Title

An Unlucky Feeling: Persistent Overestimation of Absolute Performance with Noisy Feedback

Authors

Grossman, Zachary
Owens, David

Publication Date

2011-04-11

An Unlucky Feeling: Persistent Overestimation of Absolute Performance with Noisy Feedback ☆

Zachary Grossman[a,], David Owens[b]

[a]Department of Economics, 2127 North Hall, University of California, Santa Barbara, CA 93106
[b]Haverford College, Department of Economics, 370 Lancaster Ave, Haverford, PA 19041

Abstract

How does overconfidence arise and persist in the face of experience and feedback? We examine experimentally how individuals' beliefs about their absolute, as opposed to relative, performance on a quiz react to noisy, but unbiased, feedback. Participants believe themselves to have received 'unlucky' feedback and they overestimate their own scores, but they exhibit no overconfidence in non-ego-relevant beliefs—in this case, about others' scores. Unlike previous studies of relative performance estimates, we find this to be driven by overconfident priors, as opposed to biased updating, which suggests that social comparisons contribute to biased information processing. While feedback improves performance estimates, this learning does not translate into improved estimates of subsequent performances. This suggests that people use performance feedback to update their beliefs about their ability differently than they do to update their beliefs about their performance, contributing to the persistence of overconfidence.

Keywords: overconfidence, feedback, overestimation, absolute performance, Bayesian updating, biased updating, information processing, learning transfer, cross-game learning

JEL: C91, D03, D83

☆We are grateful to Gary Charness, Muriel Niederle, Paul Niehaus, and Laura Gee, as well as seminar participants at Swarthmore College, University of Delaware, UC San Diego, and the ESA North American meetings for helpful comments and to Daniel Saunders for excellent research assistance. This research was partially funded by an SSRGP grant from the Institute for Social, Behavioral, and Economic Research (ISBER) at the University of California, Santa Barbara.

Email addresses: `grossman@econ.ucsb.edu` (Zachary Grossman), `dowens@haverford.edu` (David Owens)

1. Introduction

Life experience consistently provides information about our personal attributes and our performance in a variety of settings. A restaurateur's experience provides her with information not only about the likelihood that her restaurant will succeed, but also about her ability to successfully manage and operate a restaurant business. A Bayesian's beliefs about her own skill would grow progressively more accurate and she would use this information to inform future decisions about whether and how to undertake risky ventures.

However, psychologists and economists have found that people express overconfident estimates of their relative ability or performance across a broad across a broad range of measurements (Hoelzl and Rustichini, 2005; Burks et al., 2010), even to their own financial detriment (Malmendier and Tate, 2008; Camerer and Lovallo, 1999). This include attributes about which people receive a multitude of feedback over the course of their lives. In a commonly cited example, Svenson (1981) finds that 93% of U.S. students believe that they are better than average drivers. Most people drive on a regular basis, and each trip is an opportunity to learn about one's driving ability.

Moore and Healy (2008) and Benoit and Dubra (2007) argue that much of the evidence of overconfidence is perfectly compatible with Bayesian updating and might merely reflect a statistical bias, as opposed to a behavioral bias. However, Burks et al. (2010) reject Bayesian updating and Moebius et al. (2007) and Eil and Rao (2010) find evidence beliefs are both relatively unresponsive to information, and react to information with an ego-preserving bias. The question of whether expressed overconfidence about relative performance reflects a behavioral bias or merely a statistical bias might be resolved through a careful study of relative performance estimates. However, examining how people estimate *absolute* performance may also clarify the severity of any bias in the updating process, while shedding light on the role of social comparisons in driving the updating anomalies found by Moebius et al. (2007) and Eil and Rao (2010).

Furthermore, while in some realms, for example the competitive marketplace, relative performance is the key determinant of success, many risky decisions rely upon absolute performance as well. For example, a mortgage lender's bottom line depends upon her being a good judge of borrows' probability of default, independent of the judgments of other lenders. Thus, investigating the nature and severity of overconfidence in absolute performance measures can shed insight into how overconfidence matters in a broader range of environments including entrepreneurial and investor choice.

In this paper we examine how overconfidence about absolute performance arises and how it persists in the face of experience and feedback. Our analysis is based upon the recognition that in any task-feedback situation there are two possible sources of overconfidence in posterior beliefs: overconfident priors and biased information processing. We present an experiment that allows us to distinguish these two sources of overconfidence by examining how individuals' prior beliefs about their absolute (as opposed to relative) performance on a quiz react to noisy, but unbiased feedback.

Posterior beliefs emerging from one situation form priors for the next. Hence, informative feedback about one performance should increase the accuracy of *prior* beliefs about the *next* performance. Over time, this conflates the effects of overconfident priors versus biased information processing. To gain insight into how overconfidence may persist over a series of tasks, we examine how experience and feedback from the quiz affect subsequent prior and posterior beliefs about performance in a second, very similar quiz.

In our experiment, participants take a ten-question quiz on logic and reasoning. Four main features of the design, presented in Section 2, distinguish our experiment from others investigating overconfidence and the reaction to feedback. First, in two separate conditions, participants estimate a different target score: their own score in the *Self* condition and the score of another participant in the *Other* condition. The *Other* condition serves as a control because, while the quantity about which participants must form and state beliefs is not ego-relevant, it presents an otherwise identical task. Second, participants estimate absolute, as opposed to relative performance. Third, the nature of the feedback we give—a perturbation of the raw score, as opposed to a garbled binary signal—allows us to examine quantitatively how individuals attribute performance to external circumstances or luck. Finally, repeating the quiz/belief-elicitation protocol allows us to examine how information and experience carry over from one task or situation to the next.

Section 3 presents our results. We find that, while, on average our participants express accurate interim and posterior beliefs about others' scores, but they overestimate their own score and believe themselves to have received 'unlucky' feedback. However, this effect is largely consistent with Bayesian updating and overconfident prior beliefs, as opposed to biased information processing. We find no evidence that participants' beliefs respond conservatively to feedback about their performance. Rather, their beliefs react slightly *more* than warranted by Bayes' rule. Similarly, most participants' beliefs react fairly symmetrically to 'good news' and 'bad news'. Thus, post-feedback overconfidence is primarily driven by overconfident priors. While experience and feedback from the first quiz improve the accuracy with which participants estimate others' scores on the second quiz, we find only minimal improvement in the accuracy of participants' estimates of their own scores. This suggests that people use performance feedback to update their beliefs about their ability differently than they do to update their beliefs about their performance, which may contribute to the persistence of overconfidence.

We conclude with a discussion of these results in Section 4. Our results are consistent with recent research which has found differences in overconfidence and information processing for information that is and is not ego-relevant.[1] Charness et al. (2010) find that people make more errors in estimating their own performance on a mental ability task than on another in which feedback relates to a more abstract question. Ertac (2009) finds that whether or not information is ego-relevant affects the way it is processed.

Recent work has also found an asymmetric response to news—participants respond much less to negative information about themselves than positive information. Eil and Rao (2010) ask subjects to estimate their ranking in both an IQ task and in a measure of individual attractiveness both before and after receiving binary feedback and find similar results, labeling the asymmetric response the "good news-bad news" effect. We find limited evidence of such a bias in our results, and that the bias is largely concentrated in a small number of subjects with highly inaccurate prior beliefs. Hence, asymmetric responses to feedback found in other studies, which all elicit beliefs about *relative* performance, may depend upon mis-attribution in social comparisons and that people's response to feedback depends upon whether the performance measure is relative or absolute. We cannot reject the notion that judgments reflecting overconfidence may be compatible with Bayesian updating, as pointed out by Moore and Healy (2008) and Benoit and Dubra (2007).

Previous studies clearly establish the role of biased processing of performance feedback on generating overestimation in posterior beliefs about performance and this phenomenon is likely at play in many general settings. However, the simple and uniform nature of the noise-generating in

[1]Not all research has found exclusive overconfidence and overoptimism. Ertac (2009) finds evidence of pessimism in interpreting feedback that is incomplete (as opposed to noisy) and Clark and Friesen (2009) find zero mean error and *underconfidence* to be more prevalent in forecasts of both relative and absolute performance on a real effort task. They also find that people have less difficulty predicting relative performance than predicting absolute performance. Moore and Healy (2008) also find subjects to be under-confident, in both a relative and absolute sense, on difficult tasks.

our experiment may help our participants conform to Bayes' rule to a greater degree than in the aforementioned studies. In fact, our participants slightly *over-reacted* to the unbiased feedback, which—given their overconfident priors—might be expected to sharply curb posterior overconfidence. Though, like Clark and Friesen (2009), we find that updating is not sufficient to eliminate forecast errors, participants *do* seem to learn well about their quiz performance from the feedback as the degree of overestimation sharply drops after the feedback.

However, we do find some evidence of a bias specific to ego-relevant judgments when it comes to cross-game learning (Kagel, 1995; Cooper and Kagel, 2003, 2008, 2009). Overestimation disappears on average on the second quiz among those estimating others' scores, but among those estimating their own score, overestimation in prior beliefs barely drops from the first quiz to the second, despite the fact that participants in the second quiz have the experience and feedback from the first quiz to draw upon. Furthermore, the posterior mean-square error drops less from quiz to quiz, both in absolute terms and when compared to the error in Bayesian posteriors, when estimating one's own score. Thus, the transfer of learning from one quiz into the next appears stronger when one's own ego is not at stake.

This suggests another channel through which overconfidence may persist. Performance feedback does improve the accuracy of these participants' beliefs about the relevant performance, but not necessarily about the underlying abilities that generate similar performances. Thus, not only do people process information about themselves differently than they do information about others, they use performance feedback to update their beliefs about their ability differently than they do to update their beliefs about their performance.

2. Experimental Design

Experimental sessions took place at the Experimental and Behavioral Economics Laboratory (EBEL) at the University of California, Santa Barbara. Participants were randomly recruited from the EBEL subject pool (largely comprised of UCSB students and staff) using the online system ORSEE (Greiner, 2003). Full instructions for all conditions, as well as screenshots are presented in the Appendix.[2]

[2] The software are available from the authors upon request.

Twelve sessions were conducted, six for each condition. All sessions were conducted in pairs featuring the same two quizzes, alternating the order of the *Self* and *Other* condition in each pair. The order of the quizzes was also reversed in consecutive session pairs.[3] Each session lasted roughly 75 minutes and, with one exception, included 11 participants.[4] Participants earned $12.65 on average, which includes a $5 show-up fee and which was paid privately in cash at the end of the session. Each participant took part in only one session.

Upon arriving at the experiment, participants sat at computer terminals, were given a paper copy of the instructions, and followed along as the experimenter read them aloud. Participants then completed a protocol in which they took a quiz and were asked three times to state their beliefs about performance on the quiz. In two separate conditions participants were asked to estimate different target scores: participants in the *Self* condition were asked about their *own* score, while participants in the *Other* condition we asked about the score of an anonymous randomly selected other participant (RSOP). Beliefs were elicited three times: 1) *prior* beliefs (b_1), before participants took the quiz; 2) *interim* beliefs (b_2), immediately after taking the quiz; and 3) *posterior* beliefs (b_3), after receiving feedback about the performance being estimated.

After completing this process, participants repeated it with a second quiz, which featured new questions that were drawn from the same source as the first quiz. When expressing their beliefs, participants did not view beliefs that they had expressed previously. Further, they did not view payoffs for or scores on the *first* quiz until they had completed the second. The key distinction between the first and second round of quizzes is that participants in the second round were *experienced*, having already experienced taking the first quiz and receiving noisy feedback about their first-round target score.

2.1. Quizzes

The quizzes consisted of ten multiple-choice questions selected from a book of Mensa quizzes (Grosswirth et al., 1999). They were presented in a Microsoft Excel spreadsheet which allowed participants to select one answer from a menu of possible answers. We constructed a total of four

[3] One block of four sessions featured an error that led some participants to be given an incorrect score for one of the quizzes. As a result we omit from the data analysis all 39 observations for which the target score was miscalculated and 20 observations for a quiz following one in which the participant's target score was miscalculated.

[4] Session 9, which was in the *Self* condition, had only 9 participants.

quizzes, two of which were used in each session.[5] To incentivize effort, *for each quiz* we selected one question at random at the end of the session. Participants earned an extra $5 if their answer to this question were correct. Participants were paid a $5 show-up fee, plus the incentive payment for each quiz, plus the earnings from *one* of the three beliefs elicitations for each quiz. The elicitation was randomly-selected.

2.2. Belief Elicitation

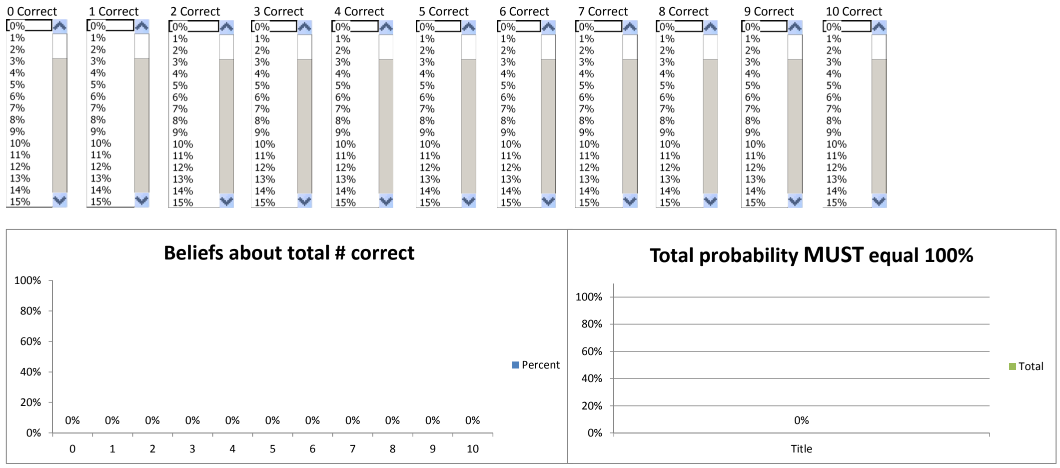
Rather than simply asking participants to guess how many questions they answered correctly, we elicited the entire distribution of their beliefs, using an interface and incentive scheme similar to those used by Moore and Healy (2008) and Eil and Rao (2010). Figure 1 presents a screenshot of our computer interface, which was programmed using Microsoft Excel. For each raw-score outcome from zero to ten correct answers, participants select the likelihood (in percentage terms) with which they believe that outcome to have occurred. The percentages initially were set to zero for each outcome and participants made adjustments by selecting numbers from the drop-down menu corresponding to each outcome.

Two charts on the same screen were provided as a visual aids to the participants. An auto-adjusting histogram provided a visual profile of the current selection and a single bar showed the running total of percentage-points selected. The experimenter verified that the total was 100% before the participant was allowed to save the estimate and close the program. We used the same interface for each elicitation, but participants did not have access to their history of belief estimates in the form of previous files.[6]

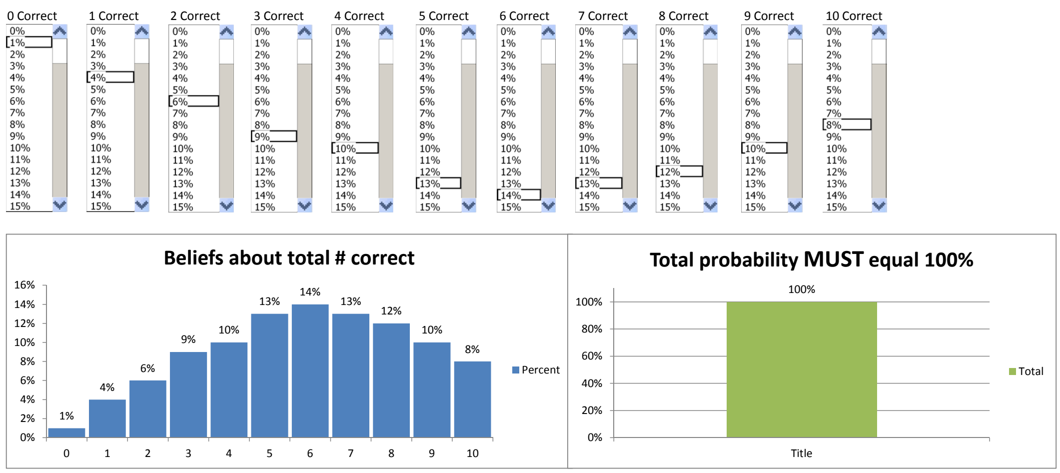
Participants' earnings for their probability estimates were determined by a quadratic scoring rule and for each quiz, one of the three estimates was selected for payment. Specifically, for each paid elicitation, a participant was earned $1 + 5p_x - w$ dollars, where p_x is the participant's estimated likelihood for the true outcome and w is the sum of squares of the likelihoods for each of the eleven possible outcomes (zero correct through ten correct). Following Moore and Healy (2008) and Eil and Rao (2010), after learning the scoring-rule formula, participants were told, "This formula may appear complicated, but what it means for you is very simple: You get paid the most when you

[5] All quizzes, plus correct answers are provided in the Appendix.

[6] Participants were allowed to write on scrap paper and may have had access to their previous estimates had they recorded the estimates on the paper.



(a) The initial appearance



(b) With entered beliefs

Figure 1: The beliefs-elicitation interface

honestly report your best guesses about the likelihood of each of the different possible outcomes." Furthermore, the instructions walked participants through four detailed examples of how the payoff for the beliefs-elicitation was calculated.[7]

[7]We submit that few participants understand the quadratic scoring rule computationally, but feel that they generally trust the experimenter when told that honestly reporting their beliefs is the best thing to do.

2.3. Feedback

For each quiz, after all participants had entered their second score estimate, each received a slip of paper, face down for privacy, with a number on it that constituted his or her feedback, f.[8] This number was generated by perturbing the target score, s^T,[9] with a noise or 'luck' term, ℓ, drawn uniformly from then set $\{-2, -1, 0, 1, 2\}$. Participants were instructed as such and the instructions emphasized the fact that the number on the paper was equally likely to be equal to $s^T - 2$, $s^T - 1$, s^T, $s^T + 1$, or $s^T + 2$.

After receiving this feedback participants entered their beliefs a third and final time. After completing this step for the first quiz, participants moved on to the second part of the session, in which they took a second quiz and entered their beliefs another three times. After completing this step for the second quiz, participants were given a statement summarizing their payment and their scores on the two quizzes, received their payment and exited.

After the elicitation of the prior beliefs distribution, b_1, participants viewed and completed the quiz, an experience that informs interim beliefs, b_2. After expressing b_2, participants observed f and expressed posterior beliefs, b_3. Bayes' rule provides no benchmark for updating from b_1 to b_2, as we do not observe the information provided by viewing and taking the quiz. It does, however, provide a specific benchmark for updating from b_2 to b_3, shown in equation 1, in which b_j^i is the expressed probability that $s^T = i$ in belief elicitation j.

$$b_3^{i*} = \begin{cases} \frac{b_2^i}{\sum_{k=f-2}^{f+2} b_2^k} & \text{if} \quad f-2 \leq i \leq f+2 \\ 0 & \text{if} \quad i < f-2 \quad \text{or} \quad i > f+2 \end{cases} \tag{1}$$

Because $|\ell| \leq 2$, b_3^{i*} is restricted to be zero outside the range $[f-2, f+2]$. Furthermore, scores within $[f-2, f+2]$ are expected in the same proportions as in the expressed interim beliefs. In particular, because of the uniformity of the noise, interim beliefs are not 'dragged' towards the feedback, f. For example, consider a participant with $b_2^6 = b_2^7 = 0.50$, who receives feedback $f = 5$. Such f could be labeled 'bad news', in the sense that the (unbiased) feedback is lower than she expected. However, her Bayesian posterior remains unchanged from b_2, $b_3^{6*} = b_3^{7*} = 0.50$. Bayesian updating results in a truncation of b_2, and no shifting of probabilities among $s^T \in [f-2, f+2]$.

[8]In the instructions, we used neutral language and avoided using the word 'feedback'. Instead we called it 'imperfect information' about the score the participant was trying to estimate.

[9]The variable s is defined as a participant's own score. In the *Self* condition, $s^T = s$.

Relative to other noise-generating processes, Bayes' rule predicts relatively mild changes in beliefs in response to feedback.

When $b_2^i = 0$ for all $i \in [f-2, f+2]$, Bayes' rule provides no benchmark for b_3. This is an important departure from previous studies, such as Moebius et al. (2007) and Eil and Rao (2010). In these designs, feedback comes in the form of a binary signal that is more likely to be correct than wrong, so all beliefs generate a Bayesian posterior. In our design, feedback is never completely wrong, in that there is a limit to how much it can differ from s^T.

3. Results

We begin the discussion of our results with a brief analysis of participants' performance on the quiz. Table 1 summarizes scores and feedback by condition, quiz order, and the exact quiz employed.[10] We present results at the individual-quiz level. Thus, each participant is represented by two separate sets of observations.

Participants answer roughly half of the ten questions correctly, on average, and those in the *Other* treatment perform marginally better, though this difference does not approach statistical significance ($t = 1.20$, $p = 0.23$). Performance improves, in both conditions, from the first quiz to the second, and there is some variance in the difficulty of the four quizzes.[11] By construction, feedback in both conditions is unbiased—on average neither an upward nor a downward perturbation of scores—so it is not surprising that f is on average very similar to s.

The data analysis proceeds as follows: first, we examine participants' beliefs for evidence that they overestimate their quiz score. Then, we explore the extent to which they find feedback on their own performance to be 'unlucky,' in spite of the fact that it is generated by a transparently unbiased process. Next, we analyze the process by which participants update their interim beliefs after receiving feedback. Finally, we evaluate the learning process between the first and second quiz in each condition.

3.1. Overestimation

Eliciting the entire distribution of participants' beliefs allows us to observe the entire profile of errors in their belief distribution. As a preliminary measure of overconfidence, however, we

[10] See the appendix for a complete list of the questions used in each quiz.

[11] Table A.9 in the appendix further breaks down scores and beliefs by quiz and by order.

Table 1: Average score (out of ten) and feedback

	Other			Self		
	s	f	N	s	f	N
Combined	5.33	5.37	99	4.91	4.93	90
	(0.25)	(0.30)		(0.25)	(0.28)	
First Quiz	5.15	5.20	54	4.68	4.74	50
	(0.34)	(0.41)		(0.31)	(0.39)	
Second Quiz	5.56	5.58	45	5.20	5.17	40
	(0.37)	(0.44)		(0.39)	(0.42)	
Quiz A	6.90	7.10	20	5.70	5.70	20
	(0.58)	(0.62)		(0.59)	(0.53)	
Quiz B	4.53	4.59	34	4.40	4.50	30
	(0.35)	(0.46)		(0.36)	(0.49)	
Quiz C	4.85	4.65	20	4.65	4.65	20
	(0.61)	(0.72)		(0.49)	(0.64)	
Quiz D	5.56	5.64	25	5.15	5.10	20
	(0.44)	(0.57)		(0.58)	(0.62)	

Standard errors in parentheses

Table 2: Mean Overestimation

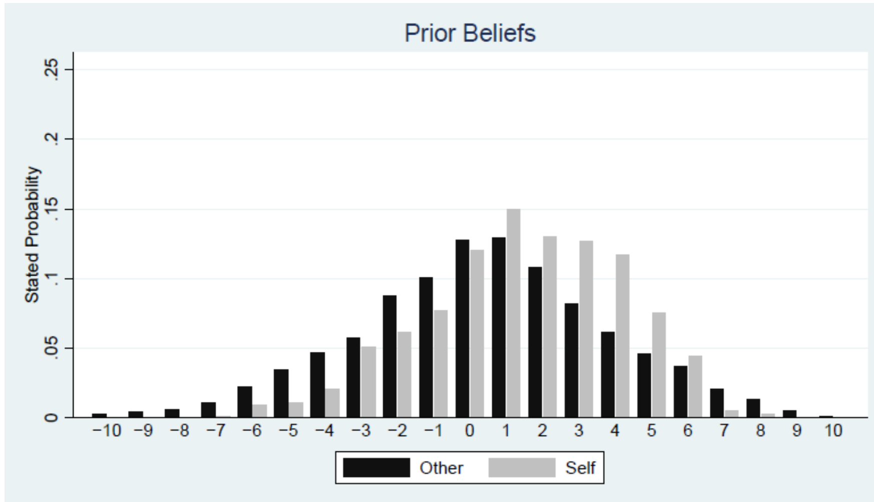
	$\bar{b}_1 - s^T$	$\bar{b}_2 - s^T$	$\bar{b}_3 - s^T$
Other	0.47	−0.23	−0.05
$N = 99$	(0.29)	(0.30)	(0.14)
Self	1.40	1.14	0.67
$N = 90$	(0.23)	(0.20)	(0.15)

Standard errors of *mean* overestimation in parentheses

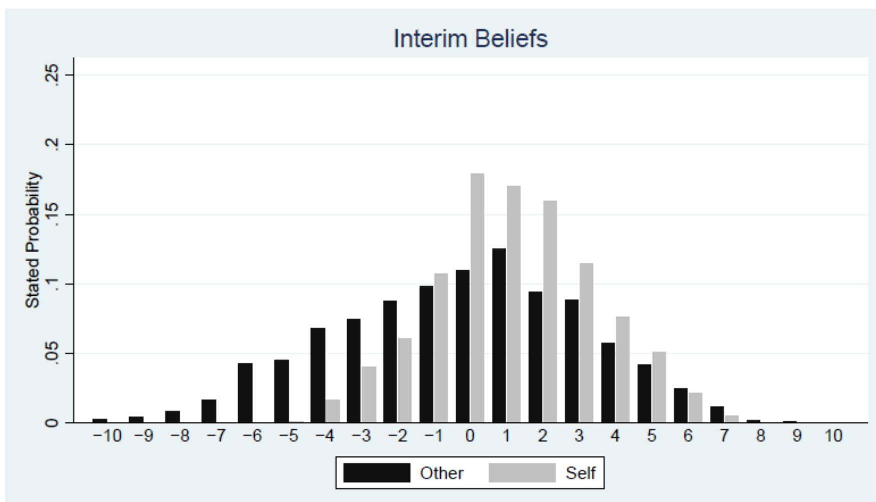
examine the error in the *mean* of each belief elicitation, which we use to proxy participants' 'best guess' about the target score. Let $\bar{b}_j \equiv \sum_{i=0}^{10} b_j^i \times i$ denote the mean of the belief distribution for elicitation j, where b_j^i is a participants' expressed belief in elicitation j that $s^T = i$. We focus on overconfidence manifested by $\bar{b}_j - s^T$, the degree to which the mean of beliefs exceeds the target score. Table 2 presents this measure by condition, separately for each elicitation.

In the *Other* condition, mean prior beliefs exhibit some overestimation. The mean of the prior beliefs distribution is on average .47 points above the target score, which corresponds to 9% of quiz scores, but the significance is only marginal ($t = 1.58$, $p = .12$). The slight overestimation in b_1 suggests that the quizzes are slightly more difficult than they expect. There appears to be no tendency for participants to overestimate the performance of others after they view the quiz or take the quiz. Interim and posterior beliefs actually exhibit slight underestimation, and do not approach significance ($t = -0.77$, $p = .44$ and $t = -0.06$, $p = .72$, respectively).

In the *Self* condition, average overestimation is significantly greater than zero across all three elicitations ($t = 5.89$, 5.62, and 4.43, respectively; $p < .001$ for all three). Further, the magnitudes of overconfidence are quantitatively meaningful, corresponding to an overestimation of 23%, 20%, and 30% of the average score, respectively. The difference in mean overestimation across the two conditions, as tested by a two-sample t-test, is highly significant ($p < .001$) for all three elicitations. The decrease in overestimation from prior to interim beliefs in the *Self* condition further supports the notion that quizzes are more difficult than participants expect when expressing b_1. Importantly, after taking the quiz and learning about its difficulty, participants continue to be overconfident about their own score, a bias that is non-existent in the estimation of a colleague's performance.



(a) Prior Overestimation: $b_1 - s^T$



(b) Interim Overestimation: $b_2 - s^T$

Figure 2: Overestimation

Because we elicit an entire distribution of beliefs, we may also characterize overestimation as a distribution, $b_j - s^T$, by subtracting the true target score, s^T, from each b_j. We aggregate this distribution across all observations and display the resulting average distributions of overestimated beliefs in Figure 2, by condition and separately for each elicitation. The extent to which overestimation in the *Self* condition exceeds that in the *Other* condition is apparent in both elicitations. On average, participants in the *Self* condition place greater weight on outcomes above the actual target score than do those in the *Other* condition, while the opposite is true for outcomes below the target score. While the participants in the *Other* condition on average do not appear to significantly overestimate the target score, both measures show that overestimation is pervasive in the *Self* condition.

Finally, we examine the financial benefit of learning across the three elicitations and the monetary cost of overconfidence. Table 3 shows the means earnings for each elicitation and the hypothetical earnings of the Bayesian posterior beliefs. The fourth column displays the mean earnings from actual posteriors beliefs when the sample is restricted to observations for which Bayesian posteriors may be calculated, i.e. for which feedback was consistent with interim beliefs, so that the comparison with Bayesian earnings is done over the same observations. While the interim-belief earnings of \$1.29 in the *Other* condition present an exception, in general the mean earnings increase across elicitations in both conditions, from \$1.42 to \$2.20 and from \$1.29 to \$1.87 in the *Other* and *Self* conditions, respectively, as the participants learn about the quiz and their score. Prior and posterior (but not interim) earnings are greater in the *Other* condition, though the difference for the posteriors appears to be driven by the participants who were surprised by their feedback, who elevate the earnings of the group in the *Other* condition, but drag down the average earnings in the *Self* condition. While participants are on average earning far less than the \$5 maximum, they do not appear to be incurring significant costs for violating Bayes rule. Actual posterior earnings of \$1.95 in the other condition are only \$0.06 below the Bayesian earnings of \$2.01 in the *Other* condition and while the actual earnings of \$2.00 in the *Self* condition are almost identical to the Bayesian benchmark. Thus, to the extent that participants's overconfidence is costing them, it does not appear to be driven by biased updating.

3.2. Estimated Luck

Next, we investigate whether participants consider noisy but unbiased feedback about their own performance to be 'unlucky.' As b_3 defines participants' post-feedback beliefs about s^T, and feedback is generated by the process $f = s^T + \ell$, the distribution $f - b_3$ implicitly defines participants' beliefs

Table 3: Mean earnings ($) for each elicitation and hypothetical earnings from Bayesian updating

	All participants			'Updateable' b_2 only	
	b_1	b_2	b_3	b_3	b_3^*
Other	1.42	1.29	2.20	1.95	2.01
$N = 99$	(0.07)	(0.07)	(0.08)	(0.06)	(0.16)
Self	1.29	1.53	1.87	2.00	2.01
$N = 90$	(0.06)	(0.07)	(0.06)	(0.15)	(0.16)

Standard errors *mean* in parentheses. Earnings theoretically range from $0 to $5. For the last two columns, $N = 81$ and $N = 80$ for the *Other* and *Self* conditions, respectively.

Table 4: Mean Estimated Luck

	Other	***Self***
$\bar{\ell}^e = f - \bar{b}_3$	0.11	-0.65
	(0.14)	(0.15)
N	99	90

Standard error of mean estimated luck in parentheses.

about their luck, ℓ.[12] We define this distribution $\ell^e \equiv f - b_3$ *estimated luck*. Table 4 summarizes the mean of estimated luck in each condition, which measures participants' estimate of the noise that contributed to their feedback.

Table 4 highlights differences across treatments in estimated luck. While participants in the *Other* condition express beliefs that s^T is perturbed upward by 0.11 questions, on average, those in the *Self* condition estimate that it is perturbed *downward* by 0.65 questions. A two-sample t-test deems this difference significant ($t = 2.57$, $p = .005$). Thus, participants appear to believe

[12]While the support of ℓ is confined to $\{-2, -1, -, 1, 2\}$, some participants express b_3 such that ℓ^e falls outside of this range.

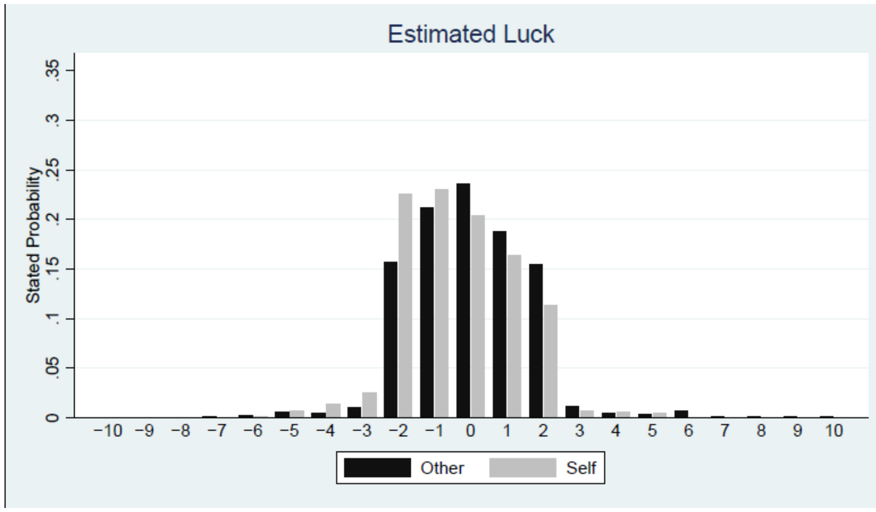


Figure 3: Estimated Luck

that feedback about their own score is *unlucky*, but that feedback about their colleague's score is relatively accurate.

Figure 3 presents the empirical distribution of $\ell = f - b_3$, aggregated across individuals. Note that in both conditions, the bulk of ℓ^e lies between -2 and 2, suggesting that most b_3 is correctly confined to $[f-2, f+2]$. The figure shows a noticeable difference in the distribution of estimated luck across conditions. In the *Other* condition, it is distributed more-or-less symmetrically around a peak of 0, suggesting that participants believe their colleagues' feedback to be neither lucky nor unlucky. In the *Self* condition, the mass of the distribution lies to the left of zero. Subjects find feedback most likely to be an unfair representation of their own performance.

3.3. Information Processing

Why might participants consistently believe that their feedback under-represents their actual performance—that they are 'unlucky'—when is it objective and unbiased? One explanation of this finding is that participants attribute positive outcomes to their performance and negative outcomes to luck.

However, the 'unlucky feeling' need not arise from a bias in the updating process. It may be perfectly consistent with Bayes rule for a participant with overconfident prior beliefs who receives relatively accurate feedback. In this section we investigate the extent to which post-feedback be-

Table 5: Summary statistics for participants who received feedback inconsistent with their interim beliefs

	s	s^T	f	$\bar{b}_1$	$\bar{b}_2$	$\bar{b}_3$
Other	4.50	6.94	7.33	5.69	4.20	6.78
$N = 18$	(0.52)	(0.76)	(1.00)	(0.42)	(0.43)	(0.62)
Self	3.25	3.25	2.42	6.19	6.43	4.75
$N = 12$	(0.70)	(0.70)	(0.78)	(0.54)	(0.53)	(0.64)

Notes: Standard error of means in parentheses. Participants summarized in this table expressed beliefs that place no probability on scores within 2 points of f.

lief updating conforms to the predictions of Bayes' rule. Below, we compare participants' actual updating process between b_2 and b_3 to the benchmark provided by Bayes' rule.

Feedback Inconsistent with Interim Beliefs. Eighteen participants in the *Other* condition, and 12 participants in the *Self* condition received feedback that was inconsistent with their interim beliefs. Table 5 summarizes the performance and beliefs of these participants.

Participants with 'un-updateable' interim beliefs differ from others along multiple dimensions. Their scores are lower, and their beliefs less accurate, than their counterparts with a Bayesian benchmark. Of the 18 participants in the *Other* condition represented in Table 5, interim beliefs lie entirely below $f - 2$ for 13. Of the 12 represented in the *Self* condition, interim beliefs lie entirely above $f + 2$ for 12. In other words, Bayes' rule makes no prediction for the participants who are most pessimistic about the performance of their peers, and most overconfident about their own performance.

While Bayes' rule provides no benchmark for b_3 for these participants, we can construct a proxy for updating in these cases. According to this proxy, b_3^P, participants expressing un-updateable interim beliefs (b_2) with $\bar{b}_2$ higher (lower) than f should express posterior beliefs (b_3) with probability 1 on $f + 2$ ($f - 2$). Intuitively, when interim beliefs are falsified, participants should believe that their score is as close to their interim beliefs as is possible, given the constraints on possible s^T imposed by f.

Table 6: Average estimated luck versus Bayesian estimated luck

	'Updateable' b_2 only		All participants	
	$(f-\bar{b}_3)$:	$(f-\bar{b}^*_3)$	$(f-\bar{b}_3)$	$(f-\bar{b}^P_3)$
	(1)	(2)	(3)	(4)
Other	0.01	-0.29	0.11	0.14
	(0.12)	(0.14)	(0.14)	(0.10)
N	81	81	99	99
Self	-0.39	-0.43	-0.65	-0.05
	(0.13)	(0.14)	(0.15)	(0.09)
N	78	78	90	90

[1] Columns (1) and (2) summarize only those beliefs that can be updated according to Bayes' rule (those *not* presented in table 5. Column (1) displays $\bar{\ell}^e$, the mean of estimated luck for these participants, and the column (2) their Bayesian benchmark.

[2] Column (3) displays the mean of estimated luck only for *all* participants, including those without a Bayesian benchmark. Column (4) displays the Bayesian benchmark for those beliefs that have one, with the Bayesian proxy b^P_3 for those who do not.

Belief-updating and its Bayesian benchmark. Table 6 presents the average estimated luck for each condition and compares it to the average estimated luck reflected in the Bayesian posteriors, b^*_3. The first two columns display the data only for participants' whose beliefs can be updated according to Bayes' rule, while columns (3) and (4) represent *all* participants, including those whose beliefs have no Bayesian benchmark. Column (4) includes $\bar{b}^*_3$ for those who have it, and the proxy $\bar{b}^P_3$ for participants without a Bayesian benchmark.

Columns (1) and (2) show that the unlucky feeling in the *Self* condition is, on average, largely consistent with Bayes' rule when it applies (-0.39 vs. -0.43, $p = .71$ according to a paired t-test). Participants in the *Other* condition, on the other hand, significantly overestimate their colleague's luck on average ($p < .01$). When un-updateable b_2 are included in the comparison, participants estimate their colleagues' luck accurately.

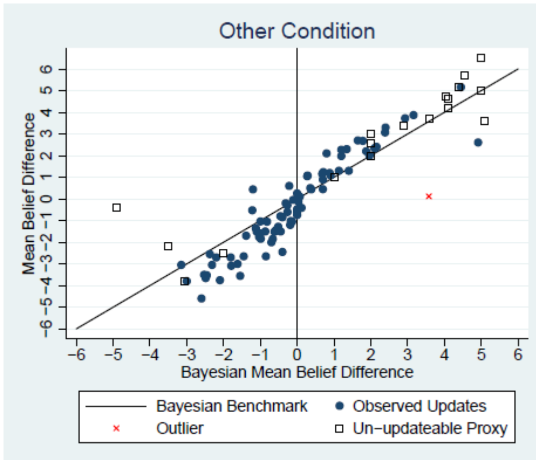
Columns (3) presents observed estimated luck for all participants. Column (4) presents a benchmark for comparison, which is $f - \bar{b}_3^*$ for b_2 with a Bayesian benchmark, and $(f - \bar{b}_3^P)$ for beliefs without a Bayesian benchmark (those represented in Table 5). When these beliefs and benchmarks are included, conclusions about estimated luck change in each condition. Estimated luck appears largely a result of updating in the *Other* condition (0.11 vs. 0.14, $p = .22$), the unlucky feeling in the *Self* condition appears far stronger than the benchmark (-0.65 vs. -0.05, $p < .0001$).

The results presented in table 6 leave the source of the unlucky feeling open to interpretation. Comparing columns (1) and (2) in the *Self* condition suggests that it is merely a result of updating overconfident priors, while comparing columns (3) and (4) hints that it is the result of an ego-preserving bias in the belief-updating process. To the extent that such a bias is present, it is largely concentrated in a small number (30) of the lowest-scoring participants with the least accurate beliefs.

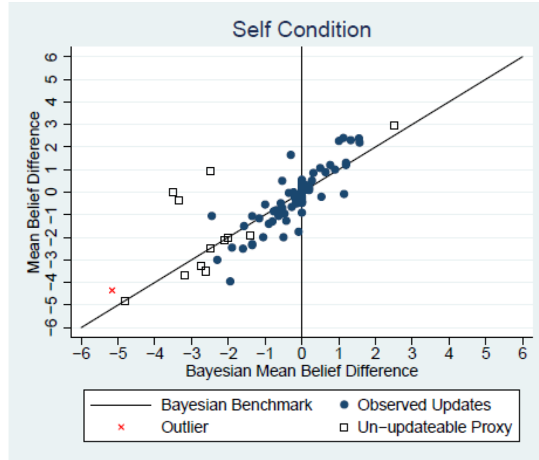
Moebius et al. (2007) find that beliefs about relative performance tend to be 'conservative', or unresponsive to new information. In addition, Eil and Rao (2010) and Moebius et al. (2007) both find belief to be asymmetrically responsive. According to this bias, participants' beliefs respond more strongly to new information if it reflects positively on them than if it reflects negatively. To test for conservatism and asymmetry in the updating beliefs about absolute performance, we compare the mean difference between interim and posterior beliefs, $(\bar{b}_3 - \bar{b}_2)$, to its Bayesian benchmark, $(\bar{b}_3^* - \bar{b}_2)$. Conservatism would suggest that $(\bar{b}_3 - \bar{b}_2) < (\bar{b}_3^* - \bar{b}_2)$. Asymmetry responsiveness would suggest cause $(\bar{b}_3 - \bar{b}_2)$ to be higher, relative to $(\bar{b}_3^* - \bar{b}_2)$, when $(\bar{b}_3^* - \bar{b}_2)$ is positive than when it is negative.

Figure 4 plots $(\bar{b}_3 - \bar{b}_2)$ vs. $(\bar{b}_3^* - \bar{b}_2)$. One observation is labeled an 'outlier' (marked with an 'X') in each condition, for reasons addressed below. The figure also plots $(\bar{b}_3 - \bar{b}_2)$ vs. $(\bar{b}_3^P - \bar{b}_2)$ for un-updateable b_2, with these observation represented as squares. The diagonal line $(\bar{b}_3 - \bar{b}_2) = (\bar{b}_3^* - \bar{b}_2)$ is included to facilitate a comparison of participants' mean difference in beliefs to its Bayesian benchmark. Observations located above the diagonal represent more favorable updating than the Bayesian benchmark, and those below represent less favorable updating.

Figure 4 presents no obvious evidence of conservatism or asymmetry in the updating process, in either condition. Conservatism, or unresponsive beliefs, would cause the observations to form a flatter line than the Bayesian benchmark. Asymmetry would suggest differences in responsiveness



(a) $(\bar{b}_3 - \bar{b}_2)$ vs. $(\bar{b}_3^* - \bar{b}_2)$: Other condition



(b) $(\bar{b}_3 - \bar{b}_2)$ vs. $(\bar{b}_3^* - \bar{b}_2)$: Self condition

Figure 4: Updating Mean Beliefs

on either side of $(\bar{b}_3^* - \bar{b}_2) = 0$. While neither is evident in the figure, responsiveness is explored more quantitatively below. A striking feature of panel 4(b) is the abundance of data points to the left of zero. This is a result of the prevalence of overconfidence manifest in b_2. Such overconfidence makes 'good news' relatively rare.

Table 7 presents a series of regressions that measure conservatism and asymmetry in the belief-updating process. The dependent variable in each regression is the difference in mean beliefs between interim and posterior beliefs, $(\bar{b}_3 - \bar{b}_2)$. Control variables include the Bayesian mean update $(\bar{b}_3^* - \bar{b}_2)$, a dummy variable, GN, indicating that feedback is 'good news' $(f > \bar{b}_2)$, and the interaction term $(\bar{b}_3^* - \bar{b}_2) \times GN$. Bayes' rule suggests that the coefficient on $(\bar{b}_3^* - \bar{b}_2)$ is one, and all others are zero. A coefficient significantly lower than one would be evidence of conservatism, while a significant coefficient on the interaction term, $(\bar{b}_3^* - \bar{b}_2) \times GN$, would be evidence of asymmetry.[13]

The regression presented in column (1) of table 7 includes all updateable interim beliefs. Column (2) is identical, except that a single outlying observation is dropped in each condition (that denoted with the red 'X' in figure 4). Column (3) includes *all* participants, with $(\bar{b}_3^P - \bar{b}_2)$ included as a

[13]Participants for whom $f = \bar{b}_2$ are categorized neither as Bad nor Good news, and are therefore omitted from Good News-Bad News analysis.

Table 7: Updating Regressions

	Other			*Self*		
	(1)	(2)	(3)	(1)	(2)	(3)
$(\bar{b}_3^* - \bar{b}_2)$	1.14	1.14	0.70	0.94	1.06	0.74
	(0.11)	(0.11)	(0.27)	(0.12)	(0.24)	$(0.12)^{***}$
GN	1.10	1.02	1.39	0.44	0.40	0.49
	$(0.23)^{***}$	$(0.23)^{***}$	$(0.33)^{***}$	$(0.14)^{***}$	$(0.16)^{**}$	$(0.11)^{***}$
$(\bar{b}_3^* - \bar{b}_2) \times GN$	-0.37	-0.24	0.22	0.30	0.18	0.45
	$(0.21)^{*}$	(0.20)	(0.29)	(0.22)	(0.31)	$(0.14)^{***}$
Constant	-0.55	-0.53	-0.93	-0.28	-0.22	-0.33
	$(0.14)^{***}$	$(0.13)^{***}$	$(0.30)^{***}$	$(0.11)^{**}$	(0.14)	$(0.09)^{***}$
Outliers Dropped	—	✓	—	—	✓	—
Bayes' Proxy	—	—	✓	—	—	✓
R^2	0.8704	0.9021	0.8734	0.7713	0.7477	0.7016
N	80	79	98	76	75	88

Notes: Dependent Variable: $(\bar{b}_3 - \bar{b}_2)$. $GN = 1$ *iff* $\bar{b}_3^* > \bar{b}_2$. Three participants for whom $f = \bar{b}_2$ are not included in this analysis, as it is considered neither good nor bad news. Asterisks applied to $(\bar{b}_3^* - \bar{b}_2)$ denote significant difference from one. Asterisks on all other coefficients denote significant difference from zero.

control variable for beliefs with no b_3^*.

Regressions (1) and (2) show no evidence of conservatism in either condition. In fact, beliefs appear slightly more responsive to information than the Bayesian benchmark predicts, as the coefficient on $(\bar{b}_3^* - \bar{b}_2)$ is greater than one. However, the inclusion of $(\bar{b}_3^P - \bar{b}_2)$ causes participants to appear significantly conservative on average, in both conditions.

The significance of GN's coefficient, particularly in the *Other* condition, suggests a discontinuity in the belief-updating process at zero. An interpretation is that beliefs are over-responsive to mildly informative feedback (that which should alter their beliefs only slightly) in both directions. This effect is present but smaller in the *Self* condition.

The coefficient of the interaction term, $(\bar{b}_3^* - \bar{b}_2) \times GN$, is our measure of asymmetry in the belief-updating process. A significantly positive coefficient would suggest that beliefs are more sensitive to good news than bad news, while a negative coefficient would suggest the opposite. Interestingly, the regression in column (1) shows significant negative asymmetry in the *Other* condition, and moderate positive asymmetry in the *Self* condition. In other words, participants' beliefs are significantly more responsive to negative information about others and marginally more responsive to positive information about their own absolute performance. This is consistent with existing experimental studies that find a positive bias in responses to feedback on *relative* performance. However, as shown in column (2), these results decrease dramatically in magnitude and significance with the removal of a single outlier in each condition, suggesting that the asymmetry is disproportionately influenced by these observations. The inclusion of $(\bar{b}_3^P - \bar{b}_2)$ changes the sign of the interaction term in the *Other* condition, and restores marginal significance to the ego-preserving asymmetry in the *Self* condition.

Conclusions about the updating process garnered from Table 7 are similar to those drawn from Table 6 about the source of estimated luck. The majority of participants in our design incorporate feedback into their beliefs in a manner consistent with Bayes' rule, and there is little evidence of conservatism or asymmetry. However, Bayes' rule provides no benchmark for a subset of participants with low scores and inaccurate beliefs. These participants appear both conservative and asymmetric in their responses to information, when compared to a proxy (b_3^P) for the Bayesian benchmark.

3.4. Learning Transfer

In the previous section, we show that deviations from Bayesian updating cannot account for our participants' post-feedback overestimation of their own scores, so the remaining overconfidence must primarily be due to overconfident priors. Among all studies cited above, beliefs do respond to feedback, even if the process is slow or biased. Thus, it remains a puzzle how overconfidence originates, and how it persists in the presence of repeated instances of tasks and feedback. The evolution of beliefs over the course of our brief experiment suggests a channel through which overconfidence may persists, even in the absence of bias in the processing of information about performance.

We examine how participants use the experience and feedback from the first quiz to improve the accuracy of their second-quiz estimates. Table 8 shows the average accuracy, measured in terms of overestimation and mean square error, exhibited in prior, interim and posterior beliefs, for the first and second quiz in each condition. The bottom row of each panel summarizes the deviation from Bayesian posteriors, on average.

In the *Other* condition, once participants take the first quiz, there is no significant overestimation, on average, with mean overestimations of -0.19, -0.10, -0.02, -0.28, and -0.01 across the subsequent 5 elicitations. However, in the *self* condition, while mean overestimation is eliminated after first-quiz feedback, on the subsequent quiz it nearly regains first-quiz levels (1.18, 1.08, and 0.65). The minor improvement that is observed may be due the experience of taking the first quiz, using the beliefs-elicitation instrument, and receiving and processing feedback, or it might reflect slightly higher average scores on the second quiz in both conditions.

Analysis of the mean squared error shows that the experience and feedback from the first quiz may be more helpful to participants in the *Other* condition. The mean squared error of the prior estimate drops from 12.52 to 11.07 in the *Other* condition and from 10.11 to 8.26 in the *Self* condition. For the posterior elicitation, mean squared error in the *Other* condition improves by 1.43, from 4.21 to 2.78, while it only improves by 0.10, from 3.71 to 3.61, in the *Self* condition. The improvement in the *Other* condition is marginally significant ($p = .052$) according to a two-sample t-test, while the improvement in the *Self* condition is not statistically significant.

Because participants in the *Other* condition estimate the performance of two different people in the two quizzes, one might expect their experience with the first quiz actually to be *less* relevant for estimating performance on the second quiz than it would be for those in the *Self* condition, who estimate the same person's performance in both quizzes. Greater improvement in the *Other*

Table 8: Average accuracy (overestimation and mean squared error) of beliefs by elicitation and quiz

		Other		***Self***	
	Beliefs	**Overest.**	**MSE**	**Overest.**	**MSE**
First Quiz	b_1	0.88	12.52	1.57	10.11
		(0.41)	(1.88)	(0.33)	(1.10)
	b_2	-0.19	11.42	1.18	6.00
		(0.41)	(1.65)	(0.26)	(0.79)
	b_3	-0.10	4.21	0.68	3.71
		(0.21)	(0.69)	(0.21)	(0.74)
	$b_3 - b_3^*$	-0.35	1.02	-0.11	0.86
		(0.17)	(0.56)	(0.11)	(0.50)
		$N = 54$		$N = 50$	
Second Quiz	b_1	-0.02	11.07	1.18	8.26
		(0.43)	(1.47)	(0.34)	(1.19)
	b_2	-0.28	11.80	1.08	6.70
		(0.45)	(1.48)	(0.32)	(1.28)
	b_3	-0.01	2.78	0.65	3.61
		(0.17)	(0.43)	(0.22)	(0.83)
	$b_3 - b_3^*$	-0.24	0.20	0.06	0.94
		(0.11)	(0.42)	(0.16)	(0.37)
		$N = 45$		$N = 40$	

Notes: Standard errors in parentheses. Comparison to Bayesian posteriors (in the fourth row of each panel) only includes observations for which feedback was consistent with interim beliefs. For these comparisons, $N = 44$ and $N = 47$, respectively, in the top panel, and $N = 37$ and $N = 31$, respectively, in the bottom panel.

condition might also be expected, though. Despite the fact that prior overestimation in the first quiz is greater in the *Self* condition, those estimating their own performance may have private information about their own ability that would dampen the impact of information and experience on beliefs relative to those in the *Other* condition.

However, comparing the MSE of actual posterior beliefs to that of the Bayesian posteriors supports the argument that the limited cross-game learning in the *Self* condition can be attributed to an ego-preserving bias, as opposed to private information. In the *Other* condition, the posterior MSE significantly exceeds the MSE of Bayesian posteriors by 1.02 ($Z = 1.82$, $p < .04$), but by the second quiz, the MSE does not significantly exceed that of the Bayesian posteriors ($Z = .48$, $p < .32$). In the *Self* condition, the posterior MSE does not decline relative to the MSE of Bayesian posteriors, increasing trivially from 0.86 to 0.94 from the first to the second quiz.

4. Discussion and Conclusion

This paper evaluates quantitatively the role subjects feel that luck plays in unbiased, but noisy assessments of their performance in an intelligence-based task. On average, they believe their feedback to underrepresent their score by some 13%. By measuring beliefs both before and after feedback, we are able to distinguish biased updating from Bayesian responses to overconfident priors. We find evidence most supportive of the latter. We find some evidence of biased information processing, but it is concentrated among the subjects with the least accurate, and most overconfident, prior beliefs. Participants accurately estimate the performance of a randomly selected colleague, and accurately believe the feedback their colleagues received to be neither lucky nor unlucky. They also update their beliefs, on average, in a manner consistent with Bayes' rule.

In one sense, our results are consistent with the findings of Eil and Rao (2010) and Charness et al. (2010), namely that overconfident posterior beliefs are more common and more severe when one's own image is at stake than when processing feedback about randomly assigned rankings and or engaging in abstract probability conditioning. However, while those studies, along with Moebius et al. (2007), find ego-preserving biases in the belief-updating process errors drive their results, we do not.

A methodological departure of our study from those mentioned above is that our subjects estimate their absolute performance, which is their score on a ten-question quiz. Our subjects estimate both their own scores and those of a colleague, but in separate treatments. Other studies use relative performance feedback, which *simultaneously* conveys information about participants' own performance and that of others. Moore and Healy (2008) explain that the belief-updating process is likely to differ when estimating one's own performance and that of others, as one is likely to have private information about their own performance.

The feedback that we employ, which is a uniform disturbance of the target score, is also a stylistic departure from the binary feedback used in related work. It restricts the range of possible scores to those close to value of the feedback, and thus is never "totally wrong." It also obviates Bayes rule for a number of our participants, who expressed overly narrow (as well as inaccurate) interim beliefs.

As stated above, we conclude that the overconfident posterior beliefs in the *Self* condition and that fact that estimated luck is negative on average are almost exclusively due to overconfident prior beliefs. We grant that degree to which participants' updating is consistent with Bayes' rule is likely aided by the simple and uniform noise process, and the fact that our updating analysis necessarily excludes participants who received feedback regarded *ex ante* as zero-probability, a group whose members on average overestimated their score more severely and performed more poorly on the quiz than the rest of the participants. Thus, our findings do not directly call into dispute the conclusions of other studies finding biased updating.

Rather, our findings refine our understanding of the mechanisms that lead overconfidence to persist in the face of repeated task-feedback experience and they highlight an important distinction that needs to be made when considering how individuals update beliefs about their performance. Our participants fail to transfer learning from context to the next. Though feedback helps them learn about their performance on one task, our participants do not appear to learn about their underlying ability, and thus exhibit almost the same degree of overestimation of their own performance prior to the second quiz as they do prior to the first quiz. Thus, our participants do appear to learn within each quiz about their own performance, but they do not appear to learn from the first quiz about the process that generates their score on the second quiz, that is, their quiz-taking ability. If the feedback that affects beliefs about performances fails to alter beliefs about the aptitudes that generate them, overconfidence with respect to these aptitudes may persist even in the face of abundant and repeated feedback.

Outside the laboratory, one has a lifetime over which to accumulate consistent and repeated feedback, from which to learn about learning about one's own performance and attributes. Despite the presence of self-serving biases that hinder one's ability to interpret information relevant to one's own performance, individuals *do* learn about their performances. Thus, overestimation should diminish as age and experience provide repeated learning opportunities.

However, our results suggest that, beyond biased updating of beliefs about performance, overconfidence may also persist because of the failure to interpret learning acquired in one context as relevant to another. In our design, we do not observe how strongly participants believe their performances are to one another. Thus, we cannot quantitatively evaluate the process by which they learn from their feedback on one quiz about their performance on the next. Still, we posit that the first quiz in our design is far more similar to the second than are most pairs of related real-world situations between which one may receive information about her abilities. Thus, even the circumstantial evidence of the lack of 'learning transfer' casts doubt about its occurrence in the real world.

Benoit, J.P., Dubra, J., 2007. Overconfidence? MPRA Paper 6017. University Library of Munich, Germany.

Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A., 2010. Overconfidence is a Social Signaling Bias. IZA Discussion Papers 4840. Institute for the Study of Labor (IZA).

Camerer, C., Lovallo, D., 1999. Overconfidence and excess entry. American Economic Review 89, 306 – 318.

Charness, G., Rustichini, A., van de Ven, J., 2010. Overconfidence, Subjectivity, and Self-Deception. Technical Report. UC Santa Barbara. Working paper.

Clark, J., Friesen, L., 2009. Overconfidence in forecasts of own performance: An experimental study. Economic Journal 119, 229–251.

Cooper, D., Kagel, J., 2008. Learning and transfer in signaling games. Economic Theory 34, 415–439. 10.1007/s00199-006-0192-5.

Cooper, D., Kagel, J., 2009. The role of context and team play in cross-game learning. Journal of European Economic Association , 1101–1139.

Cooper, D.J., Kagel, J.H., 2003. Lessons learned: Generalizing learning across games. The American Economic Review 93, pp. 202–207.

Eil, D., Rao, J., 2010. The good news-bad news effect: Asymmetric processing of objective information about yourself. American Economic Journal: Microeconomics .

Ertac, S., 2009. Does Self-Relevance Affect Information Processing? Experimental Evidence on the Resonse to Performance and Non-Performance Feedback. working paper. Koc University.

Greiner, B., 2003. An Online Recruitment System for Economic Experiments. Forschung und wissenschaftliches Rechnen 63, 79–93.

Grosswirth, M., Salny, A.F., Stillson, A., 1999. Match Wits With Mensa: The Complete Quiz Book (Mensa Genius Quiz). Da Capo Press. 1 edition.

Hoelzl, E., Rustichini, A., 2005. Overconfident: Do you put your money on it? Economic Journal 115, 305 – 318.

Kagel, J.H., 1995. Cross-game learning: Experimental evidence from first-price and english common value auctions. Economics Letters 49, 163–170.

Malmendier, U., Tate, G., 2008. Who makes acquisitions? ceo overconfidence and the market's reactions. Journal of Financial Economics 89, 20 – 43.

Moebius, M., Niederle, M., Niehaus, P., Rosenblat, T., 2007. Gender Difference in Incorporating Performance Feedback. Technical Report. Working paper.

Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychological Review 115, 502 – 517.

Svenson, O., 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychologica 47, 143 – 148.

Appendix A. Supplemental Tables and Figures

		Other Treatment				
		Obs.	s	$\bar{b}_1$	$\bar{b}_2$	$\bar{b}_3$
Quiz A	**First Quiz**	10	5.40 (0.82)	5.87 (0.22)	5.54 (0.40)	5.33 (0.46)
	Second Quiz	10	8.40 (0.50)	5.11 (0.42)	4.98 (0.45)	7.73 (0.56)
Quiz B	**First Quiz**	22	4.27 (0.49)	6.29 (0.31)	4.41 (0.32)	4.42 (0.42)
	Second Quiz	12	5.00 (0.43)	6.26 (0.31)	5.37 (0.37)	5.09 (0.49)
Quiz C	**First Quiz**	10	5.70 (1.01)	5.35 (0.38)	4.75 (0.41)	4.98 (0.96)
	Second Quiz	10	4.00 (0.61)	5.22 (0.37)	5.15 (0.33)	4.19 (0.59)
Quiz D	**First Quiz**	12	6.42 (0.50)	6.56 (0.17)	5.98 (0.34)	6.34 (0.49)
	Second Quiz	13	4.77 (0.64)	4.94 (0.36)	5.04 (0.49)	4.95 (0.64)
		Self Treatment				
		Obs.	s	$\bar{b}_1$	$\bar{b}_2$	$\bar{b}_3$
Quiz A	**First Quiz**	10	6.40 (0.72)	6.84 (0.56)	6.77 (0.71)	6.74 (0.68)
	Second Quiz	10	5.00 (0.91)	5.17 (0.83)	5.68 (0.91)	5.28 (0.91)
Quiz B	**First Quiz**	20	4.00 (0.42)	6.47 (0.28)	5.48 (0.34)	4.74 (0.41)
	Second Quiz	10	5.20 (0.63)	7.35 (0.19)	5.76 (0.39)	5.81 (0.50)
Quiz C	**First Quiz**	10	4.10 (0.67)	5.97 (0.77)	5.52 (0.88)	5.09 (0.86)
	Second Quiz	10	5.20 (0.71)	6.49 (0.58)	6.82 (0.54)	6.00 (0.69)
Quiz D	**First Quiz**	10	6.20 (0.70)	6.82 (0.32)	7.38 (0.26)	6.81 (0.46)
	Second Quiz	10	4.10 (0.82)	5.20 (0.59)	5.56 (0.47)	5.02 (0.59)

Table A.9: Performance and Belief by Quiz

Quiz A

Quiz

#1	10 mph 20 mph 30 mph 40 mph 50 mph	A man goes to visit his friend thirty miles away. He doesn't mind speeding, so he travels at 60 miles per hour and arrives in half an hour. On the way back, however, he has a little trouble with his car, and it takes him an hour to reach home. What is his average speed for the round trip?
#2	$90 $95 $100 $105 $110 $115 $120	A man went into a jewelry store and bought a $75 chain, giving the clerk a $100 bill. He returned a few moments later and bought a new watch, giving the clerk a $20 bill and receiving $5 in change. Later, the bank told the store that both the $100 bill and the $20 bill were counterfeit. Ignoring markup, overhead, cost of merchandise, etc., how much money did the store lose?
#3	3 4 5 6 7 8	You have twenty-four socks in a drawer, six each of brown, black, white, and red. How many socks must you take out of the drawer, without looking, to be sure you have a matched pair (of any color)?
#4	1 3 5 7 11 13 15	Which number in the following series of numbers is least like the others? **1, 3, 5, 7, 11, 13, 15**
#5	33.33% 40% 50% 60% 66.66%	The price of an article is cut 40% for a sale. When the sale is over, the store owner wants to bring the price back up to the original selling price. What percentage of the sale price must be added to that sale price to bring the price back up to the original selling price?

#6	8:00 AM 8:30AM 9:00AM 9:30AM 10:00AM	A girl decides to take a long walk in the country and visit a friend on tohe way. She walks at a steady pace of 2.5 miles per hour. She spends 4 hours walking over to her friend's hosue; she has a cup of coffee and a sandwich and talks to her friend, all of which occupies an hour, and then her friend runs her home in the car, over some rough road, at 20 miles per hour. She gets home at 2:30 PM. When did she leave her house?
#7	18:30 20:00 21:00 22:00 23:30	If it were two hours later, it would be half as long until midnight as it would be if it were an hour later. What time is it now?
#8	Banana Radish Strawberry Peach Lettuce	Pear is to apple as potato is to?
#9	1 4 6 10 Can't Answer	A trader buys coffee for $1,200 and sells it for $1,500. For each bag of coffee he earns a profit of $50. How many bags of coffee did he have?
#10	Canberra New York Vienna Madrid	From these four cities, pick the odd one out.

Quiz B

Quiz

#1	1 2 3 4 5	If a jet has a value of 1, and a plane has a value of 2, what is the value of a Concorde?
#2	21 22 23 24 25 26	How many minutes is it before six o'clock if fifty minutes ago it was four times as many minutes past three o'clock?
#3	CGHICOA TTOOORN IMMIA CPOEHNANGE	Which of the following scrambled words is the "odd man out" when the words are unscrambled? **CGHICOA, TTOOORN, IMMIA, CPOEHNANGE**
#4	16 days 17 days 18 days 19 days 20 days	A snail is climbing out of a well. The well is twenty feet deep. Every day the snail climbs up three feet and every night he slips back two feet. How many days will it take him to get out of the well?
#5	1 hour 1.5 hours 2 hours 2.5 hours 3 hours	Your doctor gives you six pills and tells you to take one every half hour. How long does it take you to use up all of the pills?
#6	a kind of imported handgun the Italian name for the town of Leghorn sleight of hand, conjuring tricks a former prime minister of France the title of a famous poem by Keats	Definition: Legedermain
#7	24 120 208	If *diaphanous* and *sheer* do not have the same meaning, cross out all the 9's in the line below. If they do, cross out all the 6's. If *slough* and *cough* are pronounced the same, multiply the number of 4's by 6. If not, add up all of the non-crossed-out numbers and multiply by 4. **9, 4, 6, 4, 9, 4, 6, 9, 4, 6, 9**

#8	10¢ 20¢ 30¢ 40¢ 50¢	Under certain special circumstances, a peach costs 20¢, a banana costs 30¢, and a grapefruit costs 40¢. How much will a pear cost under the same circumstances?
#9	10 11 19 20 21	If you count from 1 to 100, how many 7's will you pass on the way?
#10	36 45 46 64 99	Following the pattern shown in the number sequence below, what is the missing number? **1,8,27,?,125,216**

Quiz C

Quiz

#1	10 11 19 20 21	If you count from 1 to 100, how many 7's will you pass on the way?
#2	14 15 16 17 18	Which number comes next in this series of number? **2, 3, 5, 7, 11, 13, ?**
#3	24 120 208	If *diaphanous* and *sheer* do not have the same meaning, cross out all the 9's in the line below. If they do, cross out all the 6's. If *slough* and *cough* are pronounced the same, multiply the number of 4's by 6. If not, add up all of the non-crossed-out numbers and multiply by 4. **9, 4, 6, 4, 9, 4, 6, 9, 4, 6, 9**
#4	CGHICOA TTOOORN IMMIA CPOEHNANGE	Which of the following scrambled words is the "odd man out" when the words are unscrambled? **CGHICOA, TTOOORN, IMMIA, CPOEHNANGE**
#5	16 days 17 days 18 days 19 days 20 days	A snail is climbing out of a well. The well is twenty feet deep. Every day the snail climbs up three feet and every night he slips back two feet. How many days will it take him to get out of the well?
#6	a kind of imported handgun the Italian name for the town of Leghorn sleight of hand, conjuring tricks a former prime minister of France the title of a famous poem by Keats	Definition: Legedermain
#7	3 rungs 8 rungs 14 rungs 20 rungs 25 rungs	A man moors his boat in a harbor at high tide. A ladder is fastened to the boat, with three rungs showing. The rungs are twelve inches apart. At low tide the water level sinks twenty feet. How many rungs of the ladder are showing now?

#8

yes
no

Does Canada have a fourth of July?

#9

$240
$244
$252
$258
$264

A man bets $24 and gets back his original bet and $48 additional. He spends 25 percent of his winnings at a restaurant to celebrate, and 50 percent of his winnings to buy a present for his wife because he was so late, and his salary was $240, from which he made his original bet. How much money does he have left when he finally arrives home?

#10

Liar
Truthteller

You are at a meeting at which there are only liars and truthtellers. A woman comes up to you and says that the chairman of the meeting told her he was a liar. Is she a liar or a truthteller?

Quiz D

Quiz

#1	6 stamps 12 stamps 18 stamps 24 stamps	If there are twelve one-cent stamps in a dozen, how many two-cent stamps are there in a dozen?
#2	$0.05 $0.10 $0.50 $1.00 $1.05	You have two pencils, a good one and a cheap one. The good one cost $1.00 more than the cheap one. You spend $1.10 for both. How much did the cheap one cost?
#3	dirty, squalid, mean, despicable deaf, hard of hearing a type of acid a thing of small value, a trifle a scab	Definition: Sordid
#4	2 5 7 9 12	A horse salesman went to a horse auction with a certain number of horses. To his first customer, he sold half his horses, plus half a horse. To his second, he sold half of what he had left, plus half a horse. To the last he sold half of what he had left, plus half a horse. Each customer received whole horses, however, not half horses, and he had no horses left. How many horses did he start out with?
#5	2 3 4 5 6 7 8	Complete the series by adding the next number: **0, 0, 1, 2, 2, 4, 3, 6, 4, __**
#6	238 239 240 241 242	How many odd pages are there in a book 479 pages long?
#7	1 4 6 10 Can't Answer	A trader buys coffee for $1,200 and sells it for $1,500. For each bag of coffee he earns a profit of $50. How many bags of coffee did he have?

#8

4
9
10
12
15

Four years ago, Jane was twice as old as Sam. Four years from now, Sam will be 3/4 of Jane's age. How old is Jane now?

#9

11, 5
10, 5
10, 4
11, 6

Continue the following number series with the group of numbers below which best continues the series. **1, 10, 3, 9, 5, 8, 7, 7, 9, 6, ?, ?**

#10

23
25
28
30
32

Look at the drawing. The numbers alongside each column and row are the total values of the symbols within each column and row. What should replace the question mark?

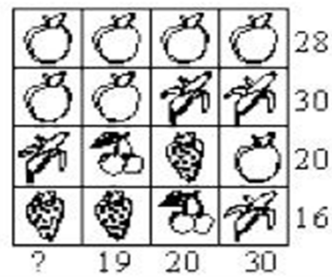


Table 10: The quiz solutions

Question\Quiz	A	B	C	D
1	40 mph	3	20	12 stamps
2	$120	26	17	$0.05
3	5	CPOEHNANGE	208	dirty, squalid, mean, despicable
4	15	18 days	CPOEHNANGE	7
5	66.66%	2.5 hours	18 days	8
6	9:00 AM	sleight of hand, conjuring tricks	sleight of hand, conjuring tricks	240
7	21:00	208	3 rungs	6
8	Radish	20	yes	12
9	6	20	$252	11,5
10	New York	64	Liar	25

Self Condition

The Quiz Experiment: Instructions

Introduction:

Thank you for participating in this experiment. Please follow along carefully as the experimenter reads through these instructions. If you have any questions at any point, please raise your hand.

This is an experiment in the economics of decision-making. A research foundation has provided funds for conducting this research.

For your participation, you will be paid privately and in cash at the end of this session. Your earnings will depend partly on your decisions, and partly on chance. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

You will receive $5 as a participation fee (simply for showing up on time). Details of how you will make decisions and gain subsequent earnings will be provided below.

Summary of the experimental procedure:

The experiment has two parts. Each part has five steps:

1. Before taking a 10 question quiz, you will estimate the number of questions, out of 10, than you will answer correctly.
2. You will take the 10 question quiz.
3. You will again estimate the number of questions, out of 10, that you answered correctly. *For the remainder of the instructions, we will refer to this number as* ***'X'.***
4. You will receive some information about the value of ***X***.
5. You will estimate ***X*** again.

The two parts are exactly the same, except that you will take a different quiz in each part. After you are done with both parts, you will learn the value of ***X*** for each part and then receive your earnings.

You can earn money both by answering questions correctly and by accurately estimating ***X***. The details of how your earnings are determined are explained below.

Detailed experimental procedure:

On the computer in front of you, a folder is open containing 4 files. Please do not touch the computer until you are instructed to do so.

1. ***Estimating X:***

Your estimation of ***X*** will be more involved than simply stating which value is the most likely. You will be asked to indicate how likely you think that ***X*** is equal to each number from 0 to 10.

When the experimenter tells you to do so, you will open the first file by double-clicking on the file with the mouse. This file has a spreadsheet that you will use to give us your estimate of ***X***. The spreadsheet has 11 dropdown menus, each corresponding to an integer between 0 and 10. In the menu below each number, you will select the likelihood with which you believe ***X*** is equal to that number. Beneath the drop-down menus are two charts that will help you. On the left is a chart titled *'**% Likelihood of total # correctly answered**.'* This will give you a graphical representation of the likelihoods that you have entered. On the right is a chart titled "***Total likelihood MUST add up to 100%".*** This chart adds up all of the likelihoods that you have selected. Please ensure that the total is 100%. Once you have finished making your selections, the experimenter will verify that your total adds up to 100%, and then instruct you on how to save and close the file. *Please do not close the file until you are instructed to do so.*

2. ***The Quiz:***

   When the experimenter tells you to do so, you will open the second file by double-clicking on the file with the mouse. This file contains a 10-question quiz. From the time the experimenter tells you to open the file, you will have exactly 10 minutes to answer the questions on the quiz. When the experimenter tells you that your time is up, please stop working on the quiz, and do not touch your computer's mouse or keyboard until given further instructions. While you are taking the quiz, you may use the back of these instructions pages for scratch work. We will not collect these pages or look at them.

3. ***Estimating X again:***
   After you complete the quiz you will estimate ***X*** a second time. The experimenter will ask you to open the third file. This file is identical to the one you used to estimate ***X*** the first time. Once again, you will enter the likelihood with which you believe that ***X*** is equal to each integer between 0 and 10, in a manner identical to that described in step 1, above.

4. ***Receiving information about X:***
   After you have made your second estimation on the value of ***X***, the experimenter will hand you a slip of paper with a number written on it. This number, ***'Y'***, is related to ***X*.** However, we call it *imperfect* information because ***Y*** will be partly determined by chance, and may or may not be equal to ***X***. Instead, ***Y*** is a number that we get when we add a randomly determined number to ***X***. Specifically, ***Y*** is equal to ***X*-2, *X*-1, *X*, *X*+1 or *X*+2,** each with equal probability.

5. ***Estimating X a final time:***

Finally, the experimenter will ask you to open the fourth file, which is identical to the files you used to estimate ***X*** the other times. Once again, you will enter the likelihood with which you believe that ***X*** is equal to each integer between 0 and 10, in a manner identical to that described in step 2, above.

6. ***Your earnings:***
You can earn money in two main ways during the course of this experiment:

1. Answering questions correctly
At the end of the experiment, *for each part*, 1 of the 10 questions will be selected at random and you will be paid an extra $5 if you answered it correctly.

2. Accurately Estimating Performance
For each part, exactly one of the three estimates that you made will be randomly selected for payment. You will be paid ($1 + 5*p_X - w$) dollars for each of your two estimations, where:

i. p_X is the likelihood that you expressed for the actual value of ***X***, and
ii. w is the sum squares of the likelihood that you attached to each of the 11 scores.

<u>This formula may appear complicated, but what it means for you is very simple</u>: *You get paid the most when you honestly report your best guesses about the likelihood of each of the different possible outcomes.* The range of your payoffs is from $0 to $5 for each set of guesses. Since we select two of these estimates for payment (one for each part), you could earn up to a theoretical maximum of $10 for two perfect guesses (but that would be very difficult to do).

Example:
As an example, imagine that, at the end of the experiment, you learn that ***X***=5.

a. If you had expressed that you were 100% certain that ***X***=5:
   - $p_X = 1.0$
   - $w = 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 = 1$
   - Earnings from estimation = $(1 + 5*1.0 - 1) = $5.00

b. If you had expressed that you were 100% certain that ***X***=4:
   - $p_X = 0$
   - $w = 0^2 + 0^2 + 0^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 = 1$
   - Earnings from estimation = $(1 + 5*0 - 1) = $0.00

c. If you expressed likelihoods of 10%, 20%, 30% and 40% for ***X***=5, ***X***=6, ***X***=7 and ***X***=8, respectively:
   - $p_X = .10$
   - $w = .1^2 + .2^2 + .3^2 + .4^2 = .30$
   - Earnings from estimation = $(1 + 5*.10 - .30) = $1.20

d. If you expressed likelihoods of 10% and 90% for ***X***=5 and ***X***=6 respectively:
   - $p_X = .10$
   - $w = .1^2 + .9^2 = .82$
   - Earnings from estimation = \$(1 + 5*.10 - .82) = \$.78

To summarize: there are two parts. ***X*** is the number of questions that you answered correctly on the quiz. In each part, you will estimate ***X*** three times: before you take the quiz, after you take the quiz, and after you receive some information about ***X***. You will be paid \$5 + your earnings from the first quiz + your earnings from one of the estimates of ***X*** from part 1 + your earnings from the second quiz + your earnings from one of the estimates of ***X*** from part 2.

Once the experimenter has answered any questions you may have, we will begin part 1.

Other Condition

The Quiz Experiment: Instructions

Introduction:

Thank you for participating in this experiment. Please follow along carefully as the experimenter reads through these instructions. If you have any questions, please don't hesitate to raise your hand.

This is an experiment in the economics of decision-making. A research foundation has provided funds for conducting this research.

For your participation, you will be paid privately and in cash at the end of this session. Your earnings will depend partly on your decisions, partly on the decisions of others, and partly on chance. If you follow the instructions and make careful decisions, you may earn a considerable amount of money.

You will receive $5 as a participation fee (simply for showing up on time). Details of how you will make decisions and gain subsequent earnings will be provided below.

Summary of the experimental procedure:

The experiment has two parts. Each part has six steps:

1. You will be randomly paired with one of the other participants in the room, whom we will call your 'colleague.' You will not learn which of the other participants is your colleague.
2. Before taking a 10 question quiz, you will estimate the number of questions, out of 10, that your colleague will answer correctly.
3. You will take the 10 question quiz. Everyone in the room will take the same quiz.
4. You will again estimate the number of questions, out of 10, that your colleague answered correctly. *For the remainder of the instructions, we will refer to this number as '**X**'.*
5. You will receive some information about the value of ***X***.
6. You will estimate ***X*** again.

The two parts are exactly the same, except that you will take a different quiz in each part. After you are done with both parts, you will learn the value of ***X*** for each part and then receive your earnings.

You can earn money both by answering questions correctly and by accurately estimating ***X***. The details of how your earnings are determined are explained below.

Detailed experimental procedure:

On the computer in front of you, a folder is open containing 4 files. Please do not touch the computer until you are instructed to do so.

1. ***Estimating X:***
   Your estimation of ***X*** will be more involved than simply stating which value is the most likely. You will be asked to indicate how likely you think that ***X*** is equal to each number from 0 to 10.

   When the experimenter tells you to do so, you will open the first file by double-clicking on the file with the mouse. This file has a spreadsheet that you will use to give us your estimate of ***X***. The spreadsheet has 11 dropdown menus, each corresponding to an integer between 0 and 10. In the menu below each number, you will select the likelihood with which you believe ***X*** is equal to that number. Beneath the drop-down menus are two charts that will help you. On the left is a chart titled *'**% Likelihood of total # correctly answered**.'* This will give you a graphical representation of the likelihoods that you have entered. On the right is a chart titled "***Total likelihood MUST add up to 100%".*** This chart adds up all of the likelihoods that you have selected. Please ensure that the total is 100%. Once you have finished making your selections, the experimenter will verify that your total adds up to 100%, and then instruct you on how to save and close the file. *Please do not close the file until you are instructed to do so.*

2. ***The Quiz:***

   When the experimenter tells you to do so, you will open the second file by double-clicking on the file with the mouse. This file contains a 10-question quiz. From the time the experimenter tells you to open the file, you will have exactly 10 minutes to answer the questions on the quiz. When the experimenter tells you that your time is up, please stop working on the quiz, and do not touch your computer's mouse or keyboard until given further instructions. While you are taking the quiz, you may use the back of these instructions pages for scratch work. We will not collect these pages or look at them.

3. ***Estimating X again:***
   After you complete the quiz you will estimate ***X*** a second time. The experimenter will ask you to open the third file. This file is identical to the one you used to estimate ***X*** the first time. Once again, you will enter the likelihood with which you believe that ***X*** is equal to each integer between 0 and 10, in a manner identical to that described in step 1, above.

4. ***Receiving information about X:***
   After you have made your second estimation on the value of ***X***, the experimenter will hand you a slip of paper with a number written on it. This number, ***'Y'***, is related to ***X*.** However, we call it *imperfect* information because ***Y*** will be partly determined by chance, and may or may not be equal to ***X***. Instead, ***Y*** is a number

that we get when we add a randomly determined number to ***X***. Specifically, ***Y*** is equal to ***X*-2, *X*-1, *X*, *X*+1 or *X*+2,** each with equal probability.

5. ***Estimating X a final time:***
Finally, the experimenter will ask you to open the fourth file, which is identical to the files you used to estimate ***X*** the other times. Once again, you will enter the likelihood with which you believe that ***X*** is equal to each integer between 0 and 10, in a manner identical to that described in step 2, above.

6. ***Your earnings:***
You can earn money in two main ways during the course of this experiment:

1. Answering questions correctly
At the end of the experiment, *for each part*, 1 of the 10 questions will be selected at random and you will be paid an extra $5 if you answered it correctly.

2. Accurately Estimating Performance
For each part, exactly one of the three estimates that you made will be randomly selected for payment. You will be paid ($1 + 5*p_X - w$) dollars for each of your two estimations, where:

i. p_X is the likelihood that you expressed for the actual value of ***X***, and
ii. w is the sum squares of the likelihood that you attached to each of the 11 scores.

<u>This formula may appear complicated, but what it means for you is very simple</u>: *You get paid the most when you honestly report your best guesses about the likelihood of each of the different possible outcomes.* The range of your payoffs is from $0 to $5 for each set of guesses. Since we select two of these estimates for payment (one for each part), you could earn up to a theoretical maximum of $10 for two perfect guesses (but that would be very difficult to do).

Example:
As an example, imagine that, at the end of the experiment, you learn that ***X***=5.

a. If you had expressed that you were 100% certain that ***X***=5:
   - $p_X = 1.0$
   - $w = 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 = 1$
   - Earnings from estimation = $(1 + 5*1.0 - 1) = $5.00

b. If you had expressed that you were 100% certain that ***X***=4:
   - $p_X = 0$
   - $w = 0^2 + 0^2 + 0^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 + 0^2 = 1$
   - Earnings from estimation = $(1 + 5*0 - 1) = $0.00

c. If you expressed likelihoods of 10%, 20%, 30% and 40% for ***X***=5, ***X***=6, ***X***=7 and ***X***=8, respectively:
   - $p_X = .10$
   - $w = .1^2 + .2^2 + .3^2 + .4^2 = .30$
   - Earnings from estimation = $(1 + 5*.10 - .30) = $1.20

d. If you expressed likelihoods of 10% and 90% for ***X***=5 and ***X***=6 respectively:
   - $p_X = .10$
   - $w = .1^2 + .9^2 = .82$
   - Earnings from estimation = $(1 + 5*.10 - .82) = $.78

To summarize: there are two parts. ***X*** is the number of questions that your colleague answered correctly on the quiz. In each part, you will estimate ***X*** three times: before you take the quiz, after you take the quiz, and after you receive some information about ***X***. You will be paid $5 + your earnings from the first quiz + your earnings from one of the estimates of ***X*** from part 1 + your earnings from the second quiz + your earnings from one of the estimates of ***X*** from part 2.

Once the experimenter has answered any questions you may have, we will begin part 1.