 3. (Color online) Signal Generating Process

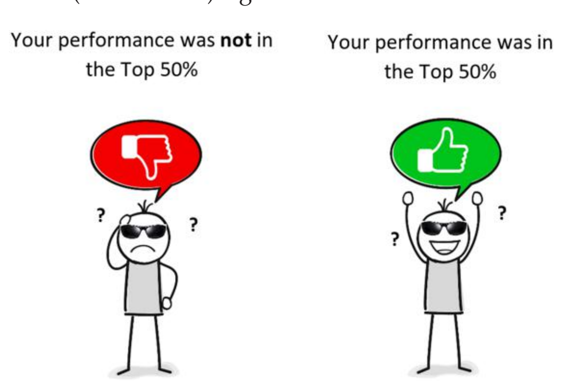


To incentivize careful reading of the articles, subjects were told that they would receive a question about the content of the article at some later stage in the experiment, providing the opportunity to win €2 if they answer the question correctly. Specifically, we asked subjects in the final questionnaire to choose the correct name of authors cited in these articles.

2.4. Signals

Subjects received two binary signals containing either positive signals or negative signals about their relative performance in the IQ test. The signals were noisy but informative with an accuracy level $q = 66.67\%$. Following Coutts (2019), subjects were told that one messenger is randomly chosen from a set of three messengers to transmit the signal as illustrated in Figure 3. Although two messengers always transmit a truthful signal, the third messenger always lies. The signal realization of both positive signals and negative signals is illustrated in Figure 4. While transmitting the signal, the messengers wore sunglasses such that individuals could not infer the reliability of the signal.

Figure 4. (Color online) Signal Realization



2.5. Questionnaire

We asked subjects to rate the importance of their performance in the IQ test for their study and job success on a seven-point Likert scale. The ratings serve as proxies for subjects' perceived ego relevance of the IQ test. In addition, we elicited subjects' self-reported effort in the IQ test on a seven-point Likert scale. We concluded the experiment with questions about the comprehensibility of the instructions and standard demographics.

2.6. Setting and Sample Size

The experiments were conducted with subjects from the laboratory for economic experiments at the Technical University Munich (ExperimenTUM) using both offline and online sessions due to the outbreak of COVID-19. We programmed the computerized experiments with the experimental software *otree* by Chen et al. (2016). Recruitment was automated using the online recruitment software ORSEE by Greiner (2015). A total of 419 subjects finished the experiment in 16 sessions (2 offline and 14 online).[5] The number of subjects in a session varied between 20 and 30.

3. Framework

In this section, we provide a stylized model of motivated beliefs in the context of our experimental setting to derive our main hypothesis. The framework follows

Engelmann et al. (2024) by modeling the benefits and costs of belief distortions as a function of direct belief utility, instrumental belief utility, and cognitive costs of belief distortions. In our experiment, subjects form beliefs about the probability of scoring in the top half of the IQ test within the session. After observing a binary signal, subjects form beliefs $\hat{\mu}$ that may deviate from objective Bayesian beliefs μ.

In our framework, subjects choose the optimal belief $\hat{\mu}$, trading off the benefits and costs of belief distortions:

$$U = \underbrace{\alpha\hat{\mu}}_{\textit{Direct belief utility}} + \underbrace{\frac{1}{2}(1 + 2\hat{\mu}\mu - \hat{\mu}^2)M}_{\textit{Instrumental belief utility}} - \underbrace{\beta(\mu - \hat{\mu})^2}_{\textit{Cognitive costs}}. \tag{1}$$

3.1. Direct Belief Utility

The first term describes that subjects derive direct utility from beliefs $\hat{\mu}$ through motives such as ego-utility (Köszegi 2006), self-esteem (Bénabou and Tirole 2002), or anticipatory utility (Brunnermeier and Parker 2005). The parameter α captures the perceived ego relevance of the underlying event.[6]

3.2. Instrumental Belief Utility

The second term describes the monetary incentives for reporting beliefs $\hat{\mu}$ under the BDM mechanism that we used in the experiment.[7] The BDM mechanism implies that subjects maximize their chance of winning a monetary price M at $\hat{\mu} = \mu$.[8]

3.3. Cognitive Costs of Belief Distortions

The third term describes that deviations of beliefs $\hat{\mu}$ from objective Bayesian beliefs μ are associated with cognitive costs of distorting reality (Bracha and Brown 2012, Coutts et al. 2024).

Maximizing Equation (1) results in the following optimal belief $\hat{\mu}$:

$$\hat{\mu} = \mu + \frac{\alpha}{M + 2\beta}. \tag{2}$$

If $\alpha = 0$, subjects form beliefs according to Bayes' rule ($\hat{\mu} = \mu$). If $\alpha > 0$, subjects derive positive direct belief utility, resulting in inflated posterior beliefs in comparison with Bayesian beliefs ($\hat{\mu} > \mu$). In our experiment, we manipulate α by providing polarizing scientific information about the importance of IQ tests in *High-Ego* and *Low-Ego* treatments, respectively ($\alpha^{High-Ego} > \alpha^{Low-Ego}$). Consequently, we expect subjects in the *High-Ego* treatment to process information more optimistically than subjects in the *Low-Ego* treatment.

4. Results

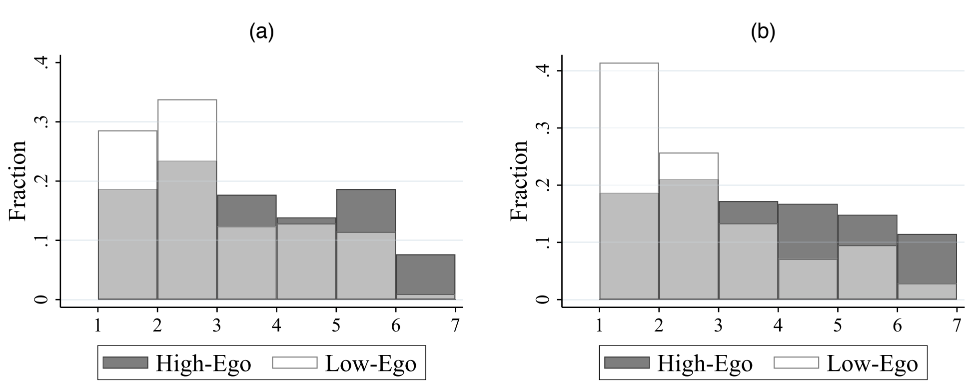
The results of our experiment are contingent on the assumption that subjects perceive the IQ test as being more ego relevant in the *High-Ego* treatment compared with the *Low-Ego* treatment. To perform a manipulation check, we compare subjects' self-reported importance of the IQ test for study and job success measured on a Likert scale (1, very low importance; 7, very high importance) between *High-Ego* and *Low-Ego* treatments.

Figure 5 illustrates the ratings for study success (a) and job success (b) separately for *High-Ego* and *Low-Ego* treatments. It shows that subjects in the *High-Ego* treatment in fact rate the importance of the IQ test higher than subjects in the *Low-Ego* treatment for both study success (Wilcoxon rank-sum test, $p < 0.001$) and job success (Wilcoxon rank-sum test, $p < 0.001$).[9]

4.1. Aggregate Beliefs

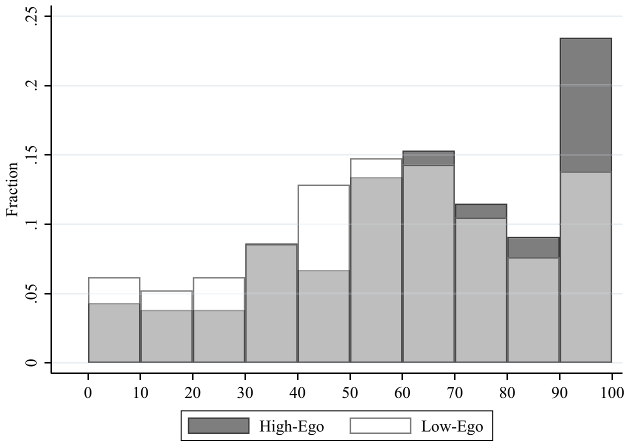
The main outcome variables from our experiment are subjects' beliefs about scoring in the top half of the IQ test within the session. Initial beliefs, measured *before* subjects received the treatment information about the importance of IQ tests, exhibit signs of overconfidence. Pooling data from both treatments, initial beliefs of being in the top half are on average 55.7% and, thus, significantly above 50% (Wilcoxon signed-rank test, $p < 0.001$). The same result holds when testing within the two treatments separately. In both treatments, initial beliefs are significantly above 50% (Wilcoxon signed-rank test, both

Figure 5. Manipulation Check



Notes. (a) Study success. (b) Job success.

Figure 6. Distributions of Final Beliefs: High-Ego vs. Low-Ego



$p < 0.05$). As expected, the distributions of initial beliefs do not differ significantly between *High-Ego* and *Low-Ego* treatments (Kolmogorov-Smirnov test, $p = 0.647$).

To study the effect of ego relevance induced direct belief utility on belief updating we compare final beliefs between *High-Ego* and *Low-Ego* treatments. Final beliefs are measured *after* subjects received the treatment information about the importance of IQ tests and the two noisy signals about their actual performance. Figure 6 depicts the distributions of final beliefs and shows that subjects in the *High-Ego* treatment form more optimistic final beliefs than subjects in the *Low-Ego* treatment (Wilcoxon rank-sum test, $p = 0.004$).

In Table 1, we quantify the average treatment effect on final beliefs, accounting for potentially confounding imbalances between treatments. Specifically, in column 1 of Table 1, we regress final beliefs on a treatment dummy (1 if High-Ego, 0 if Low-Ego), controlling for initial beliefs, gender, and IQ test scores.[10] The estimated coefficient for the treatment dummy documents that final beliefs in the *High-Ego* treatment are on average 4.81 percentage points more optimistic than final beliefs in the *Low-Ego* treatment ($p = 0.026$).

One alternative interpretation of the treatment effect on final beliefs is that the treatment induces a level shift in beliefs rather than a difference in updating behavior. This conjecture would imply that we see similar treatment differences in final beliefs independent of the signal distribution. In columns 2–4 of Table 1, we exploit the heterogeneity in signal distributions and estimate the treatment effects on final beliefs for different distributions of signals. Specifically, we run the regression analysis separately for subjects who received two negative signals, two mixed signals, or two positive signals. The results provide suggestive evidence that the treatment effect is mostly driven by the subjects who received two positive signals. This indicates that a mere level shift in beliefs that is independent of signals cannot explain the treatment effect on final beliefs.[11]

Result 1. *Initial beliefs are overconfident. Final beliefs in the High-Ego treatment are more optimistic than final beliefs in the Low-Ego treatment.*

4.2. Comparison with Bayesian Benchmark

We now extend the analysis and compare belief updating behavior to the normative benchmark of Bayes' rule using a structural empirical framework (Möbius et al. 2022). This structural framework provides several additional insights. First, in Section 4.1, we have shown that subjects in the *High-Ego* treatment form more optimistic final beliefs than subjects in the *Low-Ego* treatment, but this analysis remained agnostic about whether the belief updating process is generally optimistic or pessimistic in comparison with the Bayesian benchmark. Second,

Table 1. Final Beliefs: High-Ego vs. Low-Ego

Dependent variable: *Final Belief*	(1) Full sample	(2) Two negative signals	(3) Mixed signals	(4) Two positive signals
High-Ego	4.807**	3.667	0.563	8.074**
	(2.155)	(3.363)	(2.091)	(3.238)
Initial Belief	0.708***	0.716***	0.700***	0.572***
	(0.055)	(0.096)	(0.070)	(0.086)
Female	−2.316	2.641	−0.484	−8.936***
	(2.179)	(3.367)	(2.180)	(3.146)
IQ Test Score	1.520***	0.019	−0.238	0.203
	(0.489)	(0.896)	(0.494)	(0.791)
Constant	2.554	−8.065	22.939***	42.028***
	(4.726)	(6.762)	(5.251)	(8.985)
Observations (Subjects)	419	109	194	116
R^2	0.407	0.445	0.512	0.425

Note. Analysis uses OLS regressions with robust standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

the structural framework allows a richer description of updating behavior because we include updating in both rounds after observing each binary signal. Third, it implicitly takes initial beliefs into account and hence controls for any between-subject differences in initial beliefs. Fourth, it allows a direct comparison of subjects' responsiveness to positive and negative signals, accounting for other deviations from Bayes' rule such as conservatism or base-rate neglect.

Following Möbius et al. (2022), we use a logit transformation to derive an augmented version of Bayes' rule with indicators for positive signals $I(s_t = P)$ and negative signals $I(s_t = N)$, respectively,

$$logit(\hat{\mu}_t) = logit(\hat{\mu}_{t-1}) + I(s_t = P)log\left(\frac{q_P}{1-q_P}\right) + I(s_t = N)log\left(\frac{q_N}{1-q_N}\right). \tag{3}$$

Adding parameters δ, β_P, and β_N allows us to estimate the following empirical model, which nests Bayes' rule as a special case ($\delta = \beta_P = \beta_N = 1$):

$$logit(\hat{\mu}_{it}) = \delta logit(\hat{\mu}_{i,t-1}) + \beta_P I(s_{it} = P)log\left(\frac{q_P}{1-q_P}\right) + \beta_N I(s_{it} = N)log\left(\frac{q_N}{1-q_N}\right) + \epsilon_{it}. \tag{4}$$

The parameters β_P and β_N represent subjects' responsiveness to positive and negative signals, respectively. Conservatism implies that subjects update too little in response to both positive and negative signals ($\beta_s < 1\ \forall s \in \{P, N\}$). Optimistic belief updating is identified if subjects update their beliefs more strongly upon the receipt of positive signals compared with negative signals ($\beta_P > \beta_N$).

Table 2 shows that the estimated coefficients for subjects' responsiveness to signals are significantly below one, providing evidence for conservatism. Pooling data from both treatments shows that subjects update their beliefs more strongly on the receipt of positive signals compared with negative signals ($\beta_P > \beta_N$, $p = 0.016$). More importantly, however, this asymmetry in responsiveness to positive signals and negative signals is almost entirely driven by subjects in the *High-Ego* treatment. Although subjects in the *High-Ego* treatment are substantially more responsive to positive signals ($\beta_P^{High-Ego} > \beta_N^{High-Ego}$, $p = 0.001$), there is no such optimistic updating in the *Low-Ego* treatment ($\beta_P^{Low-Ego} > \beta_N^{Low-Ego}$, $p = 0.798$). This treatment difference in the level of optimistic belief updating is confirmed by a Chow test ($\beta_P^{High-Ego} - \beta_N^{High-Ego} > \beta_P^{Low-Ego} - \beta_N^{Low-Ego}$, $p = 0.025$).[12]

Result 2. *Subjects update their beliefs optimistically. Subjects in the High-Ego treatment update their beliefs more optimistically than subjects in the Low-Ego treatment.*

In line with previous literature, 19.8% of our subjects never update their beliefs and 10.5% update their beliefs at least once in the opposite direction that Bayes' rule would imply (Möbius et al. 2022). In the following exploratory analysis, we analyze whether these zero or wrong updates can be attributed to noise or motivated errors, that is, an extreme form of optimistic belief updating (Exley and Kessler 2024). The idea of motivated errors in our setting is that people have a higher propensity for wrong and zero updates if they *i*) receive a negative signal and *ii*) belong to the *High-Ego* treatment. In Table 3, we regress a dummy variable for zero and wrong updates on a dummy for observing a negative signal and IQ test scores. Controlling for IQ test scores, the noisy signal structure allows us to estimate the causal effect of observing a negative signal on the propensity to form zero and wrong updates in a given round. Causality is established because, conditional on

Table 2. Belief Updating

Dependent variable: *Logit Belief*	(1) Pooled	(2) High-Ego	(3) Low-Ego
δ	0.877***	0.841***	0.899***
	(0.030)	(0.055)	(0.032)
β_P	0.716***	0.796***	0.642***
	(0.048)	(0.070)	(0.067)
β_N	0.557***	0.477***	0.619***
	(0.051)	(0.073)	(0.068)
Observations	715	348	367
R^2	0.703	0.677	0.728
$\beta_P - \beta_N$	0.159	0.318	0.023
p value ($\beta_P = \beta_N$)	0.016	0.001	0.798
p value (Chow test) for ($\beta_P - \beta_N$) (Regressions 2 and 3)			0.025

Notes. Analysis uses OLS regressions with robust standard errors clustered at the individual level. Analysis includes two observations (belief updates) for each subject but excludes observations with boundary beliefs because the logit is not defined for zero or one.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 3. Motivated Errors

Dependent variable	*Zero Update*		*Wrong Update*	
	(1) High-Ego	(2) Low-Ego	(3) High-Ego	(4) Low-Ego
Negative Signal	0.279**	0.054	0.436**	−0.028
	(0.131)	(0.118)	(0.203)	(0.223)
IQ Test Score	0.050	0.010	−0.097*	0.008
	(0.036)	(0.028)	(0.054)	(0.045)
Constant	−1.030***	−0.333	−0.742	−1.424***
	(0.373)	(0.283)	(0.497)	(0.463)
Observations	395	398	284	255
Pseudo-R^2	0.011	0.001	0.067	0.000

Notes. Zero or wrong updates are dummy variables which are equal to 1 if subjects do not update in a given round or update in the wrong direction. To provide a clean comparison with correct updates, the regression analysis of zero updates excludes wrong updates and the regression analysis of wrong updates excludes zero updates. Analysis uses Probit regressions with clustered standard errors at the individual level in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table 4. Ex Post Rationalization

Dependent variable	(1) *Importance*	(2)	(3) *Effort*
	Study success	Job success	
Negative Signals	−0.306**	−0.285**	−0.266**
	(0.124)	(0.125)	(0.127)
IQ Test Score	0.094**	0.110***	0.178***
	(0.040)	(0.040)	(0.041)
Initial Belief	0.010**	0.004	0.012***
	(0.004)	(0.004)	(0.004)
High-Ego	0.679***	1.088***	0.130
	(0.177)	(0.182)	(0.178)
Observations (Subjects)	419	419	419
Pseudo-R^2	0.033	0.043	0.039

Notes. Subjects' self-reported importance of the IQ test for study and job success as well as the indicated effort are measured on a seven-point Likert scale. Analysis uses ordered logistic regressions with standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

subjects' IQ test scores, whether they observe a positive or negative signal is completely random. The results in columns 1 and 2 show that the propensity of zero updates is positively affected by observing a negative signal in the *High-Ego* treatment ($p = 0.034$), whereas it has no effect in the *Low-Ego* treatment ($p = 0.648$). Likewise, the results in columns 3 and 4 show that the propensity for wrong updates is positively affected by observing a negative signal in the *High-Ego* treatment ($p = 0.032$), whereas it has no effect in the *Low-Ego* treatment ($p = 0.899$).

Result 3. *The propensity of wrong and zero updates is increasing for negative signals in the High-Ego treatment, whereas it is independent of the valence of signals in the Low-Ego treatment.*

4.3. Ex Post Rationalization

One implicit assumption of the framework in Section 3 and the analysis thus far is that ego relevance induced direct belief utility affects the way people process information but not vice versa. We now relax this assumption and allow subjects to choose the ego relevance of the IQ test depending on what type of signals they receive (i.e., they exert some control over the shape of their direct belief utility function).

To this end, we estimate how our proxies for ego relevance, that is, subjects' self-reported importance of the IQ test for study and job success, are affected by the number of negative signals received. In addition, we also test whether subjects rationalize negative signals by indicating lower effort provision.

In columns 1 and 2 of Table 4, we regress subjects' self-reported importance of the IQ test for study and job success on the number of negative signals received and IQ test scores. Following the identification strategy in Table 3, causality is established because conditional on subjects' IQ test scores, the number of negative signals received is completely random. The results show that subjects in fact rate the importance of the IQ test for study success ($p = 0.014$) and job success ($p = 0.023$) lower as the number of negative signals increases. One interpretation of this result is that subjects ex post rationalize negative signals by downplaying the importance of the IQ test. An alternative interpretation is that people hold a certain belief about their intelligence and (rationally) update their beliefs about the reliability of the IQ test as a signal for intelligence depending on the valence of signals observed. However, the result in column 3 shows that subjects also indicate less effort provision in the IQ test when they observe more negative signals ($p = 0.036$), although we are controlling for IQ test scores. The latter result is difficult to reconcile with a rational belief updating process but rather confirms the ex post rationalization interpretation.[13]

In Online Appendix C.5, we demonstrate that ex post rationalization is stronger among subjects with pessimistic belief updating patterns (compared with Bayes) and almost vanishes for subjects with neutral or optimistic belief updating patterns. This finding suggests that ex post rationalization provides a substitute strategy for optimistic belief updating to explain away negative information. In other words, subjects have no reason to engage in optimistic belief updating if they find alternative ways to protect their ego utility.

Result 4. *Subjects rationalize negative signals about their relative performance ex post by downplaying the importance of the IQ test and pretending that they did not exert much effort in the IQ test.*

5. Conclusion

We used experiments to demonstrate the importance of ego relevance–induced direct belief utility on belief updating behavior. As opposed to a comparison of belief updating between different events with varying ego relevance, we manipulate the perceived ego relevance in a single event. This design feature allows us to study the causal effect of ego relevance on belief updating behavior while holding other properties of the updating task fixed.

Our results show that subjects update their beliefs more optimistically as direct belief utility increases. To this end, we even find evidence that subjects are more likely to update their beliefs in the opposite direction of the Bayesian prediction when they are confronted with information that negatively affects their direct belief utility. In addition, we show that subjects ex post rationalize negative information by downplaying the ego relevance of the underlying event. This ex post rationalization is more prevalent among subjects with pessimistic belief updating patterns.

From a methodological perspective, our experimental manipulation of ego relevance provides a portable paradigm to study interactions of direct belief utility with other biases in people's belief formation process. Our findings on ex post rationalization are of relevance to researchers interested in identifying motivated beliefs. For them, it is important to constrain people's ability to downplay the ego relevance of the event after receiving negative information because it undermines their motive for self-serving biases in belief formation.

From a practical perspective, our documented biases in information processing might adversely affect managerial decision making. The negative impact of overconfidence on decision making has been widely documented in the literature, and our results causally link higher overconfidence to settings with increased ego relevance. Our findings suggest that decision environments designed to downplay the ego relevance of a decision might result in less biased decision making. An alternative strategy worth exploring involves implementing data-driven decision support systems, which should not be prone to people's ego-protecting biases in the processing of information.

Acknowledgments

This experiment was pre-registered at the American Economic Association's registry for randomized controlled trials (AEARCTR-0005121) and received Institutional Review Board approval from the German Association for Experimental Economic Research. The authors thank Kai Barron, Boon-Han Koh, Yves Le Yaouanq, Tanya Rosenblat and seminar participants at the Economic Science Association Meetings, International Association for Research in Economic Psychology and the Society for the Advancement of Behavioral Economics Conference, Bavarian Graduate Program in Economics Workshop, Subjective Probability, Utility, and Decision Making Conference, German Association for Experimental Economic Research Conference, Young Economists Meeting, Workshop on Beliefs, Narratives and Memory, Matterhorn Symposium and Verein für Socialpolitic Conference for helpful comments and suggestions.

Endnotes

[1] Table A.1 in Online Appendix A maps the preanalysis plan to our paper.

[2] Direct belief utility describes a hedonic value of holding a particular belief such as deriving ego utility (Köszegi 2006) or anticipatory utility (Brunnermeier and Parker 2005) from holding positive beliefs. To this end, direct belief utility is distinct from belief utility in the Bayesian model, which is purely instrumental and indirectly derived by making the best possible decision based on accurate beliefs.

[3] Other behavioral predictions of this type of models include motivated memory (Zimmermann 2020) and motivated information avoidance (Golman et al. 2017). In this paper, we focus on optimistic belief updating in the short run but the intuition of our results also applies to these related behavioral mechanisms.

[4] Full experimental instructions are provided in Online Appendix D.

[5] Overall, 451 subjects participated, but 32 students dropped out during the online experiments.

[6] For simplicity, we assume linearity in direct and instrumental belief utility because we only need monotonicity but not risk neutrality or other properties of von Neumann-Morgenstern utility theory to derive our main hypothesis.

[7] The formula is derived from Engelmann et al. (2024) and Hill (2017).

[8] For simplicity, we abstract from the fact that individuals may derive instrumental utility from the beliefs they hold about their relative IQ from decisions they make outside the laboratory because it does not affect the qualitative predictions of the framework.

[9] The responses to the self-reported importance of the IQ test might be prone to experimenter demand effects. However, in Online Appendix B, we discuss why experimenter demand effects are only of minor concern for the main results of our experiment.

[10] Online Appendix C.1 shows that initial beliefs, gender, and IQ test scores do not differ significantly between treatments.

[11] In Online Appendix C.3, we expand this discussion in the context of the structural framework used in the following section.

[12] In Online Appendix C.2, we provide several robustness checks. First, we replicate the analysis with restricted samples excluding subjects who do not update their beliefs in the direction of the Bayesian prediction. Second, we smooth boundary priors to run the regression analysis including the most optimistic and pessimistic subjects in the sample. Third, we address the potential endogeneity concern that arises when belief updating systematically differs between subjects ranked in the top half or the bottom half of the IQ test (Barron 2021). The robustness checks confirm the results in Table 2.

[13] Table 4 shows the regression analysis for the pooled data from both treatments. In Online Appendix C.4, we run the regressions separately for *High-Ego* and *Low-Ego* treatments. The corresponding results indicate some differences in the magnitude of ex post rationalization, which are, however, not statistically significant at any conventional level.

References

Barron K (2021) Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Experiment. Econom.* 24(1):31–58.

Bénabou R, Tirole J (2002) Self-confidence and personal motivation. *Quart. J. Econom.* 117:871–915.

Benjamin DJ (2019) Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Foundations and Applications 2*, vol. 2 (Elsevier, New York), 69–186.

Bracha A, Brown DJ (2012) Affective decision making: A theory of optimism bias. *Games Econom. Behav.* 75:67–80.

Brunnermeier MK, Parker JA (2005) Optimal expectations. *Amer. Econom. Rev.* 95:1092–1118.

Buser T, Gerhards L, van der Weele J (2018) Responsiveness to feedback as a personal trait. *J. Risk Uncertainty* 56:165–192.

Camerer C, Lovallo D (1999) Overconfidence and excess entry: An experimental approach. *Amer. Econom. Rev.* 89:306–318.

Caplin A, Leahy JV (2019) Wishful thinking. NBER Working Paper No. 25707, National Bureau of Economic Research, Cambridge, MA.

Chen DL, Schonger M, Wickens C (2016) oTree—An open-source platform for laboratory, online, and field experiments. *J. Behav. Experiment. Finance* 9:88–97.

Civelli A, Deck C (2018) A flexible and customizable method for assessing cognitive abilities. *Rev. Behav. Econom.* 5:123–147.

Coffman K, Collis MR, Kulkarni L (2023) Stereotypes and belief updating. *J. Eur. Econom. Assoc.*, ePub ahead of print October 30, https://doi.org/10.1093/jeea/jvad063.

Coutts A (2019) Good news and bad news are still news: Experimental evidence on belief updating. *Experiment. Econom.* 109:21.

Coutts A, Gerhards L, Murad Z (2024) What to blame? Self-serving attribution bias with multi-dimensional uncertainty. *Econom. J.* Forthcoming.

Drobner C (2022) Motivated beliefs and anticipation of uncertainty resolution. *Amer. Econom. Rev. Insights* 4:89–105.

Eil D, Rao JM (2011) The good news-bad news effect: Asymmetric processing of objective information about yourself. *Amer. Econom. J. Microeconom.* 3:114–138.

Engelmann J, Lebreton M, Salem-Garcia N, Schwardmann P, van der Weele JJ, Chang LA (2024) Anticipatory anxiety and wishful thinking. *Amer. Econom. Rev.* Forthcoming.

Ertac S (2011) Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *J. Econom. Behav. Organ.* 80:532–545.

Exley CL, Kessler JB (2024) Motivated errors. *Amer. Econom. Rev.* Forthcoming.

Feiler D, Tong J (2022) From noise to bias: Overconfidence in new product forecasting. *Management Sci.* 68:4685–4702.

Galasso A, Simcoe TS (2011) CEO overconfidence and innovation. *Management Sci.* 57:1469–1484.

Golman R, Hagmann D, Loewenstein G (2017) Information avoidance. *J. Econom. Literature* 55:96–135.

Greiner B (2015) Subject pool recruitment procedures: Organizing experiments with ORSEE. *J. Econom. Sci. Assoc.* 1:114–125.

Grossman Z, Owens D (2012) An unlucky feeling: Overconfidence and noisy feedback. *J. Econom. Behav. Organ.* 84:510–524.

Hill SJ (2017) Learning together slowly: Bayesian learning about political facts. *J. Politics* 79:1403–1418.

Invernizzi AC, Menozzi A, Passarani DA, Patton D, Viglia G (2017) Entrepreneurial overconfidence and its impact upon performance. *Internat. Small Bus. J.* 35:709–728.

Karni E (2009) A mechanism for eliciting probabilities. *Econometrica* 77:603–606.

Köszegi B (2006) Ego utility, overconfidence, and task choice. *J. Eur. Econom. Assoc.* 4:673–707.

Kozakiewicz M (2022) Belief-based utility and signal interpretation. Working paper, University of Bonn, Bonn, Germany.

Malmendier U, Tate G (2005) CEO overconfidence and corporate investment. *J. Finance* 60:2661–2700.

Malmendier U, Tate G (2008) Who makes acquisitions? CEO overconfidence and the market's reaction. *J. Financial Econom.* 89:20–43.

Möbius MM, Niederle M, Niehaus P, Rosenblat TS (2022) Managing self-confidence: Theory and experimental evidence. *Management Sci.* 68:7793–7817.

Ortoleva P, Snowberg E (2015) Overconfidence in political behavior. *Amer. Econom. Rev.* 105:504–535.

Schumacher C, Keck S, Tang W (2020) Biased interpretation of performance feedback: The role of CEO overconfidence. *Strategic Management J.* 41:1139–1165.

Trautmann ST, van de Kuilen G (2015) Belief elicitation: A horse race among truth serums. *Econom. J. (London)* 125:2116–2135.

Van der Weele JJ, Siemens FA (2020) Bracelets of pride and guilt? An experimental test of self-signaling. *J. Econon. Behav. Organ.* 172:280–291.

Zimmermann F (2020) The dynamics of motivated beliefs. *Amer. Econom. Rev.* 110:337–361.