To follow the herd or break away? Overconfidence and Social Learning*

Mir Adnan Mahmood[+] and Jason Paulo Tayawa[+]

[+]Department of Economics, The Ohio State University

Draft: November 20, 2022

Latest draft available here

Abstract

We study the effects of overconfidence in a sequential social learning setting. In a lab experiment, we let subjects form beliefs about their own and others' quality of information by tying the accuracy of their signal to their score on a trivia quiz. Their beliefs about the expected scores allow us to measure and study the effects of confidence on social learning. Our results show that there are two distinct effects of confidence manifesting in their behavior of breaking herds. First, subjects that exhibit more confidence about their *relative* quiz performance are more likely to follow their signal than the herd. Second, subjects who realize that their *absolute* performance is better than expected are also more likely to follow their signal. The *relative overconfident* subjects are more likely to benefit from following their signals in easy quizzes, while *absolute underconfident* subjects are more likely to benefit in hard quizzes after scores are revealed. These findings can be partially explained by a model of social learning where rational agents have information structures that induce overconfidence about their relative signal accuracy.

Keywords: Overconfidence; Belief Updating; Social Learning; Herding

JEL Classification: C92; D83; D91

*Mahmood: mahmood.69@osu.edu; Tayawa: tayawa.1@osu.edu. Both authors contributed equally to the manuscript. We would like to thank Yaron Azrieli, P.J Healy, John Kagel, John Rehbeck, Hyoeun Park, Changkuk Im, participants at the OSU Theory/Experimental Reading group and participants at the ESA North American Regional Meetings, the Midwest Economic Association Annual Meetings, and the Stony Brook International Conference on Game Theory. Funding for this project was provided by the JMCB Small Research Grant and the Burton Abrams Dissertation Award. This study was determined exempt under OSU's IRB with protocol numbers 2020E1126 and 2021E0985. All errors are our own.

1 Introduction

In his widely regarded experiment, Svenson (1981) shows that more than half of the US participants considered themselves to be safer drivers than 80% of the group and more than 90% ranked themselves to be better drivers than half of the group. These results have been attributed to *overconfidence*, a behavioral bias which manifests in people who have skewed beliefs about their ability, relative performance, difficulty of task undertaken, or chance of success (Moore and Healy, 2008; Huffman et al., 2022). In economics, an overconfident agent forms extreme beliefs about an ego-relevant domain, which stems from a miscalibration of the quality of their information (see, for example, Möbius et al., 2022). This bias has been shown to have detrimental effects in several domains. In managerial decision making, overconfident managers engage in aggressive corporate policies (Ben-David et al., 2013; Malmendier and Tate, 2005). In politics, overconfident individuals exhibit ideological extremeness and partisan identification (Ortoleva and Snowberg, 2015). In market settings, overconfident consumers pay for more goods than what they consume (Grubb, 2015). In laboratory experiments, overconfident subjects enter excessively in market entry games (Camerer and Lovallo, 1999), misdetect lies that propagate misinformation (Serra-Garcia and Gneezy, 2021; Kartal and Tyran, 2022), and exacerbate free-riding in public goods games (Yin et al., 2019). However, overconfidence need not always lead to bad outcomes. One such setting where it can play a positive role is social learning.

In this paper, we study the effects of overconfidence in a sequential social learning setting. Here, each agent receives two forms of information, one is in a form of a private signal and another is the publicly observable history of actions of her predecessors. When an individual observes her predecessors' actions, the public information will eventually be strong enough that she will ignore her private signal. Consequently, her action no longer adds to the informativeness of the publicly observed history. This holds true for all subsequent movers, so only the earliest movers' information is encoded in the public information. Such phenomenon is termed as *informational cascades* and *herding* (Bikhchandani et al., 1992; Banerjee, 1992; Bikhchandani et al., 2021). In the worst case, this can lead to the society "herding" on the wrong action if the earliest predecessors receive incorrect signals.[1] Overconfidence can help mitigate this problem. Since overconfident individuals overweight their private signal relative to the public information, they follow their signals more often. And so, the society also learns their private signal. Hence, overconfident agents can provide a positive information externality for subsequent agents. Furthermore, agents that are overconfident can potentially benefit from it. If we consider an expert that is mistakenly underconfident and thus believes her signal is less accurate than others, then she would likely follow a herd of individuals with less accurate signals. If instead she is overconfident, then she would follow her signal and break away from a potentially incorrect herd.

[1]In what follows we use the convention that *informational cascade* or *cascade* is said to occur when the optimal action of an agent is to follow their predecessors regardless of their private signal, while *herding* is said to occur when a sequence of agents choose the same action.

Our experiment builds on the social learning setting of Anderson and Holt (1997) by incorporating a stage that induces overconfidence. We focus on two forms of overconfidence: *relative* and *absolute*. An agent with relative overconfidence believes she has a more accurate signal than the average. An agent with absolute overconfidence believes her signal is more accurate than it is. To induce overconfidence, we have subjects take part in a trivia quiz prior to them participating in the social learning task. The accuracy of their information in the social learning task is tied to their performance on the quiz. We map the quiz score to the accuracy such that higher scores lead to more accurate signals. Following Moore and Healy (2008), we vary the difficulty of quizzes, either easy or hard, and delay the revelation of their own score to generate and observe the different degrees of relative and absolute overconfidence.[2]

In the social learning portion, a color is randomly assigned to a group and their task is to guess whether the color is red or blue. In a random order, each subject in the group sequentially makes their guesses, which are observed by subsequent subjects. They report their belief about the probability that the assigned color to the group is red. They observe two forms of information: 1) binary indications of their predecessors' guesses presented as a sequence of red and blue colors, and 2) their own private signal.[3] We ask subjects to report their belief twice, once after observing only their predecessors' guesses and again after they receive their private signal. The change in the first and second reported belief allows us to observe whether the subject conforms to or breaks away from a herd. Furthermore, this allows us to observe if the break from a herd leads to a positive or negative change in payoffs. We then compare the variation in a subject's relative and absolute confidence to changes in herd breaking behavior, individual welfare, and information aggregation.

Easy quizzes induce subjects to exhibit *relative overconfidence* and *absolute underconfidence* on their quiz performance. Hence, in expectation, a subject that has taken an easy quiz believes that they have a more accurate signal than others. Furthermore, after revealing their own score in an easy quiz, a subject realizes that they have a more accurate signal than initially expected. The hard quizzes induce the opposite where subjects exhibit *relative underconfidence* and *absolute overconfidence*. As a result, we expect subjects to break herds more often after taking an easy quiz than after taking a hard quiz. As a control, we also let subjects face a sequential social learning game where the accuracy of the signal is randomly assigned. This is to observe their behavior when the subject's belief about their own score is independent of their beliefs about their group's score.

We find that when a private signal opposes the herd, subjects follow their private signal 44.5% of the time after an easy quiz, 30.7% after a hard quiz, and 38.4% when the accuracy is randomly assigned. We measure overconfidence by having subjects report their expected score and the score of a randomly selected individual that has taken the same quiz. Looking at the degree of relative confidence, we find a

[2]Moore and Healy (2008) use different terminologies for their measures of overconfidence. Relative overconfidence is synonymous to their definition of overplacement and absolute overconfidence is synonymous to their definition of overestimation.

[3]The binary indication is based on the reported beliefs of the earlier subjects in the sequence. If the reported belief is above 50%, then the indicated guess is red. Otherwise, it is blue. The exact details are provided in Section 3.

5% marginal effect on the probability that a subject follows their signal when it opposes the herd in easy quizzes, but no significant effect in hard quizzes. For absolute confidence, we find a -6% marginal effect on the probability that a subject follows their signal when it opposes the herd after their own scores are revealed in a hard quiz, but no significant effect in an easy quiz.

From the change in behavior due to overconfidence, we look at the potential welfare effects at the individual and social level. At the individual level, overconfident subjects with high scores stand to gain by following their signal whenever the true average score is relatively lower. We find that subjects exhibiting higher degrees of relative confidence move away from incorrect herds more often in easy quizzes. Those exhibiting lower degrees of absolute confidence are more likely to move away from incorrect herds after realizing their true score in a hard quiz. At the social level, we look at the aggregated information through the guess on the assigned color by the last mover of the group. We find that the last movers correctly guess the assigned color most often in easy quizzes (82.2%), then in the control block of randomly assigned accuracy (59.6%), and least in hard quizzes (58.8%). More interestingly, for the case before scores are revealed, the last movers in hard quizzes correctly guess the color 61.8% of the time, which is more than their average accuracy of 56.5%. This is in contrast with the rate of correct guesses (50.8%) in the control where the average accuracy is higher (67.8%). This suggests that overconfidence can have a positive information externality on subsequent movers.[4]

We highlight three points that contrast with previous models of social learning with overconfidence (see, for instance Odean, 1998; Bernardo and Welch, 2001; Angrisani et al., 2021). First, researchers often attribute the overweighting of private signal to overconfidence. That is, agents process their information in a miscalibrated fashion that puts more weight on their private signal than the actions of their predecessors. Our experiment measures the degrees of overconfidence, relative and absolute, and we show that these measures predict subjects' propensity to follow their signals that contradict the herd. This validates the use of overconfidence to model agents' miscalibrated beliefs. Second, overweighting of the private signal, termed as overprecision, is shown to be beneficial for information aggregation, but at the cost of the overconfident individual (Bernardo and Welch, 2001). In our setting, overconfidence stems from the endogenous determination of signal accuracy. That is, after taking a quiz, they form beliefs about the quality of their signal and the quality of the signal of a randomly selected subject that has taken the same quiz. We find that this form of overconfidence can benefit the individual. Hence, overconfidence that is based on *acquired* expertise, not on misperception or miscalibration, can help information aggregation without decreasing individual welfare.[5] Lastly, to the best of our knowledge,

[4]The effects of overconfidence on informational cascades and information aggregation are discussed in more detail in Appendix A. There, we show that a higher degree of relative confidence in the population weakly increases the length of a herd before an informational cascade occurs. Furthermore, we show that a lower degree of absolute confidence for the experts can weakly increases this length. For the information aggregation, we point out the cases where these two can lead to an increase in the likelihood of a correct cascade. Such a situation is when a small increase in relative overconfidence (or decrease in absolute confidence) triggers the experts to postpone the cascade.

[5]Our notions of relative and absolute overconfidence do not necessarily mean that an agent is biased since they may be generated by a Bayesian updating agent. That is, after taking a quiz, a subject receives a signal, unobserved by the experimenter,

we are the first to study the effects of both types of confidence in a social learning environment. In Kogan (2009), they assign the accuracy of a subject based on their relative ranking in a quiz and measure the relative overconfidence implicitly through the subject's reported estimate of an asset value. Our method differs from Kogan (2009) since we directly measure the subject's beliefs, which allows us to differentiate the correlations of relative overconfidence and absolute overconfidence with herd breaking behavior.

This paper contributes to the large literature on social learning experiments which has advanced considerably since the seminal work of Anderson and Holt (1997). Early works on social learning have expanded on this framework by looking at longer horizons (Kübler and Weizsäcker, 2005; Goeree et al., 2007), the presence of outside options (Cipriani and Guarino, 2005), endogenized private information acquisition (Kübler and Weizsäcker, 2004), and distinguishing and identifying cascades from herds (Çelen and Kariv, 2004). Weizsäcker (2010), in a meta-analysis of social learning experiments, finds that subjects tend to overweight their private information. Recent developments have looked at social learning in networks (Chandrasekhar et al., 2020; Grimm and Mengel, 2020; Dasaratha and He, 2021) and social learning with repeated interactions (Agranov et al., 2022).

There are several experimental works that are closely related to our setting. First, Nöth and Weber (2003) conduct a social learning experiment in which they consider two levels of information accuracy, which are exogenously assigned to subjects. They find that subjects overweight their private information. Our setting departs from theirs in that our information accuracy is endogenously determined by the subjects' performance on trivia quizzes. Angrisani et al. (2021) look at a social learning environment in a continuous action space in which, rather than reporting guesses about a binary state, subjects' actions coincide with their subjective posterior beliefs about the state. In line with previous experiments, they find that subjects tend to overweight their private information.[6] They show that a model of overconfidence in which agents believe that their signal is relatively more accurate than others can capture this behavior. Our setting departs from theirs in that we directly induce overconfidence by tying the signal accuracy to a subjects' quiz performance.

Our paper also adds to the literature on behavioral social learning. Several authors have looked at naive agents that neglect the correlation of observed actions (Eyster and Rabin, 2010; Eyster et al., 2015), altruistic agents that improve information aggregation (March and Ziegelmeyer, 2020), and heterogeneous populations of conformists and non-conformists (Duffy et al., 2021). In relation to behavioral learning, Bohren and Hauser (2021) provide a framework to evaluate misspecified social learning with heterogeneous agents. The model closest to our setting of endogenous expertise is Chen (2022). Their paper reexamines the formation of informational cascades in the presence of ambiguity in the data gen-

and updates her belief about her own performance and the average performance of the group. After her own score is revealed, she then updates her belief about the average performance of the group. See Benoît and Dubra (2011) for a discussion of a population of Bayesian agents that manifests apparent overconfidence.

[6]In a similar paper, De Filippis et al. (2022) look at a restricted two-person setting of Angrisani et al. (2021) and find that the second mover in the sequence overweights their private signal when it contradicts their prior.

erating process of the predecessors' signals. Although we do not model agents with ambiguity aversion, we present the conditions for a cascade to occur in Appendix A. Our analysis follows that of Wu (2015) with experts and laymen. We modify their setting and include a joint distribution of agent types. This joint distribution captures the presence of relative and absolute confidence in the population.

The paper is structured as follows. Section 2 provides an illustrative example discussing the implications of overconfidence in a social learning context. Section 3 presents the experimental design. Section 4 describes the results. Section 5 contains a discussion and final remarks.

2 An Illustrative Example

Suppose that a new technology comes out that can potentially boost a firm's productivity. Since the technology is untested, there is a chance that it can fail. Suppose, for simplicity, that there is an equal chance of $Success$ or failure. Several firms are interested in the technology and need to decide on whether or not they should adopt it. Suppose the payoffs from adopting the technology is equal to 1 if the technology is a $Success$ and 0 otherwise. Conversely, the payoffs from rejecting the technology is 1 if the technology fails and 0 otherwise.

Consider the sequential learning environment as in Bikhchandani et al. (1992) and Banerjee (1992). Firms, in a sequence, make a decision whether to adopt the technology (A) or to reject it (R). Firms can observe the decision of their predecessors when making a decision. Furthermore, they are able to get a noisy signal about the likelihood of the technology's success. One can think of this signal as a recommendation, either adopt (a) or reject (r), coming from research conducted by the firm. Following the convention used in the social learning literature, suppose each firm's signal is identically and independently distributed conditional on the state with the signal's accuracy set at $p > 0.5$.

The case of the first two firms is simple. The first firm always follows its recommendation. The second firm can infer the signal received by the first firm. If its recommendation coincides with the first firm's action then it follows the recommendation. If the recommendation goes against the first firm's action then the second firm randomizes (with equal probability) between choosing A or R.[7]

Consider now the case of the third firm in the sequence. Suppose it observes conflicting decisions from its predecessors (i.e. one chose A and the other chose R). In this case, the third firm can perfectly infer its predecessors' signals. Under Bayes rule, the firm's belief of the probability of $Success$, based on just the predecessors' actions is 50%. Thus, the third firm follows the recommendation based on its own signal. Subsequent firms can then infer the first three firm's signals and act accordingly.

Suppose that the third firm observes the first two firms deciding on the same action. Assume that the first two firms choose A; the case for both choosing R is symmetric. The third firm can perfectly

[7]In our experiment, we randomize between the action taken with equal probability if the subject has reported a belief of 50%.

infer that the first firm's signal was a. However, it cannot perfectly discern the signal received by the second firm. It could have either received a signal a, or received a signal r and randomly chose A with 50% probability. From the perspective of the third firm, the probability of the technology succeeding is given by $\Pr(Success|A, A) = \frac{p(1+p)}{p(1+p)+(1-p)(2-p)} > 0.5$. The third firm in this case will choose A regardless of the signal it receives, because the firm's belief about the probability of $Success$ is too extreme to be overturned by a contradicting signal.[8] As a consequence, the third firm's decision does not provide any information to subsequent firms. Thus, any succeeding firm will act from the perspective of the third firm and always choose to adopt the technology regardless of their signal. Thus, all subsequent firms fall into an *informational cascade* (Banerjee, 1992; Bikhchandani et al., 1992), where their predecessors' actions result in a belief too strong to be overturned by their private information. As a consequence, social learning stops early on, resulting in firms herding on the same action.

One drawback of informational cascades is that the firms may end up herding on the wrong action: If the first two firms receive an incorrect recommendation, this can cause subsequent firms also to choose the wrong action. This can cause significant losses to all the firms.

2.1 Overconfidence and breaking from the herd

Bikhchandani et al. (1992) show that cascade formation and herding are always possible in sequential social learning settings where agents' private signals are conditionally independent and identically distributed. However, certain biases can improve the social learning outcome by postponing the formation of cascades.[9] One such bias that can improve this outcome is overconfidence wherein agents have inflated beliefs about their own expertise and information accuracy.

To see how overconfidence can influence social learning, consider a modification of the previous setup where there are three types of firms, L, M, and H, with corresponding accuracy of 55%, 65%, or 90%, respectively.[10] Furthermore, each firm type forms a belief about the accuracy of a randomly selected firm, which is based from some prior distribution of firm types. Consider the second firm. Suppose that if the second firm is an L-type, it expects that the average accuracy of other firms is 60%. After observing A, it is immediate that the second firm is already in a cascade since it has a lower accuracy than the expected accuracy of the first firm. The same argument holds for an L-type third firm after observing history A, A.

Now suppose that if the third firm is an M-type, it expects that a random firm is type L with 50% chance, M with 40% chance, and H with 10% chance, independently across firms. From observing the history A, A, the third firm knows that an L-type firm follows its signal in the first position but not in the second position. Furthermore, both the M- and H-type firms in the second position will follow

[8]Note that $\frac{p(1-p)}{(1-p)(2-p)} > \frac{p}{1-p}$, which implies that a signal r cannot overturn the public belief after observing two As.

[9]For a discussion on the effects of psychological biases on social learning, see Bikhchandani et al. (2021).

[10]This follows the signal accuracy of the experimental design.

their signal. Given that the true state is $Success$, the probability of observing A, A is $0.625 \times (0.50 \times 1 + 0.40 \times 0.65 + 0.10 \times 0.90) = 0.53125$. If the true state is $Fail$, then the probability of observing A, A is $0.375 \times (0.50 \times 1 + 0.40 \times 0.35 + 0.10 \times 0.10) = 0.24375$. Hence, the probability of $Success$ after observing A, A is about 68.55%, which places the third firm in a cascade (recall that the M-type firm has 65% accuracy).

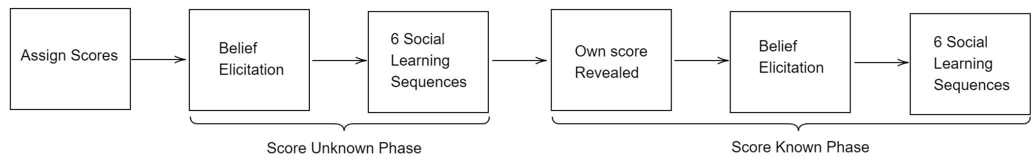
We now look at a population that exhibits more relative confidence. If the distribution of firms allows the M-type firms to be more confident in *relative* terms, that is, M-type firms expect that the average accuracy of another firm is less than 65%, then M-type firms must believe that there are more L-type firms in the population. Let us now suppose that in this new distribution, an M-type third firm expects that a random firm is type L with 70% chance, M with 20% chance, and H with 10% chance. The probability of $Success$ after observing A, A is 64.37%, less than the accuracy of the third firm. Hence, the third firm is not in a cascade and will optimally choose to follow its realized signal.

The same can be argued for a population that exhibits less absolute confidence. Suppose that a firm's belief is symmetric such that Pr(own type = L, other's type = M) = Pr(own type = M, other's type = L). If the distribution of firms allows the M-type firms to be less confident in *absolute* terms, then the unconditional expectation of any firm's own accuracy is less than 65%—a high chance of being L-type. By symmetry of the M-type firm's belief, initially believing that it has a high chance of being L-type implies that it believes there is a high chance that other firms are L-type after observing its own type to be M. From the third firm's perspective, if there is enough mass on the L-type firms, then it will again choose to follow its signal. This follows from the fact that the M-type firms maintain the belief that other firms are still of a lower type even after realizing their own type is higher than expected. Notice that the effect of absolute confidence enters through the belief of firms' relative accuracy, i.e., an increase in the degree of relative confidence.

The illustration shows that gains from following one's signal increases when a population allows for either more relative confidence or less absolute confidence. In terms of information aggregation, we look at the probability of being in a correct cascade. For the third firm, it is seemingly better off if the population is only composed of M-type firms because the probability of $Success$ after A, A is 69.4% compared to 64.37% in the relative overconfident population example. However, a fourth firm cannot learn from the third firm's action in the former case. Hence, the probability of being in a correct cascade is 69.4% after observing A, A, A. In the relatively overconfident population, the third firm's action reveals its signal, and so the fourth firm learns from it. Similarly, the fourth firm is in cascade after observing A, A, A. However, the probability of being in a correct cascade is 74%, which is higher than in a population of just M-type firms.

This illustration suggests that overconfidence in the population can have potential welfare implications. This follows from the idea that overconfidence can induce agents to hesitate and break away from

Figure 1: Timeline in a Block



the herd in favor of their signal. Hence, a herd with overconfident agents carries more information since those overconfident agents are more likely to break the herd, in turn, revealing their privates signal to succeeding agents. A more detailed discussion is found in Appendix A.

3 Experimental Design

In order to test if overconfidence can affect social learning, we focus on the idea of *inducing* overconfidence in the social learning setting of Anderson and Holt (1997). The goal is to induce overconfidence over the accuracy of their own private signals in both relative and absolute terms. Following Moore and Healy (2008), we opt to induce overconfidence by tying the accuracy of each subject's private signal to their performance on trivia quizzes.

3.1 Overview

The experiment consists of 5 blocks in total with subjects randomly assigned into groups of 6 between blocks. Each block proceeds in 2 phases. Figure 1 shows the timeline of each block. Subjects are first assigned a score in integer values from 0 to 10. In Block 1, scores are assigned based on a random draw from a uniform distribution. In Blocks 2-5, scores are based on the subject's performance on a 10-question quiz taken at the beginning of the block. All subjects within a group take the same quiz.

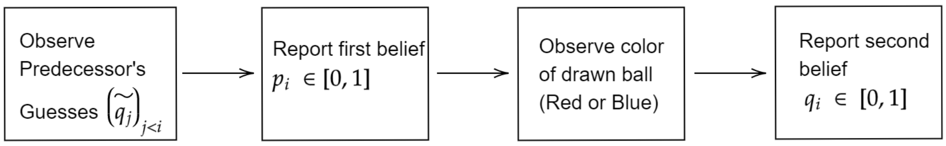
In the *Score Unknown* phase, subjects are not provided with any information about the scores. In the *Score Known* phase, subjects are notified about their own score at the start of the phase, but not about the scores of other group members. Furthermore, for each block, we elicit beliefs about quiz performance at the start of each phase. Following this, subjects then take part in 6 social learning sequences in each phase.

3.2 Social Learning Environment

Each social learning sequence proceeds as follows. First, nature draws a color, either Red or Blue, with equal probability. Subjects only know the prior probability of the colors being drawn. Following this, the first mover reports their *first* posterior belief p_1 for the color Red being assigned to the sequence.[11] The first mover then receives a private signal which takes the form of either a Red ball or a Blue ball that is drawn from a subject specific jar. Then, they report their *second* posterior belief q_1 about the color Red being assigned.

The second mover then sees a binary indication that reflects the first mover's "guess" about the color. This binary indication $\tilde{q}_1$ is either Red or Blue and is based on the first mover's second reported belief q_1. Specifically, if $q_1 > 0.5$, the guess is Red, if $q_1 < 0.5$, then it is Blue, and if $q_1 = 0.5$, then the computer randomly chooses between Red and Blue with equal probability. The second mover then follows the same process of reporting their first belief p_2, receiving a private signal, and reporting their second belief q_2. Subsequent movers follow the same pattern of reporting their first and second beliefs after observing a history of their predecessors' guesses $h_i = (\tilde{q}_j)_{j<i}$ where $\tilde{q}_j \in \{Red, Blue\}$. This pattern is illustrated in Figure 2. The sequence continues until the 6th mover reports their second belief and subjects then move on to the next sequence.

Figure 2: Belief Timeline in Position i



Sequences occurred simultaneously. First, every subject reported their beliefs for the sequence they held the first position in. They then reported their beliefs for the sequence they held the second position and so on. The sequences were ordered to ensure that each subject report their beliefs in every position in the sequence, i.e., each subject i reported their beliefs in each position $j \in \{1, 2, 3, 4, 5, 6\}$.[12] To minimize learning effects on the true scores of their group, subjects were not shown the outcome of any given sequence.

3.3 Inducing Overconfidence

As mentioned above, our goal is to induce overconfidence in subjects to see if it will influence subjects' behavior in a social learning setting. To induce overconfidence, we vary the difficulty of the trivia quizzes, which generates different degrees of relative and absolute overconfidence. Since we want

[11] Reported beliefs were elicited as integers ranging from 0 to 100. The question asked was "What do you think is the probability (from 0 to 100%) that the color assigned to your group is Red for this sequence?"

[12] This design choice cuts down on the time taken for each block. We would like to thank OSub Kwon for his help in coding the sequences in this way.

overconfidence to translate directly over expectations regarding the accuracy of a subject's information, we tie the subject's accuracy to their performance on the trivia quiz.

3.3.1 Quizzes

At the beginning of blocks 2-5, subjects first take part in a 10-question general knowledge trivia quiz. Questions were centered around 4 topics: Geography, History, Science and Entertainment.[13] Subjects had 20 seconds to answer each question by typing their response into a text box. Subjects were notified at the start of each quiz that any spelling errors would be counted as an incorrect answer.

Quizzes are classified as either *Easy* or *Hard*. In the experiment, subjects take part in 2 easy and 2 hard quizzes. This classification was done ex-ante.[14] Subjects faced either one of these two sequences of quizzes: Easy 1, Hard 1, Hard 2, Easy 2 or Hard 1, Easy 1, Easy 2, Hard 2. The order of the quizzes was fixed for all subjects in a given session. The difficulty and the order of quizzes was unknown to subjects. The details of the questions used in the quizzes are given in Appendix D.

3.3.2 Tying overconfidence to beliefs about accuracy

Recall that each subject receives a private signal when making their guesses. This private signal takes the form of the color of a ball drawn from a subject specific jar. The color of the jar matches the color assigned to the group for the given sequence. Each jar is composed of 100 balls, the colors of which are either Red or Blue. If a subject scores anywhere from 0 to 3 points on the trivia quiz, their subject specific Red (Blue) jar contains 55 Red (Blue) and 45 Blue (Red) balls resulting in a signal accuracy of 0.55. If the subject scores anywhere from 4 to 7 points, their Red (Blue) jar contains 65 Red (Blue) and 35 Blue (Red) balls resulting in an accuracy of 0.65. Finally, if a subject scores anywhere from 8 to 10 points, their Red (Blue) jar contains 90 Red (Blue) balls and 10 Blue (Red) balls resulting in accuracy of 0.9. For block 1, the scores are randomly assigned based on a uniform distribution over 0 to 10, with this being common knowledge among subjects.

A consequence of this setup is that the true signal accuracies are nearly identical across subjects for a given quiz, with most of the variation coming from a subject's beliefs about accuracies.

3.3.3 Measuring Overconfidence

To get a reliable measure of overconfidence, we elicit a subject's beliefs not only about their own performance on the quiz, but also about one randomly selected group member's performance (henceforth, GM) on the same quiz. Specifically, we elicit entire distributions: For each score S from 0 to 10,

[13] Questions were sourced from Moore and Healy (2008) and online trivia repositories.

[14] We first conducted a trivia quiz using a separate pool of 88 subjects where participants took part in a 60-question quiz. We had 2 separate quiz banks, each of 60 questions. 44 subjects took part in one quiz and a separate pool of 44 subjects took part in the other quiz. We then sampled questions from each bank. For each easy quiz, we sampled questions that yielded an average score of 9 out of 10. For hard quizzes, we sampled questions that yielded an average of 3 out of 10.

we elicited beliefs that the subject's or the GM's score was S.

We elicit these distributions over scores twice in each quiz block. In the Score Unknown Phase, we elicit distributions for the subject's own performance as well as the performance of a GM after taking the quiz but before starting the social learning sequences. Since subjects do not know their own score, eliciting distributions about their own performance allows us to obtain a measure of *absolute* overconfidence. In the Score Known Phase, we only elicit distributions about the performance of a GM after showing the subject their realized score but before starting the social learning sequences. Comparing a subjects' beliefs about a GM's performance to their own score allows us to get a measure of *relative* overconfidence.

3.4 Incentives and Payment

Our experimental design required incentivizing 4 key components: Quiz participation, eliciting overconfidence measures (both relative and absolute) and reported beliefs in the social learning task. We adopted the random payment mechanism to pay for each of these components. Subjects were first paid for their performance on one randomly chosen quiz. Subjects could earn \$10 $\times$ r where r is the subject's percentile score when pooling their score with the scores of 44 people who took the quiz prior to running sessions.[15] For the same quiz, we then paid subjects for their beliefs about their own and the GM's performance. We randomly chose one score (from 0 to 10) and paid subjects for their beliefs that their own score was equal to the randomly chosen number. We used the BDM mechanism (see Holt and Smith, 2016, for details) under which subjects could earn \$2. We adopted the same procedure to pay them for their beliefs about their GM's quiz performance. Finally, we paid subjects for both of their reported beliefs about the color Red in one randomly selected social learning sequence using the BDM mechanism. Subjects could earn \$12 for each of their reported beliefs in the sequence chosen for payment. Subjects also received a \$5 show-up fee.

Since subjects were paid for both beliefs in the same sequence, they could try to hedge by reporting beliefs that contradicted each other. We ran 3 additional sessions in which we changed the payment scheme for the social learning task. In these sessions, subjects were only paid for one randomly chosen reported belief regarding the color being Red. We again used the BDM mechanism but increased the payout from \$12 to \$18. We also increased the show up fee to \$8 but kept the incentives for the other three components the same.

[15] We opted to not have the incentives for the quizzes tied to the performance of the group members since we wanted to distinguish incentives between eliciting measures of overconfidence and the quizzes.

3.5 Procedures

Sessions were conducted in person at the Ohio State University Experimental Economics Laboratory. Subjects were recruited from the undergraduate and graduate student population using ORSEE (Greiner, 2015). We conducted 8 sessions using the initial payment scheme in the months of October and November 2021. To check for the possibility of hedging, we conducted 3 additional sessions in the month of April 2022 using the second payment scheme outlined above.

Instructions were read out loud at the beginning of the experiment and subjects were allowed to ask clarification questions. Subjects were provided with paper copies of the instructions. Instructions were also programmed into the software, so after the oral description, subjects were able to read the instructions again at their own pace prior to starting the experiment.[16] Sessions lasted between 2 and 2.5 hours, with subjects earning around $25 on average. The experiment was programmed using oTree (Chen et al., 2016).

3.6 Data Description

We collected data for 114 subjects over 11 sessions as described above. We targeted 12 participants per session, but due to low show ups in some instances, we ran the experiment with 6 subjects.[17] 78 subjects participated in the experiment under the initial payment scheme and 36 subjects participated under the second payment scheme. We were unable to find evidence of potential hedging so we pool the data for all sessions.[18]

4 Results

The preview of the results is as follows. We first show that the quiz difficulty and revelation of a subject's own score affect the degree of relative and absolute confidence. Second, we show the effects of varying quiz difficulty and revelation of scores on individual herd breaking behavior and changes in welfare. Third, we show how the degree of relative and absolute confidence vary with these outcomes. Lastly, we look at the composition of herds and the level of information aggregated in a sequence.

4.1 Quiz Outcomes

Table 1 shows the average scores of subjects in easy and hard quizzes. The patterns are consistent with the expectations of quiz difficulty on measures of overconfidence. Subjects, on average, expect their score to be lower than their true score in the Easy quiz, 7.735 points compared to 8.292 points. This

[16] The instructions are reproduced in Appendix C.

[17] For one session, due to experimenter error, we were unable to collect beliefs about quiz performance in the last block of the experiment so we discard the data for the entire block for that session.

[18] Appendix E.1 provides some descriptive evidence against potential hedging.

indicates that subjects, on average, exhibit *absolute underconfidence* in the Easy quiz, underestimating their performance by 0.557 points ($p < 0.0001$). This trend is reversed in the hard quiz where subjects, on average, expect their score to be 3.059 points which is higher than the true average of 2.509 points indicating that subjects exhibit *absolute overconfidence*, overestimating their performance by 0.55 points ($p < 0.0001$).

Table 1: Quiz Scores and Overconfidence

	Easy Quiz		Hard Quiz	
	Score Unknown	Score Known	Score Unknown	Score Known
Own score		8.292		2.509
		(1.838)		(1.093)
E_i(own score)	7.735		3.059	
	(1.984)		(1.439)	
E_i(other score)	7.467	7.388	3.685	3.585
	(1.474)	(1.754)	(1.593)	(1.560)
Relative Confidence	0.268	0.904	-0.626	-1.076
	[.0915]	[.123]	[.112]	[.122]
Absolute Confidence		-0.557		0.550
		[.089]		[.091]
Observations	216	216	228	228

Mean coefficients; standard deviation in the parentheses; standard errors in the parentheses square brackets

On relative performance, we find that in Easy quizzes, subjects consistently overplace their performance relative to others. In the Score Unknown phase, subjects, on average, report a lower score for the GM by 0.268 points ($p < 0.01$), indicating *relative overconfidence*. The degree of relative overconfidence increases in the Score Known phase to 0.904 points ($p < 0.0001$). Again, the patterns reverse for the hard quizzes. Subjects, in the Score Unknown phase, exhibit *relative underconfidence* at 0.626 points ($p < 0.0001$) which increases to 1.076 points ($p < 0.0001$) in the Score Known phase. We summarize these findings below.

Result 1. *On average, subjects are overconfident about their relative performance and are underconfident about their absolute performance in easy quizzes. In hard quizzes, subjects are underconfident about their relative performance and are overconfident about their absolute performance.*

4.2 Overconfidence and Individual Behavior

This section investigates the degree to which overconfidence, in both relative and absolute terms, influences the degree of herd breaking behavior. Intuitively, the degree of overconfidence, relative or absolute, reflects the degree to which subjects believe in the accuracy of their information. As a consequence, subjects that report a higher degree of overconfidence should, in theory, be more likely to follow their own signal. For the purposes of the analysis, we focus on a subsample of observations in which the subjects' first reported beliefs satisfy two conditions. First, their reported belief follows their immediate predecessor's observed action. This condition captures the idea that the subject's immediate predecessor summarizes the information aggregated up to that point and the initial belief accounts for this. Second, their signal contradicts their reported belief. This condition allows us to observe the subject's opportunity to break away from the herd and follow their signal instead. Since we look at the effects on *herding behavior*, we restrict the analysis to when the subject is on the third position in the sequence or later.

This subsample identifies observations in which subjects break away from herds. Essentially, an observation is classified as a *herd break* if the subject's second reported belief, after observing her signal, crosses the 50% threshold in the direction in line with her private signal. As an example, consider a subject who observes her immediate predecessor's action and initially believes that there is a 70% chance that the color assigned is red. Suppose that she receives a blue signal and then revises her belief to 40%. This behavior is classified as a herd break since the subject is revising her belief in the direction opposite to her initial belief and going against the information from her predecessor.

Going a step further, herd breaking behavior allows us to construct a measure of welfare that captures potential gains and losses from breaking away from existing herds. A herd break in the direction of the correct state is considered to be a welfare *gain* of 1, i.e., a change of 1. A herd break in the direction of the incorrect state is considered to be a welfare *loss* of 1, i.e., a change of -1.[19] Consequently, no herd breaks lead to no change, which is set at 0. For instance, consider the same example as before where a subject's blue signal contradicts her initial belief that the color is red with 70% chance. If her second report is below 50%, then she observes a welfare gain of 1 if the state is blue and a welfare loss of 1 if the state is red. If her report stays above 50%, then she observes no welfare change. This measure of welfare change is a way to count breaking away from incorrect herds positively and breaking away from correct herds negatively. We interpret this to mean that the average change in welfare captures the benefits or losses incurred from breaking away the herd relative to conforming to the herd.

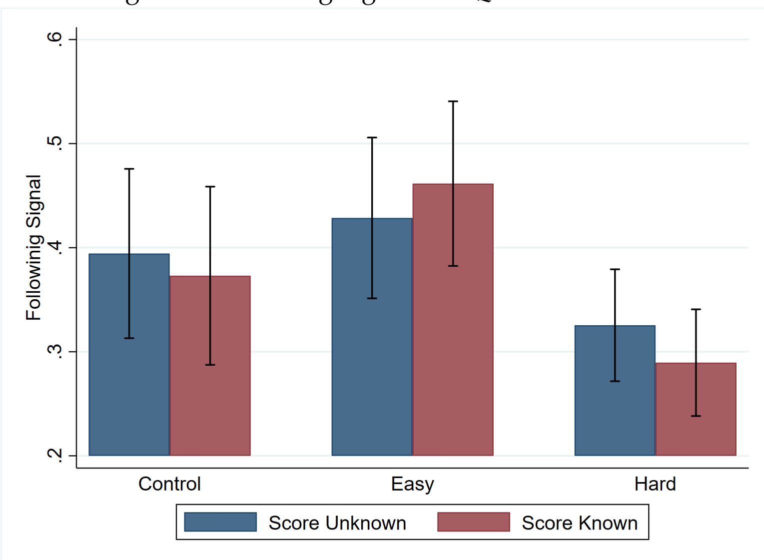
In the following subsections, we investigate the effects of overconfidence on herd breaking behavior and welfare change in tandem. First, we compare the aggregate effects of Quiz blocks versus the Control block, as well as the aggregate effect of subjects knowing their own score. Second, we look at the relationship between relative overconfidence and herd breaking behavior. Finally, we look at the relationship between absolute overconfidence and herd breaking behavior.

[19]If the subject's second reported belief is exactly 50%, we set the measure equal to 1 if the change in belief is towards the correct state and the computer selected a reported belief that matched the correct state. It is set to 0 if the computer selected a reported belief that did not match the state. For the cases when q = 50% but the state and signal do not match, the measure is -1 if the change in belief is towards the incorrect state and is 0 if the change in belief is towards the correct state.

4.2.1 Quiz Difficulty and Score Revelation

Based on the results of overconfidence on quiz performance in Result 1, we expect an aggregate increase in herd breaking behavior in the Easy quiz blocks compared to the Hard quiz blocks due to the higher degree of relative confidence in the former. Figure 3 shows the empirical frequency of subjects changing their belief to match their signal across all blocks, and consequently breaking away from the existing herds. Subjects follow their signals most often in the Easy quiz blocks (mean 44.5%), then in the Control block (mean 38.4%), and least often in the Hard quiz blocks (mean 30.7%). Controlling for phase, order, and session, we find a significant difference between the Control and Hard quiz blocks (regression, $p < 0.1$), and between the Easy and Hard quiz blocks (regression, $p = 0.001$). We do not find a significant difference between the Control and Easy quiz blocks. Before revealing scores, we find that those with high scores, 8-10 points, broke the herd 16 percentage points more (regression, $p < 0.1$) in Easy quiz blocks than in the Control block. Among those with low scores, 0-3 points, subjects broke the herd 19.5 percentage points less (regression, $p < 0.05$) in the Hard quiz blocks than in the Control block. Hence, subjects are responsive to their quiz performance even before their scores are revealed. The difference in responsiveness between the Control block and Quiz blocks suggests that this is due to the presence of relative and absolute confidence in the Quiz blocks.

Figure 3: Following Signals for Quiz and Round



When looking at the effect of score revelation, we see in Figure 3 that there is no significant change in behavior for a given block. This is verified by running a linear regression of herd breaking behavior on phases in Control (regression, $p = 0.894$), Easy (regression, $p = 0.5333$), and Hard (regression, $p = 0.4609$) quiz blocks. Unsurprisingly, this is driven by the heterogeneity in scores among the subjects. That is, subjects with higher scores should be more inclined to follow their signals after their scores are revealed compared to those that scored lower. Between the subjects with low scores, 0-3, and middle scores, 4-7, those with middle scores broke the herd 15 percentage points more (regression, $p < 0.05$) than those with low scores after scores were revealed. This difference is even larger between subjects with low and high scores with 21 percentage points (regression, $p < 0.01$). There is no significant difference in the change of herd breaking behavior for subjects with high and middle score (regression, $p = 0.560$). These effects, however, are mostly coming from the change in behavior in the Control block. This suggests that there is no aggregate difference in behavior for Quiz blocks. Even so, the succeeding discussion shows that the effect of revealing the score on herd breaking behavior is driven by absolute confidence.

Now, we look at the aggregate effects of quiz difficulty and score revelation on welfare. That is, do the aggregate changes in herd breaking behavior necessarily translate to the aggregate improvements in individual welfare? Table 2 compares the average welfare change for Control, Easy and Hard quiz blocks. Since the quiz difficulty affects both the beliefs about scores and the true scores and, in turn, affects welfare, we control for the average score of a subject's group members. Hence, the different levels of quiz difficulty capture the aggregate effect of overconfidence on welfare. Column 1 compares welfare in the Easy quiz blocks to the Control block. We find a positive and significant effect (0.251, $p < 0.1$) of being in an Easy quiz block, which indicates that subjects are more likely to break away from incorrect herds in the Easy quiz block than in the Control block. In comparing the Control block to Hard quiz blocks (column 2), we find no significant difference (0.011, $p > 0.1$). Column 3 compares Easy to Hard quiz blocks. We find that subjects in the Hard quiz blocks are significantly worse off (-1.258, $p < 0.001$) indicating that they are less likely to break away from incorrect herds.

Across all specifications, the effect of realizing own signal accuracy on welfare is not significant. Comparing the subjects with low and middle scores, we find that there is a positive effect of score revelation for those with middle scores than those with low scores (coefficient 0.144, $p < 0.05$). Although we find a difference in behavior between those with low and high scores, we do not find an increase in welfare for the subjects with high scores after scores are revealed (coefficient 0.0249, $p > 0.1$). This suggests that subjects with high scores are following their signals often enough that they break away from correct herds after scores have been revealed.

Result 2. *With respect to quiz difficulty and score revelation:*

a) Subjects break herds more in the Easy and Control blocks compared to the Hard quiz block. Subjects that have middle or high scores break herds more often than those with low scores after scores are revealed.

b) Subjects are more likely to break away from an incorrect herd in Easy quiz blocks than Hard quiz or Control blocks when controlling for the average score. Those with middle scores break away from incorrect herds more often that those with low scores after scores are revealed.

Table 2: Block and Phase Effect on Welfare Change

	Control v Easy	Control v Hard	Easy v Hard
Easy	0.251*		
	(0.127)		
Score Known	-0.0178	-0.0235	-0.0682
	(0.0684)	(0.0665)	(0.0711)
Easy x Score Known	-0.0628		
	(0.108)		
Ave Others' Score	-0.0509*	-0.0176	-0.214***
	(0.0286)	(0.0280)	(0.0524)
Hard		0.0112	-1.258***
		(0.0824)	(0.312)
Hard x Score Known		0.00608	0.0493
		(0.0842)	(0.0872)
Constant	0.378*	0.169	1.960***
	(0.209)	(0.173)	(0.451)
Observations	585	867	916

Regressions comparing the effect of Quiz and Phases on the welfare measure. Standard errors clustered at the subject level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

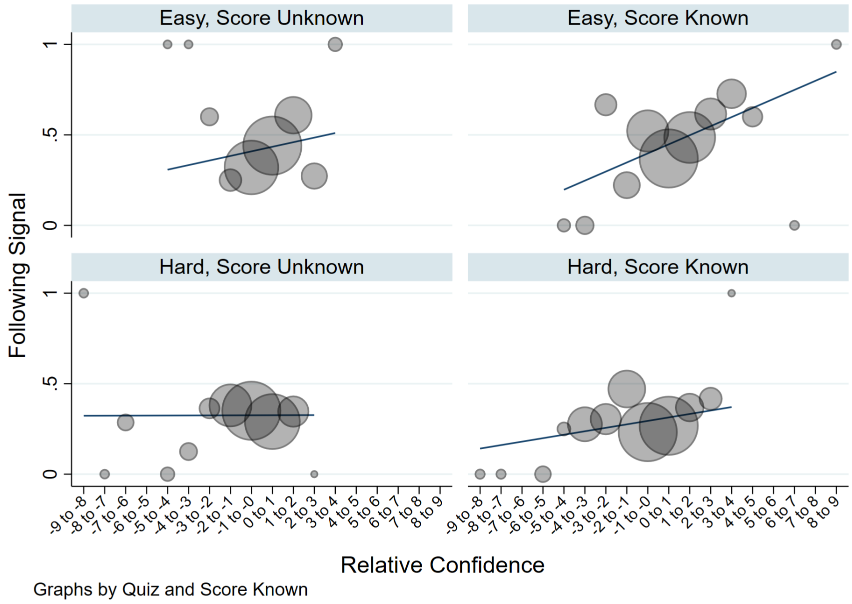
4.2.2 Relative Confidence

Recall that the easy quizzes induce relative overconfidence and absolute underconfidence, while hard quizzes induce the opposite. However, these aggregate effects of quiz difficulty do not capture the degree to which overconfidence varies with the subject's propensity to follow their signal against the herd–that is, breaking a herd. In the following, we look at how our measures for the two forms of overconfidence directly relate to changes in herd breaking behavior and welfare.

Following our illustration in Section 2, relatively overconfident subjects believe that other subjects have a lower accuracy, on average. Hence, we expect that the more relative overconfidence a subject exhibits, the more likely they will follow their own signal. Figure 4 illustrates this relationship of relative confidence and herd breaking behavior for a given quiz difficulty and phase. The fitted line shows this to be a positive linear relationship, i.e., subjects that exhibit more confidence about their relative quiz performance are more likely to follow their signal against the herd. Column 1 in Table 3 shows that this relationship is significant when pooling for both quiz blocks and controlling for the subject's expectation of other subjects' score. That is, a subject that expects to have scored 1 point higher than her peers follows her signal 3.28 percentage points more. This positive relationship is seemingly more robust in Easy quiz blocks than in Hard quiz blocks. For Easy quizzes, a one point increase in the measured relative confidence translates to a 4.95 percentage points ($p < 0.001$) increase in herd breaking, and, for hard quizzes, it translates to 1.62 percentage points increase ($p > 0.1$).

Figure 5 plots the average change in welfare for a given level of relative confidence. Similar to

Figure 4: Following Signals and Relative Confidence



the behavior in herd breaks, there is a weakly increasing relationship between relative confidence and welfare change, which is most apparent in Easy quiz blocks. This positive relationship is statistically significant as shown in columns 1-3 of Table 4. We control for the true average score of other subjects in the sequence, which effects the likelihood that a subject correctly matches the state. The first column pools for both quiz difficulty while the second and third columns look at the effect in the Easy quiz blocks and the Hard quiz blocks, respectively. In Easy quizzes, a 1 point increase in relative confidence translates to an 8.89 percentage point increase when a subject follows their signal against the herd. This effect is halved in the Hard quiz blocks with a 3.93 percentage point increase. Overall, this shows that subjects who exhibit more relative overconfidence are moving away from incorrect herds more often.[20]

Result 3. *With respect to the degree of relative confidence:*

a) *Subjects that exhibit more confidence about their relative performance in easy quizzes are more likely to break herds.*

b) *Subjects that exhibit more confidence about their relative performance are more likely to move away from incorrect herds particularly in easy quizzes.*

[20]Note that we are using a linear regression with a categorical dependent variable. This naively treats the difference between positive welfare change and no change the same as the difference between no welfare change to negative welfare change. In the appendix, the third specification in Table E.2 presents a multinomial logit model with the baseline of no welfare change. It shows that subjects with higher measured relative confidence are moving away from incorrect herds more often and moving towards incorrect herds less often, which are jointly significant (p < 0.001).

Figure 5: Relative Confidence on Welfare

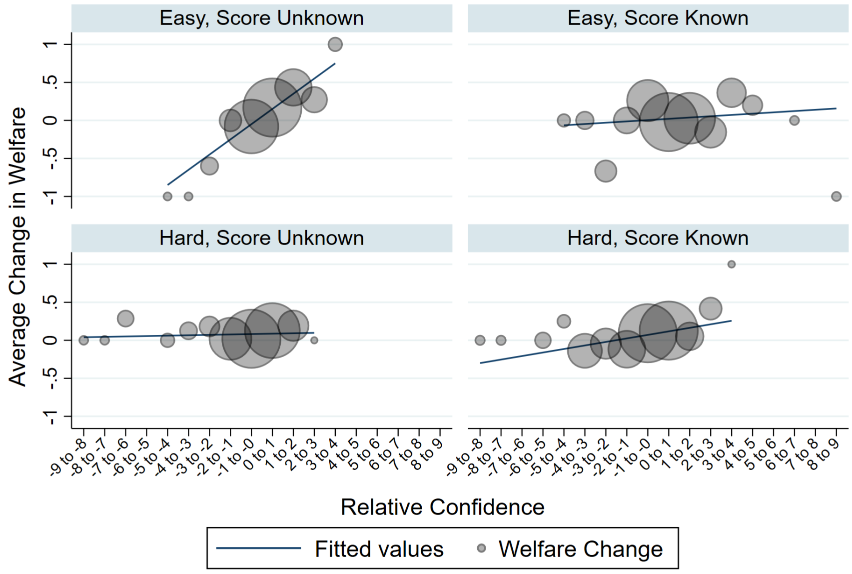


Table 3: Effect of Confidence on Following Signal

	(1)	(2)	(3)	(4)	(5)	(6)
Relative Confidence	0.0328***	0.0495***	0.0162			
	(0.0118)	(0.0186)	(0.0221)			
E_i(other score)	0.0182**	0.0118	-0.00489			
	(0.00846)	(0.0207)	(0.0251)			
Absolute Confidence				-0.0529***	-0.0315	-0.0576**
				(0.0160)	(0.0300)	(0.0253)
E_i(own score)				0.0195**	0.0301*	-0.00370
				(0.00768)	(0.0181)	(0.0224)
Constant	0.504***	0.463***	0.608***	0.500***	0.355**	0.616***
	(0.0792)	(0.168)	(0.115)	(0.0777)	(0.156)	(0.112)
Observations	916	317	599	916	317	599

The table looks at the effects of several variables of interest on following signal against the herd. Columns 1-3 columns present the effect of relative confidence; and columns 4-6 present the effect of absolute confidence. Columns 2 and 3 look at the effect of relative confidence fixing the quiz difficulty, easy quiz for column 2 and hard quiz for column 3. Columns 5 and 6 look at the effect of absolute confidence fixing the quiz difficulty, easy quiz for column 5 and hard quiz for column 6. All specifications control for session and order effects.
Standard errors clustered at the subject level in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2.3 Absolute Confidence

Unlike relative confidence which may change between phases for a given subject, absolute confidence is fixed in a quiz block. Recall, it is measured as the difference, in expectation, between the subject's own score and their actual score for the quiz block. From the illustration in Section 2, subjects that exhibit absolute underconfidence maintain their belief that other subjects have lower accuracy even after realizing they have a higher score than expected. Analogously, those that exhibit absolute overconfidence

Table 4: Effects of Confidence on Welfare

	(1)	(2)	(3)	(4)	(5)	(6)
Hard	-0.648*			-0.645*		
	(0.337)			(0.334)		
Score Known	-0.0855	-0.0735	-0.00869	-0.0849	-0.0387	0.0235
	(0.0693)	(0.0739)	(0.0474)	(0.0727)	(0.0728)	(0.0471)
Hard x Score Known	0.0946			0.100		
	(0.0858)			(0.0908)		
Ave Others' Score	-0.160***	-0.187***	-0.104*	-0.162***	-0.168**	-0.0983
	(0.0505)	(0.0694)	(0.0618)	(0.0511)	(0.0661)	(0.0641)
Relative Confidence	0.0721***	0.0889***	0.0393*			
	(0.0161)	(0.0288)	(0.0224)			
E_i(other score)	0.0622***	0.102***	0.0174			
	(0.0188)	(0.0290)	(0.0260)			
Absolute Confidence				-0.0774***	0.000883	-0.0896***
				(0.0248)	(0.0477)	(0.0303)
E_i(own score)				0.0660***	0.110***	0.0199
				(0.0155)	(0.0222)	(0.0218)
Constant	1.037**	1.022	0.349	1.022**	0.816	0.324
	(0.495)	(0.696)	(0.218)	(0.484)	(0.655)	(0.212)
Observations	916	317	599	916	317	599

The table looks at the effects of several variables of interest on welfare change. Columns 1-3 columns present the effect of relative confidence; and columns 4-6 present the effect of absolute confidence. Columns 2 and 3 look at the effect of relative confidence fixing the quiz difficulty, easy quiz for column 2 and hard quiz for column 3. Columns 5 and 6 look at the effect of absolute confidence fixing the quiz difficulty, easy quiz for column 5 and hard quiz for column 6. All specifications control for session and order effects.
Standard errors in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

maintain their belief that others have higher accuracy after realizing their score is lower than expected. This translates to a higher degree of relative confidence after scores are revealed for those subjects that exhibit absolute underconfidence and a lower degree of relative confidence for those that exhibit absolute overconfidence. Hence, we expect that those exhibiting absolute overconfidence to follow their signals less while those exhibiting underconfidence to follow their signals more after scores are revealed.

We observe that absolute confidence has an effect on herd breaking behavior after subjects realize their own score. For the Easy quiz block in Figure 6, subjects that are underconfident about their absolute performance initially follow their signal less as indicated by the upward sloping solid line. However, after realizing their true score, those that are initially underconfident follow their signal more as indicated by the flatter dashed line. This pattern is also observed in the Hard quiz block. Here, subjects that are overconfident in their absolute quiz performance do not generally follow their signal any differently than those that are less confident about their absolute performance, which is depicted by the weakly decreasing solid line. Once they realize their true score, those subject that are initially overconfident start to break herds less than those that are initially underconfident as shown by the steeper downward sloping dashed line. Columns 4-6 in Table 3 verify these findings. Pooling for both quiz blocks, we have a significant negative effect of absolute confidence on herd breaking behavior. That is, a 1 point increase in the measured degree of absolute confidence translates to a drop of 5.29 percentage points ($p < 0.001$) in the likelihood of following a signal. This effect is pronounced in the Hard quiz block ($p < 0.05$) but not in the Easy quiz block ($p > 0.1$).

Figure 6: Following Signals and Absolute Confidence

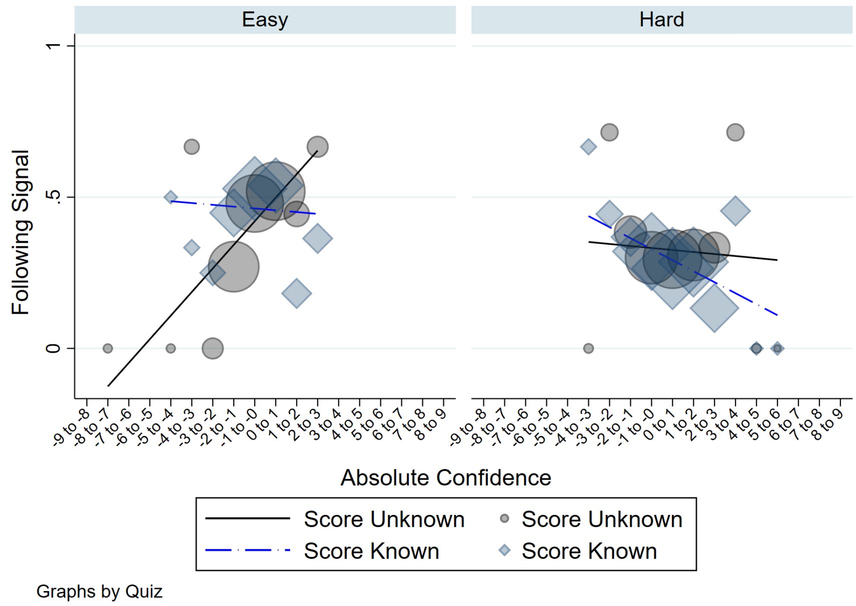
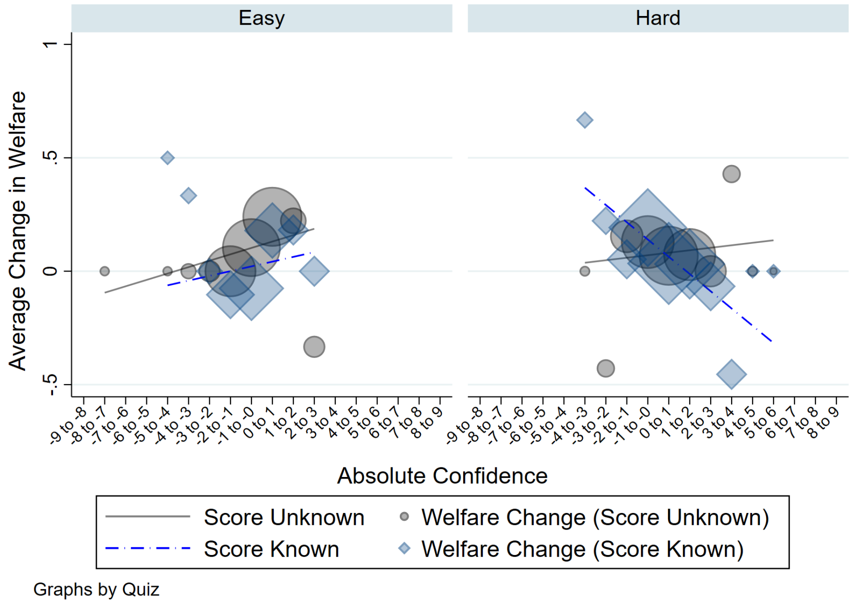


Figure 7 plots the measure of absolute confidence against the change in welfare for Easy and Hard quiz blocks. The solid line shows the positive relationship between absolute confidence and welfare change in the Score Unknown phase. This positive relationship weakens in the Easy quiz blocks and reverses in the Hard quiz blocks in the Score Known phase as illustrated by the dashed line. Columns 4 to 6 in Table 4 quantify this relationship when controlling for the average score of the group. In particular, Column 4 shows a significant negative effect of absolute confidence on welfare change indicating that subjects with a higher degree of absolute confidence are less likely to move away from an incorrect herd after realizing their performance. Confirming the qualitative findings in the figure, Column 5 shows that there is no significant relationship between welfare change and absolute overconfidence in Easy quiz blocks in the Score Known phase. In Column 6, we find this negative relationship in the Hard quiz blocks wherein a 1 point increase in the measure of absolute overconfidence translates to an 8.96 percentage point decrease when a subject follow their signal against the herd.[21]

[21]The fourth multinomial logit specification in E.2 similarly verifies that subject's that are more confident in absolute terms are less likely to move away from incorrect herds after realizing their own score.

Figure 7: Absolute Confidence on Welfare



Result 4. *With respect to the degree of absolute confidence:*

a) Subjects that exhibit more confidence about their absolute quiz performance are less likely to break herds after observing their own score.

b) After realizing their quiz performance, subjects that exhibit more confidence about their absolute performance are less likely to break away from incorrect herds particularly in hard quizzes.

4.3 Herd Behavior and Information Aggregation

In the following, we ask how does overconfidence affect the formation of herds and information aggregation. We define the length of a herd to be the length of a sequence of publicly observed actions that are the same. For instance, if we see a sequence of R, R, R, B, B, R, then we say that a herd of length three started from the first mover and a herd of length two started from the fourth mover.

4.3.1 Herd formation in a sequence

To understand the structures of herds in a sequence, we look at three indications of herd formation: herds that start from the first mover, herds that end at the last mover, and the total number of switches in a sequence. As an example, consider again the sequence above. We say that there is a herd of length 3 that starts from the first mover, a herd of length 1 that ends at the 6th mover and a total of 3 switches in the sequence. Tables 5, 6, and 7 provide the frequencies of the lengths and switches for a given block.

Table 5 tabulates the frequencies of herd lengths starting from the first mover, which we interpret as a measure of the persistence of herds. We see that about half of the herds are short, with lengths of at

Max Herd Length	Control	Easy	Hard	Total
1	41.67	32.41	40.57	37.63
2	17.98	13.19	14.69	14.78
3	8.77	7.87	9.21	8.60
4	3.07	7.18	5.48	5.65
5	5.26	5.32	4.61	5.02
6	23.25	34.03	25.44	28.32

Table 5: Percentage of Herd Lengths beginning at the first mover

Max Herd Length	Control	Easy	Hard	Total
1	28.95	25.23	29.17	27.60
2	17.11	12.50	14.25	14.16
3	12.72	10.42	10.75	11.02
4	8.77	8.80	8.99	8.87
5	9.21	9.03	11.40	10.04
6	23.25	34.03	25.44	28.32

Table 6: Percentage of Herd Lengths ending at the 6th mover

Number of switches	Control	Easy	Hard	Total
1	23.25	34.03	25.44	28.32
2	16.67	18.06	19.30	18.28
3	33.33	30.56	31.58	31.54
4	15.35	10.42	12.72	12.37
5	10.53	5.79	9.21	8.15
6	0.88	1.16	1.75	1.34

Table 7: Percentage of number of switches

most two. Short herds are more prominent in Control and Hard quiz blocks than in Easy quiz blocks. Among the longer herds, the most observed herd length is 6, which is more prominent in the Easy quiz blocks than in Control or Hard quiz blocks. Regressing the length on herds on block dummies verifies that, on average, Easy quiz blocks have longer herds that start from the first mover than the Control block (F-test, $p < 0.01$) and Hard quiz block (F-test, $p < 0.01$), respectively.

We see a similar pattern when looking at herds that end at the final mover. In Table 6, the herd lengths are longer in Easy quiz blocks than in Control and Hard quiz blocks. Although most of the herds are either of length 1 or 6 in the aggregate, the modal length differs when looking at the blocks. Control and Hard quiz blocks have the modal length of 1 while Easy quiz blocks have a modal length of 6. Qualitatively, this says that herds are more persistent in Easy quiz blocks and that final movers are seemingly switching more often in Control and Hard quiz blocks.

Table 7 looks at a complementary point where it tabulates the total number of switches of sequences in a block. We see that the modal number of switches in the Control and Hard quiz blocks is three while the modal number of switches is 1 for the Easy quiz block. Running a simple regression of the number of switches on the block dummies show that the Easy quiz blocks have less switches than the Control block (F-test, $p < 0.01$) and the Hard quiz block (F-test, $p < 0.01$). Furthermore, the average number of switches in Control and Hard quiz blocks are not statistically different (F-test, $p > 0.1$).

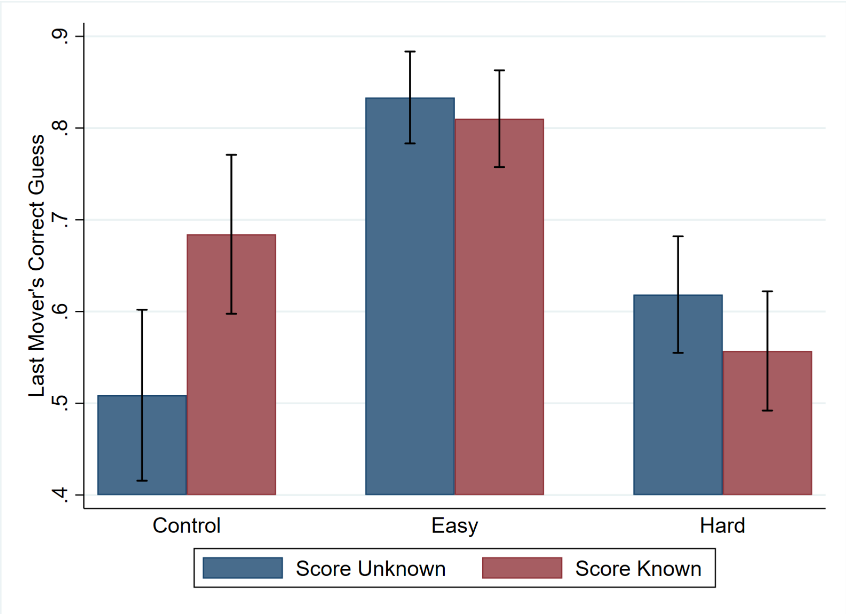
Comparing these results to the individual herd breaking behavior, it is seemingly inconsistent that we find more subjects following their own sign in Easy quiz blocks, but also have shorter herds in the Control or Hard quiz blocks. That is, we observe longer herds in the Easy quiz block even though subjects follow their signals more often. One possible resolution is that signals that contradict both the immediate predecessor and initial guess are less often observed in Easy (18.34%) than in Control (29.39%) and Hard (32.84%).

Result 5. *Herds are consistently longer in Easy quiz blocks than the other blocks. Furthermore, there are fewer switches in Easy quiz block than the other blocks. There is no significant difference of herd lengths nor switches between the Control and Hard quiz blocks.*

4.3.2 Information Aggregation

This section looks at the effects of the treatment variations on the degree of information aggregation. We are interested in the rate at which the final mover in the sequence aggregates information and correctly guesses the state. Figure 8 compares the rate at which the final mover guesses the state correctly, conditioning on the blocks and the phase. There are striking differences in behavior when comparing the Control blocks to the Quiz blocks. Firstly, when moving from the Score Unknown phase to the Known phase, the proportion of final movers guessing correctly significantly increases in the Control block. The pattern is reversed, however, for the Easy and Hard quiz blocks. Regression results show that this change is significant in the Control block but not in the Easy and Hard blocks.

Figure 8: Information Aggregation



In comparing across the blocks, the 6th mover is, on average, correctly guessing the state most often in the Easy quiz blocks. This is supported by the regressions in Table 8, columns 1 and 3. Pooling for both phases, the final mover guesses correctly 82.18% of the time in Easy quiz, 59.65% in Control, and 58.77% in Hard quizzes. When comparing behavior across the Control and Hard blocks, the final mover correctly guesses the state 10.06% more ($\chi^2,\ p < 0.1$) in the Hard quizzes in the Score Unknown phase. This trend, however, reverses in the Score Known phase, where the final mover correctly guesses 12.72% more ($\chi^2,\ p < 0.05$) in the Control block. Column 2 in Table 8 provides support that these differences are significant. It is worth to note the comparison between the Hard quiz and Control blocks. Before subjects know their own score, the last movers in the Hard quiz block correctly guess the color 61.8% of the time compared to 50.8% in the Control block. However, the average accuracy in the Hard quiz block is significantly lower (56.5%) compared to the average accuracy in the Control block (67.8%). This suggests that overconfidence can have a positive information externality on subsequent movers.

Table 8: Effects of Quiz on Information Aggregation

	Control v Easy	Control v Hard	Easy v Hard
Easy	0.328***		
	(0.0545)		
Score Known	0.175***	0.175***	-0.0231
	(0.0640)	(0.0648)	(0.0372)
Easy x Score Known	-0.199***		
	(0.0731)		
Hard		0.110*	-0.217***
		(0.0628)	(0.0395)
Hard x Score Known		-0.237***	-0.0383
		(0.0771)	(0.0568)
Constant	0.496***	0.626***	0.859***
	(0.0629)	(0.0615)	(0.0460)
Observations	660	684	888

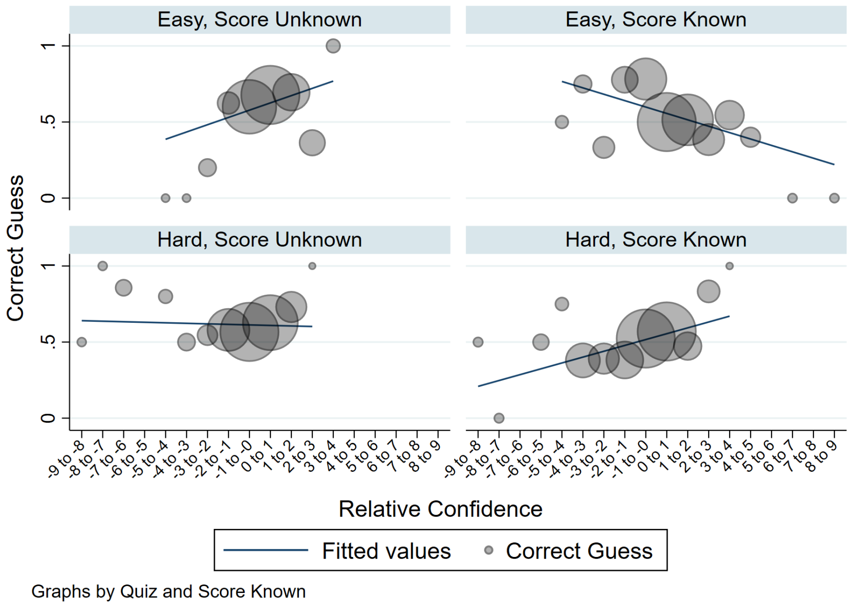
Regressions comparing the effect of Quiz and Phases on the last mover guessing correctly. Standard errors clustered at the group-phase level in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Result 6. *The final movers in a sequence correctly guess the state the most in Easy quiz blocks, second most in the Hard quiz blocks, and least in the Control block when the scores have not been revealed. After revealing their own score, they correctly guess the state most often in Easy quiz blocks, then second in the Control block, and least in the Hard quiz blocks. In the Control block, they correctly guess the state more often after realizing their own score. There is a negative differential effect of knowing the score in the Quiz blocks versus the Control block.*

4.3.3 Individual Final Guesses

Figure 9 shows the relationship between relative overconfidence and the rate of subjects correctly guessing the state across Quizzes and Phases. When looking at Easy quiz blocks, we find a positive and significant relationship between relative overconfidence and guessing the state correctly (logistic regression, $p < 0.1$) when subjects do not know their own score. In contrast, relative confidence and correct guesses have a negative and significant relationship in Easy quizzes when subjects know their own score (logistic regression, $p < 0.1$). The results are qualitatively different in Hard quizzes. We find no significant relationship between relative confidence and correctly guessing the state when subjects do not know their own score (logistic regression, $p > 0.1$). On the other hand, there is a positive and significant relationship between relative confidence and correct guessing when subjects know their own score (logistic regression, $p < 0.05$).

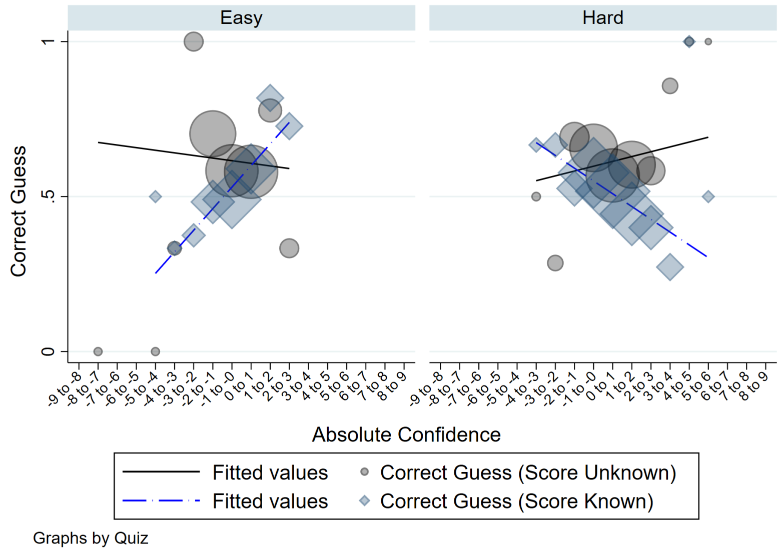
Figure 9: Relative Confidence on Final Guesses



Result 7. *In Easy quiz blocks, subjects that exhibit more confidence about their relative performance correctly guess the state more often before knowing their own score and less often after knowing their score. In Hard quiz blocks, subjects that exhibit more relative confidence guess the state more often after knowing their score. However, it does not affect the chance of correctly guessing the state before knowing their score.*

For absolute confidence, recall that the degree of absolute overconfidence for each subject is fixed within each block. Figure 10 plots the relationship between absolute confidence and guessing the correct state for Easy and Hard quiz blocks. In Easy quizzes, for a fixed level of absolute *underconfidence*, subjects are more likely to guess correctly when they do not know their own score. For a fixed level of absolute *overconfidence*, subjects are more likely to guess correctly when they know their own score. Furthermore, the effect becomes more pronounced as subjects exhibit higher degrees of absolute confidence (logistic regression, $p < 0.1$). The qualitative patterns reverse for Hard quizzes. Subjects who exhibit absolute overconfidence are less likely to guess correctly when they know their own score. Subjects who are absolutely underconfident are more likely to guess correctly when they know their own score. The

Figure 10: Absolute Confidence on Final Guesses



difference becomes more pronounced in the degree of absolute confidence (logistic regression, $p < 0.1$).

Result 8. *After realizing their quiz performance, subjects that exhibit more confidence about their absolute performance correctly guess the state more often in the Easy quiz blocks and less often in Hard quiz blocks.*

5 Discussion

Our results show that overconfidence can improve welfare in social learning settings. We do so by allowing subjects to form beliefs about their own and others' information based on their scores in a trivia quiz. Their beliefs on expected performance allow us to measure and study the effects of the two forms of confidence–relative and absolute confidence–on their welfare and belief updating behavior. Our experimental findings show that subjects exhibiting a higher degree of relative overconfidence are more likely to move away from incorrect herds. After revealing their own performance on the quiz, subjects that have overestimated their absolute performance are more likely to move away from correct herds than those who have underestimated their absolute performance. Hence, the welfare gain or loss from overconfidence depends on the type confidence the subject exhibits. We find that these results coincide with changes in herd breaking behavior. Subjects exhibiting more confidence about their relative quiz performance followed their signals more often. Subjects that have overestimated their absolute performance followed their signals less often after they realize their true performance. Overall, the easier quiz difficulty, which induced relative overconfidence and absolute underconfidence, leads to higher information aggregation.

Our paper departs from previous experimental works on overconfidence in social learning in several key ways. The main distinction is that previous research looks at "intrinsic" overconfidence in the sense that subjects tend to overweigh their private signal more than they should, even though the signal accuracy is assigned exogenously. This overconfidence can be thought as subjects assuming that they are better than others in processing the same information (Angrisani et al., 2021). This overconfidence leads to more information aggregation, but is costly at the individual level since subjects tend to break away from correct herds on average. In contrast, our paper *induces* overconfidence in a more natural way that manipulates the beliefs subjects have about their own relative expertise. Essentially, the subjects who exhibit overconfidence are the ones who themselves have high signal accuracy (experts). This leads to welfare gains, not only at the social level due to the positive information externality, but also at the individual level since the experts are the ones who are most likely to break away from herds.

A point of note is that most of the benefit from overconfidence is observed in settings where subjects do not know their own signal accuracy. One conjecture for this is that subjects are more moderate in their degree of relative overconfidence. Not knowing their own score tempers their expectations about other peoples' scores. This "cautious" overconfidence moderates the degree to which they break away from herds. On the other hand, the degree of overconfidence is exacerbated when subjects know their own score. This removes the tempering effect causing subjects to overweight their private signals more and break away from potentially correct herds. This is supported in Appendix B where we look at the the updating behavior of subjects. Subjects update their beliefs more extremely the more they exhibit higher degrees of relative confidence, and more so after they realized that they have underestimated their performance.

References

Agranov, M., Lopez-Moctezuma, G., Strack, P., and Tamuz, O. (2022). Learning through imitation: an experiment. Technical report, National Bureau of Economic Research.

Anderson, L. R. and Holt, C. A. (1997). Information cascades in the laboratory. *The American Economic Review*, 87(5):847–862.

Angrisani, M., Guarino, A., Jehiel, P., and Kitagawa, T. (2021). Information redundancy neglect versus overconfidence: a social learning experiment. *American Economic Journal: Microeconomics*, 13(3):163–97.

Banerjee, A. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Ben-David, I., Graham, J. R., and Harvey, C. R. (2013). Managerial miscalibration. *The Quarterly journal of economics*, 128(4):1547–1584.

Benoît, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Bernardo, A. E. and Welch, I. (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, 10(3):301–330.

Bikhchandani, S., Hirshleifer, D., Tamuz, O., and Welch, I. (2021). Information cascades and social learning. Technical report, National Bureau of Economic Research.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy*, 100(5):992–1026.

Bohren, J. A. and Hauser, D. N. (2021). Learning with heterogeneous misspecified models: Characterization and robustness. *Econometrica*, 89(6):3025–3077.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American economic review*, 89(1):306–318.

Çelen, B. and Kariv, S. (2004). Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, 94(3):484–498.

Chandrasekhar, A. G., Larreguy, H., and Xandri, J. P. (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, 88(1):1–32.

Chen, D., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9.

Chen, J. Y. (2022). Sequential Learning under Informational Ambiguity. SSRN Scholarly Paper ID 3480231, Social Science Research Network, Rochester, NY.

Cipriani, M. and Guarino, A. (2005). Herd behavior in a laboratory financial market. *American Economic Review*, 95(5):1427–1443.

Dasaratha, K. and He, K. (2021). An experiment on network density and sequential learning. *Games and Economic Behavior*, 128:182–192.

De Filippis, R., Guarino, A., Jehiel, P., and Kitagawa, T. (2022). Non-bayesian updating in a social learning experiment. *Journal of Economic Theory*, 199:105188.

Duffy, J., Hopkins, E., and Kornienko, T. (2021). Lone wolf or herd animal? information choice and learning from others. *European Economic Review*, 134:103690.

Eyster, E. and Rabin, M. (2010). Naive herding in rich-information settings. *American economic journal: microeconomics*, 2(4):221–43.

Eyster, E., Rabin, M., and Weizsacker, G. (2015). An experiment on social mislearning. *Available at SSRN 2704746*.

Goeree, J. K., Palfrey, T. R., Rogers, B. W., and McKelvey, R. D. (2007). Self-correcting information cascades. *The Review of Economic Studies*, 74(3):733–762.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.

Grimm, V. and Mengel, F. (2020). Experiments on belief formation in networks. *Journal of the European Economic Association*, 18(1):49–82.

Grubb, M. D. (2015). Overconfident consumers in the marketplace. *Journal of Economic Perspectives*, 29(4):9–36.

Holt, C. A. and Smith, A. M. (2016). Belief elicitation with a synchronized lottery choice menu that is invariant to risk attitudes. *American Economic Journal: Microeconomics*, 8(1):110–39.

Huffman, D., Raymond, C., and Shvets, J. (2022). Persistent overconfidence and biased memory: Evidence from managers. *American Economic Review*, 112(10):3141–75.

Kartal, M. and Tyran, J.-R. (2022). Fake news, voter overconfidence, and the quality of democratic choice. *American Economic Review*, 112(10):3367–97.

Kogan, S. (2009). Distinguishing the effect of overconfidence from rational best-response on information aggregation. *The Review of Financial Studies*, 22(5):1889–1914.

Kübler, D. and Weizsäcker, G. (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies*, 71(2):425–441.

Kübler, D. and Weizsäcker, G. (2005). Are longer cascades more stable? *Journal of the European Economic Association*, 3(2-3):330–339.

Malmendier, U. and Tate, G. (2005). Ceo overconfidence and corporate investment. *The journal of finance*, 60(6):2661–2700.

March, C. and Ziegelmeyer, A. (2020). Altruistic observational learning. *Journal of Economic Theory*, 190:105123.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2022). Managing self-confidence: Theory and experimental evidence. *Management Science*.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological review*, 115(2):502.

Nöth, M. and Weber, M. (2003). Information aggregation with random ordering: Cascades and overconfidence. *The Economic Journal*, 113(484):166–189.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The journal of finance*, 53(6):1887–1934.

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2):504–35.

Serra-Garcia, M. and Gneezy, U. (2021). Mistakes, overconfidence, and the effect of sharing on detecting lies. *American Economic Review*, 111(10):3160–83.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta psychologica*, 47(2):143–148.

Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, 100(5):2340–60.

Wu, J. (2015). Helpful laymen in informational cascades. *Journal of Economic Behavior & Organization*, 116:407–415.

Yin, X., Li, J., and Bao, T. (2019). Does overconfidence promote cooperation? theory and experimental evidence. *Journal of Behavioral and Experimental Economics*, 79:119–133.

Appendix A A model of overconfidence in social learning

We consider a sequential social learning model where a player has beliefs about their own accuracy and the accuracy of other players. The environment has an infinite sequence of players that sequentially make guess about the state of the world $\omega \in \{R, B\}$ after observing their own private signal $s \in \{r, b\}$ and the history of their predecessors' guesses. Guessing correctly earns the player a payoff of one and guessing incorrectly gives zero. Each guess made is publicly observed but the private signal and individual payoff is observed by the player that made the guess. Each private signal can have an accuracy q^L, q^M, or q^H where $\frac{1}{2} < q^L < q^M < q^H$. The accuracy of a generic player i's private signal depends on the signal environment $E \in \{E_1, E_2\}$. In particular, the accuracy of player i q_i in E_1 first-order stochastically dominates accuracy q_i in E_2 where $\Pr(q_i \geq x|E_1) \geq \Pr(q_i \geq x|E_2)$ for any $x \in R$ and $\Pr(q_i \geq x|E_1) > \Pr(q_i \geq x|E_2)$ for some x. Hence, the average accuracy is higher in environment E_1 than E_2.

Table 9 shows the joint distribution of accuracy and signal environment we consider. We say that players have symmetric beliefs if $\Pr(q_i = x, q_j = y|E) = \Pr(q_i = y, q_j = x|E)$. Following Table 9, a symmetric belief implies $\alpha_2 = \beta_1$, $\alpha_3 = \gamma_1$, and $\beta_3 = \gamma_1$.

q_i	q_j	E_1	E_2
q^H	q^H	α_1	γ_3
q^H	q^M	α_2	γ_2
q^H	q^L	α_3	γ_1
q^M	q^H	β_1	β_3
q^M	q^M	β_2	β_2
q^M	q^L	β_3	β_1
q^L	q^H	γ_1	α_3
q^L	q^M	γ_2	α_2
q^L	q^L	γ_3	α_1

Table 9: Joint distribution of accuracy and environment

There are two different levels of informativeness, either players know the signal environment or know the signal environment as well as their own accuracy. In the first level of informativeness, player i has the same expected accuracy for his own and a random player j in the same signal environment where $\mathbb{E}(q_i|E) = \mathbb{E}(q_j|E)$. In the second level of informativeness, subject i is also informed about their own accuracy q_i and may use this to infer the average accuracy of the players in the same environment $\mathbb{E}(q_j|E, q_i)$. We allow the case where player i's own accuracy does not inform him of the average player's accuracy given the environment, which gives $\mathbb{E}(q_j|E, q_i) = \mathbb{E}(q_j|E)$.[22] In the more general case of $\mathbb{E}(q_j|E, q_i) \neq \mathbb{E}(q_j|E)$, we show that overconfidence can affect the public belief and the formation of informational cascades.

[22]This case follows the setup of Wu (2015) where there is a known distribution of high accuracy and low accuracy type players. In their model, a player's accuracy is independent from another player's accuracy. We show that assuming independence of accuracies cannot explain the change in public likelihood ratios after players observe their own accuracy.

We say that a sequence of n consecutive identical actions is a *herd of length n* and a herd of arbitrary length as a *herd*. An informational cascade for type q is said to have occurred if it is optimal for the player of type q to follow the herd for any possible signal realization. A full informational cascade has occurred if an informational cascade has occurred for all types.

We denote the public likelihood ratio after observing the actions of the first n players conditioned on the signal environment E and observer's type q^i as l_{n,E,q^i} where $l_{n,E,q^i} = \frac{\Pr(A_1,A_2,...,A_n|R,E,q^i)}{\Pr(A_1,A_2,...,A_n|B,E,q^i)}$. If there is no confusion, we simply denote this by l_n. The dynamics of l_n depends on the combination of accuracy types and signal realizations. There are six relevant thresholds for public likelihood ratio conditional on the true state being R. If l_n is greater than the first threshold $\frac{q^H}{1-q^H}$, there is no accuracy type and signal realization that can overturn the public belief. Any player type in position $n+1$ is in a cascade where it is optimal to choose R. Hence, the action of the player in position $n+1$ is no longer informative and $l_{n+1} = l_n$. If l_n is in between thresholds $\frac{q^H}{1-q^H}$ and $\frac{q^M}{1-q^M}$, then only a type q^H with signal b can optimally choose B with the corresponding public likelihood ratio $l_{n+1} = l_n \frac{1-q^H}{q^H}$. If instead action R in position $n+1$ is observed, then this happens probability $1 - \Pr(q^H|E, q^i)q^H$, i.e., the case where the player in position $n+1$ is not a high type with signal b. The corresponding public likelihood ratio is $l_{n+1} = l_n \frac{1-\Pr(q^H|E,q^i)(1-q^H)}{1-\Pr(q^H|E,q^i)q^H}$. If l_n is in between thresholds $\frac{q^M}{1-q^M}$ and $\frac{q^L}{1-q^L}$, then types q^H and q^M with signal b can optimally choose B. This case happens with probability $\Pr(q^H|E, q^i)(1-q^H) + \Pr(q^M|E, q^i)(1-q^M)$. The corresponding public likelihood ratio is $l_{n+1} = l_n \frac{\Pr(q^H|E,q^i)(1-q^H)+\Pr(q^M|E,q^i)(1-q^M)}{\Pr(q^H|E,q^i)q^H+\Pr(q^M|E,q^i)q^M}$. If R is observed in position $n+1$ then $l_{n+1} = l_n \frac{1-\Pr(q^H|E,q^i)(1-q^H)-\Pr(q^M|E,q^i)(1-q^M)}{1-\Pr(q^H|E,q^i)q^H-\Pr(q^M|E,q^i)q^M}$. If l_n is in between $\frac{q^L}{1-q^L}$ and $\frac{1-q^L}{q^L}$, then no type is in a cascade and the player in position $n+1$ will choose according to their realized signal. If R is observed in position $n+1$ then $l_{n+1} = l_n \frac{\mathbb{E}(q|E,q^i)}{1-\mathbb{E}(q|E,q^i)}$. If B is observed in position $n+1$ then $l_{n+1} = l_n \frac{1-\mathbb{E}(q|E,q^i)}{\mathbb{E}(q|E,q^i)}$. Analogous arguments hold for l_n between $\frac{1-q^L}{q^L}$ and $\frac{1-q^M}{q^M}$, $\frac{1-q^M}{q^M}$ and $\frac{1-q^H}{q^H}$, and above $\frac{1-q^H}{q^H}$.

The following results are extensions of the results in Wu (2015). In particular, our setup introduces a middle type and allows for different beliefs over the accuracy of other players given their own type.

Proposition 1. *Consider a herd that starts with the player in the first position. If players have symmetric beliefs conditional on the signal environment, then a full informational cascade occurs after a herd of two in the first level of informativeness. In the second level of informativeness, type q^L is in a cascade after a herd of length 1, type q^M is in a cascade after a herd of length M, and type q^H is in a cascade after a herd of length H where M is the smallest integer that satisfies*

$$M \geq \frac{\ln\left(\frac{q^M}{1-q^M}\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}\right)}{\ln\left(\frac{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}\right)} + 1 \tag{1}$$

and H is the smallest integer that satisfies

$$H \geq \frac{\ln\left(\frac{q^H}{1-q^H}\left(\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}\right)\left(\frac{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}\right)^{M-1}\right)}{\ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)} + M \tag{2}$$

Proof. Consider a the first level of informativeness where players only observe the signal environment. Suppose that players symmetric beliefs where $\Pr(q_i = x, q_j = y|E) = \Pr(q_i = y, q_j = x|E)$. For player i, his expected accuracy $E(q_i|E)$ is equal to his expectation about an arbitrary player j's accuracy $E(q_j|E)$. Consider a herd of length two with guess R. The public likelihood ratio conditional on signal environment E is then

$$\left(\frac{\mathbb{E}(q_j|E)}{1-\mathbb{E}(q_j|E)}\right)\left(\frac{1-\frac{1}{2}(1-\mathbb{E}(q_j|E))}{1-\frac{1}{2}(\mathbb{E}(q_j|E))}\right).$$

The first factor is the likelihood ratio that any type has a received a signal R. This is equal to $\left(\frac{\mathbb{E}(q_i|E)}{1-\mathbb{E}(q_i|E)}\right)$ since players have symmetric beliefs. The second factor is likelihood ratio based on the event that any player type that has not recieved a signal B and chose R based on the tie-breaking rule with equal probability. This second factor is greater than 1 for any nondegenerate distribution over q_j conditional on E. Hence, player i in the third position observes that it is optimal to follow the herd of two with guess R regardless of his realized private signal. The same argument holds for a herd of two with guess B.

Now, consider the the second level of informativeness where players observe the signal environment and their own accuracy type. For the lowest accuracy type q^L in the second position, he observers a herd of one with guess R so the public likelihood ratio conditional on his type $q_i = q^L$ and signal environment E is

$$l_1 = \frac{\mathbb{E}(q_j|E, q^L)}{1-\mathbb{E}(q_j|E, q^L)}.$$

Notice that l_1 is greater than $\left(\frac{q^L}{1-q^L}\right)$ for any nondegenerate distribution over q_j given the signal environment. This implies that an information cascade has occurred for type q^L in the second position.

Let the player i in the $M+1$ position be an accuracy type q^M. The public likelihood ratio after a herd of length M with guess R conditional on $q_i = q^M$ and signal environment E is

$$l_M = \left(\frac{E(q_j|E,q^L)}{1-E(q_j|E,q^L)}\right)\left(\frac{1-\Pr(q_j^H|E,q^M)(1-q^H)-\Pr(q_j^M|E,q^M)(1-q^M)}{1-\Pr(q_j^H|E,q^M)q^H-\Pr(q_j^M|E,q^M)q^M}\right)^{M-1}.$$

The first factor is assessed based on the expectations of type q^L and the second factor is assessed based on the expectations of type q^M.[23] The condition needed for an informational cascade to occur for type

[23]In computing the conditional public likelihood ratio for a given type q_i in some position, we assume that q_i assesses the occurrence of a cascade based on the expectations of types weakly lower than their own type. For instance, a type q^L in the second position is already in a cascade. Type q^M players understand this and assess the public likelihood ratio for the first position based on the expectations of type q^L that caused this cascade. The remaining actions are then assessed based on type q^M's beliefs.

q^M is the inequality $l_M > \frac{q^M}{1-q^M}$. This condition is satisfied if and only if Equation 1 is satisfied. To see this, we have

$$\begin{aligned}
& l_M > \frac{q^M}{1-q^M} \\
\iff & \left(\frac{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}\right)^{(M-1)} > \left(\frac{q^M}{1-q^M}\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}\right) \\
\iff & (M-1)\ln\left(\frac{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}\right) > \ln\left(\frac{q^M}{1-q^M}\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}\right) \\
\iff & M > \frac{\ln\left(\frac{q^M}{1-q^M}\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}\right)}{\ln\left(\frac{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}\right)}+1
\end{aligned}$$

where the second line follows from multiplying $\frac{1-\mathbb{E}(q_j|E,q^L)}{\mathbb{E}(q_j|E,q^L)}$ on both sides, the third line follows from taking the natural log of both sides, and the last line follows from isolating M on the left-hand side.

An analogous argument holds for type q^H. Let the player in the $H+1$ position have an accuracy type q^H. The public likelihood ratio after a herd of length H with guess R is

$$l_H = l_M \cdot \left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)^{H-M}.$$

The condition needed for an informational cascade to occur for type q^H is the inequality $l_H > \frac{q^H}{1-q^H}$. This condition is satisfied if and only if Equation 2 is satisfied. To see this, we have

$$\begin{aligned}
& l_H > \frac{q^H}{1-q^h} \\
\iff & \left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)^{H-M} > \left(\frac{q^H}{1-q^H}\frac{1}{l_M}\right) \\
\iff & (H-M)\ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right) > \ln\left(\frac{q^H}{1-q^H}\frac{1}{l_M}\right) \\
\iff & H > \frac{\ln\left(\frac{q^H}{1-q^H}\frac{1}{l_M}\right)}{\ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)}+M
\end{aligned}$$

where the second line follows from multiplying $\frac{1}{l_M}$ on both sides, the third line follows from taking the natural log of both sides, and the last line follows from isolating H on the left-hand side. □

Remark 1. For a nondegenerate joint distribution of accuracy types and signal environments, it follows from Equations 1 and 2 that the necessary and sufficient condition for $H > M > 1$ is the ordering

$$\frac{q^H}{1-q^H} > l_M > \frac{q^M}{1-q^M} > l_1.$$

Furthermore, when $M = 1$ we have $l_M = l_1$. Since $\frac{q^H}{1-q^H} > \frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)}$, H is at least 2. Hence, a full cascade may occur after a herd of length 2.

Proposition 2. *Suppose a herd of length H starting from the player in the first position is needed for a full cascade to occur. Let the public likelihood ratio in position k conditional on signal environment E and type q^L be $l_k \in \left[\frac{1-q^L}{q^L}, \frac{q^L}{1-q^L}\right]$. A herd of at least H length and at most $H+1$ starting from position $k+1$ is needed for a full cascade to occur.*

Proof. We prove the statement for a herd with guess R. An analogous proof can be made with a herd with guess B. If l_k is equal to one, then exactly a herd of length H leads to a full cascade. First, consider $l_k \in \left(1, \frac{q^L}{1-q^l}\right]$. Then, a cascade has occurred for type q^L in position $k+2$ since we have

$$l_{k+1} = l_k \frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)} > \frac{q^L}{1-q^L}$$

where the inequality follows from $l_k > 1$ and $\frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)} > \frac{q^L}{1-q^L}$. A similar argument follow for types q^M and q^H where a cascade has occurred for type q^M in position $k+M+1$ and for type q^H in position $k+H+1$. Therefore a full cascade occurs after a herd of length H.

Now, consider $l_k \in \left[\frac{1-q^L}{q^L}, 1\right)$. A type q^L in position $k+2$ observes a guess R in position $k+1$. The the public likelihood ratio in l_{k+1} is

$$l_{k+1} = l_k \frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)}.$$

However, it is not necessarily the case that l_{k+1} is greater than $\frac{q^L}{1-L}$. The first case is $l_{k+1} > \frac{q^L}{1-L}$. In this case, the previous argument holds. The second case is $l_{k+1} \leq \frac{q^L}{1-L}$. Here, a cascade for type q^L in position $k+2$ has not occurred. Consider a type q^L in position $k+3$. The public likelihood ratio after observing a herd of two is

$$l_{k+2} = l_k \left(\frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)}\right)^2 = l_{k+1} \frac{\mathbb{E}(q_j|E,q_i^L)}{1-\mathbb{E}(q_j|E,q_i^L)}.$$

Since, l_{k+1} is weakly greater than 1, $l_{k+2} > \frac{q^L}{1-L}$. Hence, a cascade occurs for type q^L in position $k+3$. It follows that a cascade occurs for type q^M in position $k+M+2$ and for type q^H in position $k+H+2$. Therefore a full cascade occurs after a herd of length $H+1$. □

Lemma 1. *Given that a full informational cascade has occurred from a herd that started from the first mover, the* *log odds probability that it is correct is given by the equation*

$$\begin{aligned}\ln\left(\frac{\Pr(Correct|Cascade)}{1-\Pr(Correct|Cascade)}\right) &= \ln\frac{\mathbb{E}(q_j|E,q^L)}{1-\mathbb{E}(q_j|E,q^L)} \\ &+ (M-1)\ln\left(\frac{1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)}{1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M}\right) \\ &+ (H-M)\ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right),\end{aligned} \tag{3}$$

where H and M are the integers that satisfy Condition 2 and Condition 1.

Proof. Let the true state be R. An analogous argument holds for B. Consider a herd of length H with action R. The probability that an agent chooses R for the first mover is given by $\mathbb{E}(q_j|E,q^L)$, i.e., the probability that that the first mover receives a r signal conditional that the true state is R. The probability that the second to the M-th mover choose R is given by $\left((1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)\right)^{M-1}$, which is the probability that these $M-1$ movers are not q^M nor q^H types with signal b. Lastly, the probability that $M+1$-th mover to the H-th mover choose R is $\left(1-\Pr(q_j^H|E,q_j^H)(1-q^H)\right)^{H-M}$, which is the probability that these movers are not q^H types with signal b. The product of all these probabilities gives the probability of a cascade occurred for R given by

$$\mathbb{E}(q_j|E,q^L)\left((1-\Pr(q_j^H|E,q_j^M)(1-q^H)-\Pr(q_j^M|E,q_j^M)(1-q^M)\right)^{M-1}\left(1-\Pr(q_j^H|E,q_j^H)(1-q^H)\right)^{H-M}.$$

The probability that a cascade occurred for B is just

$$1-\mathbb{E}(q_j|E,q^L)\left((1-\Pr(q_j^H|E,q_j^M)q^H-\Pr(q_j^M|E,q_j^M)q^M\right)^{M-1}\left(1-\Pr(q_j^H|E,q_j^H)q^H\right)^{H-M}.$$

Taking the ratio of these two is equal to the odds ratio of the probability that the cascade is correct conditional that a cascade occurred. Then, taking the natural log of this ratio gives Eq 3. □

Now, we introduce the definitions of absolute and relative overconfidence for player i. Player i with information $\mathcal{I}_i$ exhibits relative overconfidence if he believes that his accuracy is higher than player j. That is, given her information $\mathcal{I}_i$ we have $\mathbb{E}(q_i|\mathcal{I}_i) > \mathbb{E}(q_j|\mathcal{I}_i)$. Player i exhibits absolute overconfidence if he believes that his accuracy is higher than what it actually is. That is, given her information $\mathcal{I}_i$ we have $\mathbb{E}(q_i|\mathcal{I}_i) > q$. The definitions for relative and absolute underconfidence are the reverse of the inequelities of the overconfidence counterparts.

Proposition 3. *If an arbitrarily small increase in relative confidence by the q^H type leads to a higher H, then the probability of correct informational cascades that starts from the first mover is higher.*

Proof. Fix the beliefs of types q^L and q^M. Suppose that q^H is more overconfident in relative terms, i.e.,

$q^H - \mathbb{E}(q_j|E, q^H)$ increases. This implies a decrease in $\Pr(q_j^H|E, q_j^H)$. Notice that in Condition 2, the first term increases from the decrease in $\Pr(q_j^H|E, q_j^H)$. Hence, this weakly increases H. Since an arbitrarily small change in relative confidence leads to an increase in herd length of $H+1$, it implies that Condition 2 exactly binds. Thus, the change in probability of a correct cascade in log odds is given by

$$(H+1-M)\ln\left(\frac{1-\Pr'(q_j^H|E, q_j^H)(1-q^H)}{1-\Pr'(q_j^H|E, q_j^H)q^H}\right) - (H-M)\ln\left(\frac{1-\Pr(q_j^H|E, q_j^H)(1-q^H)}{1-\Pr(q_j^H|E, q_j^H)q^H}\right), \tag{4}$$

where $\Pr'(\cdot)$ denote the new beliefs due to higher degree of relative confidence. Since any arbitrarily small decrease in $\Pr(q_j^H|E, q_j^H)$ increases H, then it must be the case that Condition 2 exactly binds at $\Pr(q_j^H|E, q_j^H)$ for H. A slight decrease in $\Pr(q_j^H|E, q_j^H)$ leads to a unit jump from H to $H+1$. By continuity, $\ln\left(\frac{1-\Pr'(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr'(q_j^H|E,q_j^H)q^H}\right) \approx \ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)$. Thus Eq 4 is equal to $\ln\left(\frac{1-\Pr(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr(q_j^H|E,q_j^H)q^H}\right)$, which is positive.

□

Proposition 4. *If an arbitrarily small decrease in absolute confidence by the q^M type leads to a higher H but not a change in M, then the probability of correct informational cascades that starts from the first mover is higher.*

Proof. Fix the beliefs of types q^L and q^H. Suppose that q^M is less confident in absolute terms, i.e., $E(q_j|E) - q^M$ decreases. This implies that there is a increase in $\Pr(q_j^L|E, q^M)$. Let the new belief be denoted by $\Pr'(q_j^L|E, q^M)$. Since this an arbitrarily small increase in $\Pr'(q_j^L|E, q^M)$ increases H but not M, then Condition 2 binds while Condition 1 is slack. Then, the change in probability of a correct cascade in log odds is given by

$$\begin{aligned}
&(M-1)\ln\left(\frac{1-\Pr'(q_j^H|E, q_j^M)(1-q^H)-\Pr'(q_j^M|E, q_j^M)(1-q^M)}{1-\Pr'(q_j^H|E, q_j^M)q^H-\Pr'(q_j^M|E, q_j^M)q^M}\right)\\
&-(M-1)\ln\left(\frac{1-\Pr(q_j^H|E, q_j^M)(1-q^H)-\Pr(q_j^M|E, q_j^M)(1-q^M)}{1-\Pr(q_j^H|E, q_j^M)q^H-\Pr(q_j^M|E, q_j^M)q^M}\right)\\
&+\ln\left(\frac{1-\Pr'(q_j^H|E, q_j^H)(1-q^H)}{1-\Pr'(q_j^H|E, q_j^H)q^H}\right).
\end{aligned} \tag{5}$$

Since the change in $\Pr(q_j^L|E, q^M)$ is arbitrarily small, the first two terms are approximately equal by continuity. Hence, the change in probability of a correct cascade is $\ln\left(\frac{1-\Pr'(q_j^H|E,q_j^H)(1-q^H)}{1-\Pr'(q_j^H|E,q_j^H)q^H}\right)$, which positive.

□

Appendix B Belief updating

The previous results show that overconfidence affects subjects' gain or loss in breaking away from a herd even when we control for the degree of difficulty of the quizzes and the revelation of their own score. The following results show the effects of the degree of confidence on their updating behavior moving from their initial belief to the final belief for a given sequence. All analyses on belief updating controls for the initial belief in absolute log odds terms. Table 10 shows that a subject deviates more from her initial belief when she holds a more extreme initial belief.

Figure 11: Quiz Difficulty on Belief Updating

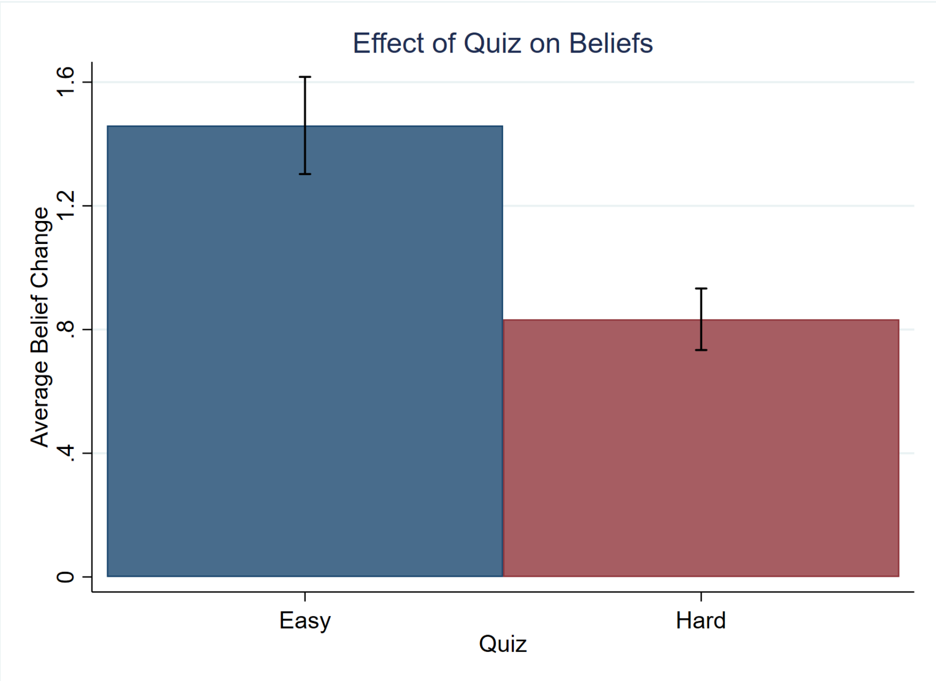


Figure 11 shows the difference of subjects' belief updating between easy and hard quizzes. The average change in beliefs is significantly higher in easy quizzes than hard quizzes with a mean difference of 0.6266 in log odds terms. Column 1 in Table 10 shows a significant difference in updating for the two quiz difficulties controlling for initial belief and score revelation. It further shows that the differential effect of revealing the subject's own score in belief updating. In easy quizzes, revealing the score increases the deviation from initial belief as indicated by the positive coefficient of Score Known variable. The negative coefficient of the interaction of Hard and Score Known variables indicates that the revelation of the score decreases the subject's deviation from her initial belief. The following result summarizes the effects of two treatments on belief updating.

Result 9. *Subject are more modest in updating their beliefs in hard quizzes. Subjects are even more modest in updating their beliefs after realizing their actually quiz performance in hard quizzes. Subject are less modest in updating their beliefs in easy quizzes after observing their quiz performance.*

Table 10: Effects on Belief Updating

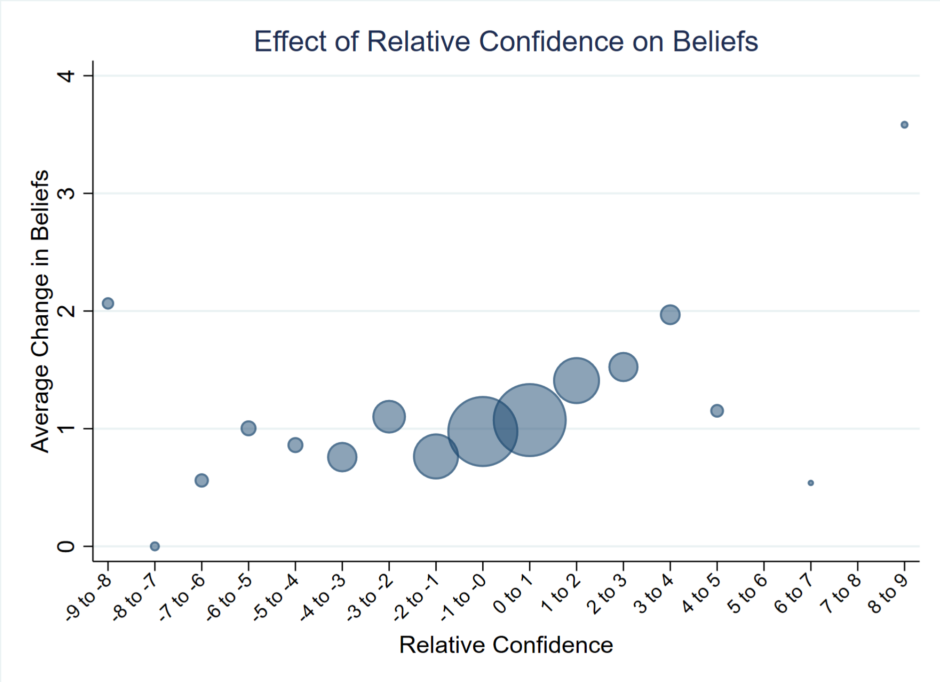
	(1)	(2)	(3)	(4)
Initial Belief	0.515***	0.506***	0.409***	0.409***
	(0.0805)	(0.0802)	(0.0972)	(0.0967)
Hard	-0.239*	0.217	0.262	0.298
	(0.137)	(0.194)	(0.210)	(0.210)
Score Known	0.322**	0.321**	0.367**	0.385**
	(0.149)	(0.144)	(0.158)	(0.157)
Hard x Score Known	-0.411**	-0.358**	-0.442**	-0.468***
	(0.164)	(0.160)	(0.172)	(0.169)
Relative Confidence		0.0689**		-0.0264
		(0.0326)		(0.0233)
E_i(other score)		0.110**		
		(0.0432)		
1(Absolute Confidence$\geq$ 1)			-0.286**	-0.316**
			(0.115)	(0.120)
1(Absolute Confidence$\geq$ 1) x Initial Belief			0.165	0.166
			(0.109)	(0.109)
E_i(own score)			0.105***	0.116***
			(0.0344)	(0.0378)
Constant	1.021***	0.217	0.376	0.315
	(0.187)	(0.341)	(0.319)	(0.324)
Observations	1254	1254	1254	1254

The dependent variable is the absolute difference between log odds ratio of the second reported belief and log odds ratio of first reported belief. The variable Initial Belief is the absolute log odds ratio of the initial belief of the subject. Column 1 presents the total effect of quiz difficulty; column 2 presents the effect of relative confidence; and columns 3 presents the effect of absolute confidence using the dummy variable Overestimate which pools the individuals that exhibit absolute overconfidence.

Standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 12: Relative Confidence on Belief Updating

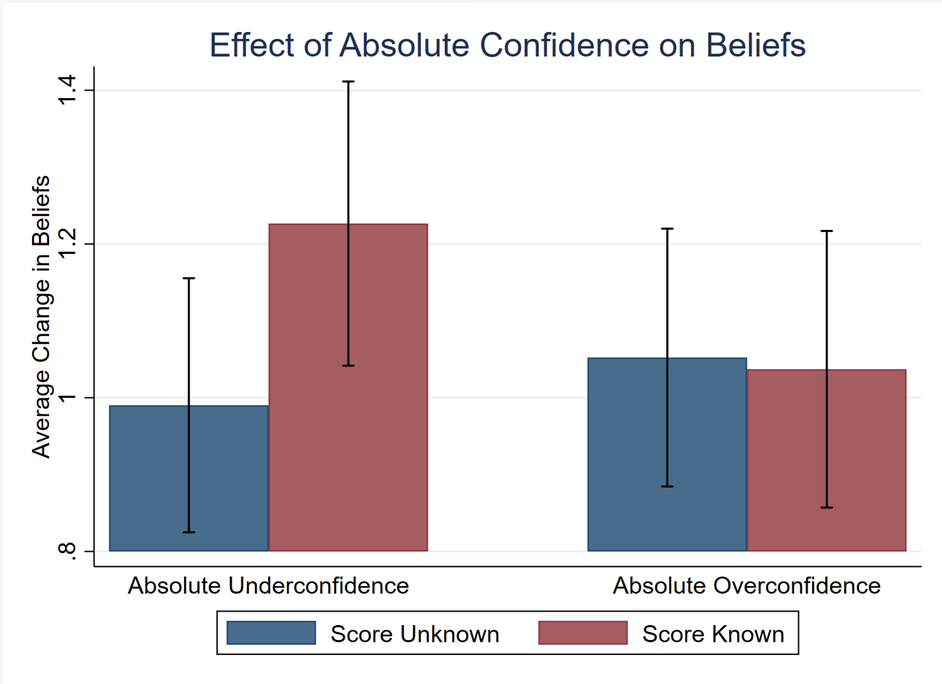


We look at the effects of relative confidence on belief updating. Figure 12 plots the average change in beliefs given the subject's degree of relative confidence. We find a positive linear relationship of relative confidence and belief updating in log odds terms with estimated slope of 0.0837 (s.e. 0.0301). Table 10 shows the positive linear relationship is robust with estimated coefficient 0.0689 (s.e. 0.0326) controlling for initial belief belief, quiz difficulty, and score revelation. We include the variable E_i(other score) to capture the baseline level of belief updating for when a subject does not exhibit relative over- or underconfidence. For such a subject, we find that the baseline behavior belief updating increases with the subject's expectation of a GM's score.

Result 10. *Subjects that exhibit more confidence about their relative performance deviate more from their initial belief.*

Our final result shows the relationship of subjects' confidence regarding their absolute performance on belief updating. Figure 13 shows the different effects of score revelation on belief updating between subjects that exhibit absolute overconfidence and absolute underconfidence. For subjects that exhibit absolute underconfidence, there is a significant increase in belief updating of 0.2363 (s.e. 0.1266) in log odds terms. However, there is no significant difference in updating behavior (mean 0.0152, s.e. 0.1251) for subjects that exhibit absolute overconfidence after revealing their scores. We use the dummy variable Overestimate that is set to 1 when a subject exhibits absolute overconfidence. In Table 10 Column 3, we control for initial belief, quiz difficulty, and score revelation. The estimated coefficient for Overestimate is negative (-0.286) and significant (s.e. 0.115). We find this relationship to be robust even when controlling for relative confidence as presented in Column 4. We use the subject's initial expectation of their score

Figure 13: Absolute Confidence on Belief Updating



to capture the baseline behavior of belief updating when they do not exhibit any degree of absolute confidence. We find that the increase in their expected score increases the deviation from their initial belief. We also interact the effect of initial belief with subjects that exhibit absolute overconfidence. We find a positive relationship given by the estimate coefficient (0.166). We use a joint hypothesis test on Initial Belief and its interaction with Overestimate and find that there is a statistically significant difference ($p < 0.0001$) on the effect of initial belief between absolute over- and underconfident subjects. The following result summarizes relationship of belief updating behavior and absolute confidence.

Result 11. *Subjects exhibiting absolute overconfidence deviate less from their initial belief after observing their quiz performance than those exhibiting absolute underconfidence.*

Appendix C Instructions for the Experiment

The instructions for the experiment are provided below.

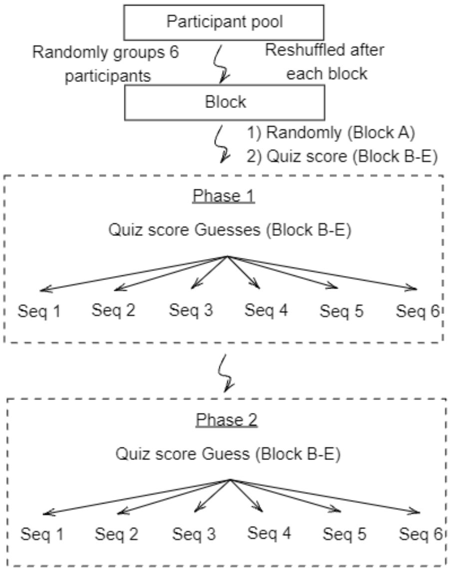
Instructions

Welcome and thank you for taking part in this Economics Experiment. This experiment will last for around 2-2.5 hours. If you read the instructions carefully, you can earn a considerable amount of money depending on your decisions and chance. Your earnings will be paid out to you via Cash at the end of the experiment.

Before we begin, we ask that you turn off your cell phones for the duration of this experiment. We also ask that you do not communicate with anyone for the duration of this experiment and only use the software provided to you on your devices. Failure to comply with these rules may result in dismissal from this experiment and you will not be paid any earnings you may have otherwise received.

Overview

1. This experiment has 5 blocks. Each block has 2 phases, and each phase has 6 sequences. Your main task is to make guesses about the chance that the color assigned to your group is Red in each sequence.
2. Before each block, you will be randomly assigned into groups of 6.
3. You will be assigned a score for a block. In Block A, the score will be randomly assigned to you. In Blocks B-E, this will be based on your score on the quiz you will take at that beginning of the block.
4. You will only take a quiz at the start of Blocks B, C, D, and E.
5. There are two phases for each block. In Phase 1, we will not reveal your assigned score. In Phase 2, we will reveal your assigned score.
6. In Phase 1, we will ask about your beliefs on your own and a randomly selected group member's quiz performance.
7. In Phase 2, we will show you your score and ask you about your belief on a randomly selected group member's quiz performance.
8. A higher score means your private signal is more likely to match the color assigned to your group for the sequence. A lower score means it's less likely to match the assigned color.
9. You will proceed with your main task of making guesses about the assigned color of the sequence, 2 guesses per sequence, which means 12 guesses per phase with a total of 24 guesses per block
10. A color, Red or Blue, will be randomly assigned to your group with equal chance for each sequence.
11. You will make two guesses about the probability that the assigned color is Red for the sequence.
12. Before making your 1st guess, we will show you an indication of other's guesses for the sequence. We will show you nothing if you are the first mover in the sequence, but you will still know that Red is assigned with 50% chance.
13. Before making your 2nd guess, we will show you your private signal drawn from your jar. You will still see the same indications from guess 1.
14. You will start as the 1st mover in the first sequence you face, then as 2nd mover in the second sequence you face, ..., and lastly as the 6th mover in the sixth sequence you face.
15. You will proceed to the next block until you finish all 5 blocks.



Next

Instructions

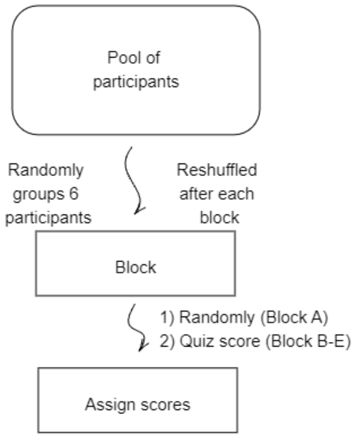
Blocks

This experiment consists of a series of decision rounds spread out over **5 blocks**. At the beginning of each block, you will be **randomly** assigned to a group with 6 participants (including you). Neither you, nor your group members will know each other's identities. Your main task is to make guesses about the chance that the color assigned to your group is Red.

You will be given a signal to help you make your guesses. At the beginning of each block you will be assigned a score from 0 to 10. The accuracy of the signal will depend on the score. The purpose of a higher score is to have a more accurate signal to help you make your guess. How the scores are assigned to each one in your group depends on the block.

In Block A - You will be randomly assigned a score from 0 to 10, each equally likely.

In Blocks B-E - Your group will first take part in a general knowledge trivia quiz. Each quiz contains 10 questions. Your score in the quiz determines your assigned score for the block. **Everyone in your group will take the same quiz per block.**



Back

Next

Instructions

Phases

Each block will have 2 phases.. In Phase 1, you will not know the score assigned to you. In Phase 2, you will be notified about your own score. **However, you will never know the score of any of your group members.**

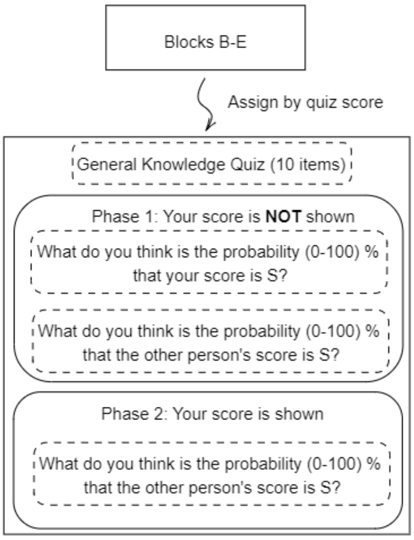
For Blocks B to E, you will make some additional guesses at the start of each phase about your and a randomly selected group member's quiz performance. In Phase 1, after taking the quiz, for each score S (from 0 to 10), you will guess the probability that the score is S by answering the following question:

What do you think is the probability (0-100)% that the score is S?

You will make 2 sets of these 11 guesses. One set is for your own score on the quiz. The other set is for the score of one randomly selected member of your group who has taken the same quiz.

In Phase 2, you will be shown your own score and will once again make the same guesses for the randomly selected member.

Note that for each set, all probabilities must sum up to 100%!



Back

Next

Instructions

Sequences:

There are **6 sequences in a given phase** with a total of 12 sequence in a block. In each of these sequences, your group will be assigned a color, Red or Blue, both equally likely. The color assigned to your group in each sequence is chosen **independently** of the colors in other sequences. Neither you nor your group members will know the actual color assigned to your group

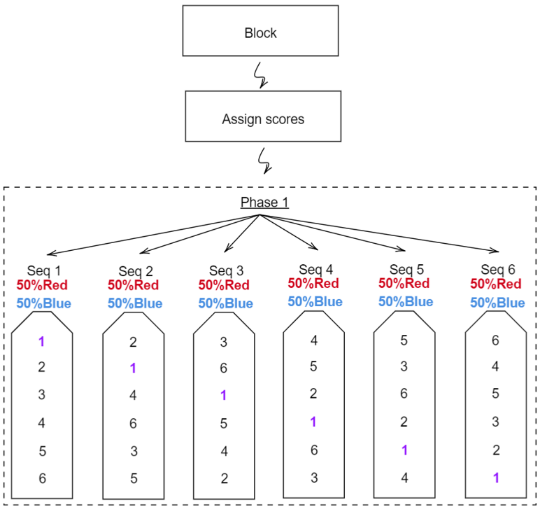
Your main task is to guess the probability that the color assigned to your group is Red for each sequence. You will make 2 guesses per sequence by answering:

What do you think is the probability (0-100)% that the color is Red for this sequence?

For each sequence, every group member will make their guesses in a randomly determined order. The timeline for each set of guesses is as follows:

1. When it is your turn, you will first see an indication of what the previous movers in your group have done. If you are the first mover, you will see no indication.
2. You will then make your first guess (Guess 1) about the probability that the color assigned to the group is Red based on the indication of the previous movers. If you are the first mover in the sequence, then you make your guess based only on the fact that the color assigned is either Red with 50% chance or Blue with 50% chance.
3. You will then see a private signal that provides some hint as to your group's assigned color. You will then make your second guess (Guess 2) about the probability that the color assigned to the group is Red based on the past indications and your signal.

The following is a sample illustration of six sequence of Phase 1 in a given block.



Back

Next

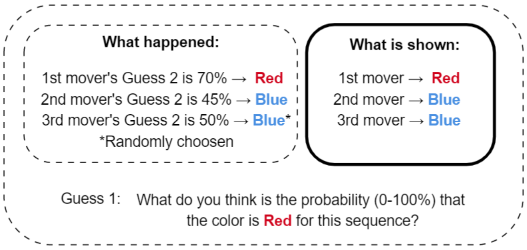
Instructions

Guesses

Indications

Before making the first guess (Guess 1), you will be able to see an indication of what other people in your group have guessed for this sequence. This indication can be either Red or Blue and will depend on their second guess (Guess 2). You won't be shown their exact percentage, you will only know if their guesses are above or below 50%. If someone before you guessed greater than 50%, the indication will be Red. If they guessed less than 50%, the indication will be Blue. If they guessed 50%, then the computer will randomly choose an indication of either Red or Blue, each equally likely. If you are the first mover, there are no previous guesses and hence no indications. Again, you only know the fact that the color assigned is either Red with 50% chance or Blue with 50% chance.

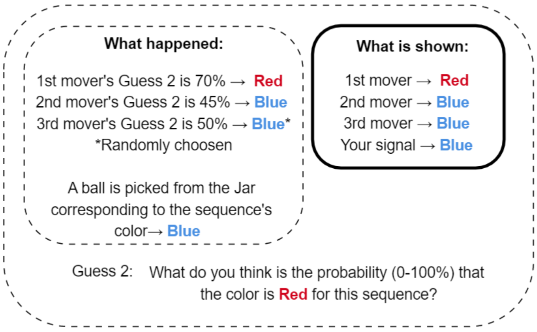
The following is a sample illustration of a history of indications for Guess 1.



Private Signal

Before you make the second guess, you will also receive a private signal. This signal will take the form of a ball drawn from a jar that is assigned to you based on your group's color. If the group's color is Red, your ball will be drawn from a "Red Jar" and if the group's color is Blue, the ball will be drawn from a "Blue Jar." Each jar will be composed of 100 balls. The Red jar contains more red balls than blue balls. The Blue jar contains more blue balls than red balls. The composition of the jars depends on your assigned score for the block. **Note that your private signals are drawn independently of the signals of other participants**.

The following is a sample illustration of a history of indications for Guess 2.



Back

Next

Instructions

Jars

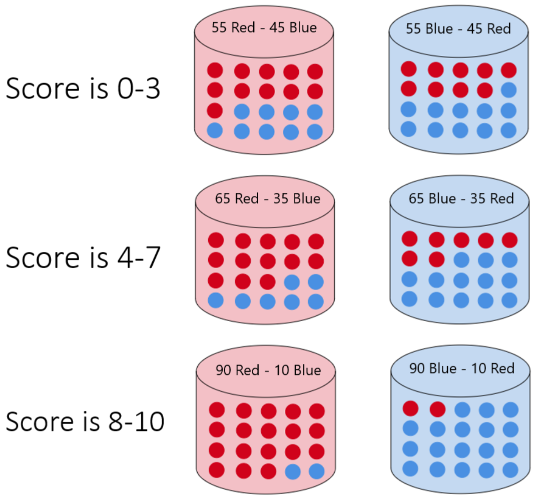
Each of you gets their private signal drawn from their own assigned jar. The composition of your assigned jar depends on your score (from 0 to 10) for the block.

A higher score will lead to a higher number of Red balls and a lower number of Blue balls in the Red jar. Likewise, higher scores will lead to a higher number of Blue balls and a lower number of Red balls in the Blue jar.

A score of 0-3 gives a composition of 55 red and 45 blue balls in the Red jar and 45 red and 55 blue balls in the Blue Jar. A score of 4-7 gives a composition of 65 red and 35 blue balls in the Red jar and 65 blue and 35 red balls in the Blue jar. A score of 8-10 results in 90 red and 10 blue balls in the Red jar and 90 blue and 10 red balls in the Blue jar.

Note - The color of everyone's jar is the same as the color of the group. However, the composition of everyone's jar is dependent on their own score. You will not know the color of the jar the ball is drawn from, neither will you know the color or composition of the jars of other people in your group. Your score has no influence on the chance of the color being assigned to your group, rather a higher score makes it more likely for your ball to be the same color as that of the group.

The following is an illustration of the Jar composition given a score.



Back

Next

Instructions

Order of Sequences

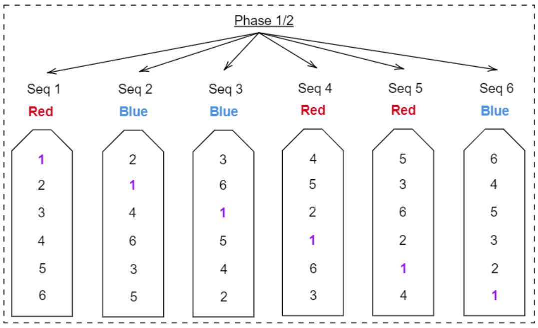
In the beginning of a phase, each of you in your group will simultaneously act as the first mover and make guesses in different sequences. After everyone makes their guesses as the first mover, you will then randomly move to another sequence following the first mover of that sequence in making your guesses. When everyone in your group finishes making their second set of guesses, you move on to another randomly chosen sequence and make your two guesses following the first and second movers of that sequence. This process repeats until all of you in the group make your two guesses as the 6th mover.

We repeat this structure twice per block, once for each phase.

A few points to note:

1. The color assigned to the group remains the same for each mover in a given sequence. However, the color assigned in one sequence has no impact on the color assigned in another sequence. Knowing the color in one sequence does not help you in making your guesses in another sequence
2. You will never face the same sequence twice i.e. you will make only one set of guesses in a given sequence
3. You will never follow the same person twice meaning you will follow a given person at most once in any given sequence. Also, you will never follow yourself.

The following is a sample illustration of the six sequences in a given phase.



Back

Next

Instructions

Payment

Your payment will depend on the following:

1. Your performance on one randomly chosen quiz (from Blocks B, C, D and E) (up to $10)
2. Your guess on one of your scores for the selected quiz in 1) (up to $2))
3. Your guess on one of the scores of your group member's performance for the selected quiz in 1) (up to $2)
4. Your two guesses in one randomly chosen sequence in Blocks A to E. (up to $24)

After each block, you will be shown the *possible* payment you can receive provided the block or quiz is chosen for payment.

For 1) we will pool your quiz score with the scores of 44 people who have previously taken the same quiz. Your earnings for your performance on the quiz will be $10 * r where r is your percentile on the quiz.

For 4), the computer will first randomly choose a sequence for the block. Each of the 12 sequences is equally likely to be chosen. The computer will then draw a random number Y1 from 0 to 100. **If Y1 is less than or equal to your reported Guess 1**, then you will be paid **$12 if the assigned color is Red. If Y1 is greater than your guess,** then you will be paid based on a lottery in which **you can earn $12 with Y1% chance.**

The computer will draw another random number Y2 from 0 to 100 for Guess 2. Then, the same payment scheme is implemented for your reported Guess 2 in the chosen sequence. This can earn $12.

For your guesses about your own and random member's quiz performance (parts (2) and (3)), the computer will randomly choose a score S from the selected quiz in part (1). The computer will then generate a number Y from 0 to 100. If Y is less than your guess, then you will be paid $2 if the score is S. If it less than Y, then you will be paid based on a lottery than can earn $2 with Y% chance. Each guess on the two selected scores can earn $2.

At the end of the experiment, the computer will randomly draw a block from A to E and a quiz from blocks B to E. You will be paid for the chosen block and quiz as stated above. In addition to this, you will also be paid a $5 show-up fee.

Note: The payment scheme is designed in a way to ensure that the best way to maximize your earnings is to report your true guess for the probability. For example, for guesses regarding the color Red, you maximize your earnings by reporting your true beliefs regarding the probability of the color being Red.

Back

Next

Appendix D Quiz Questions

Easy 1		
No.	Question	Answer
1	What continent lies directly south of Europe?	*Africa*
2	In what North American country is the city of Toronto located?	*Canada*
3	In what western U.S. state is the Silicon Valley?	*California*
4	What is the smallest prime number greater than 3?	*5*
5	What is the chemical symbol for water?	*H2O*
6	What African predator is the fastest land animal?	*Cheetah*
7	On what country did the United States drop atomic bombs on during World War 2?	*Japan*
8	The Red Scare was a fear of what political system?	*Communism*
9	Cinderella, The Little Mermaid, Aladdin, and The Lion King are all films produced by what famous entertainment company?	*Disney*
10	John Lennon, Paul McCartney, George Harrison and Ringo Starr were the four members of what famous classic rock band?	*The Beatles*

Hard 1		
No.	Question	Answer
1	The most famous 'tea party' of the American Revolution took place in what city?	*Boston*
2	What is the only country whose capital begins with 'Z'?	*Croatia*
3	The psychoactive ingredient in marijuana is THC. What does THC stand for?	*Tetrahydrocannabinol*
4	The study of the structural and functional changes in cells, tissues and organs that underlie disease is called what?	*Pathology*
5	Who was the leader of Germany's Nazi party during World War 2?	*Hitler*
6	The Korean War was fought in which decade?	*1950-1960*
7	Who wrote and directed Kill Bill Volumes 1 and 2?	*Quentin Tarantino*
8	Who is the only actress to have been nominated for the "Best Actress" Academy Award 17 times?	*Meryl Streep*
9	What actress performed the voice of Bo Peep in the film Toy Story?	*Annie Potts*
10	Who is credited with inventing the wristwatch in 1904?	*Louis Cartier*

Easy 2		
No.	Question	Answer
1	What is the capital of and largest city in Japan?	*Tokyo*
2	An octopus has how many arms?	*8*
3	What is the major pumping organ of the human circulatory system?	*Heart*
4	On what date in 2001 were airplanes crashed into New York's World Trade Centers, destroying them? (Month then date)	*September 11*
5	At the beginning of what year, known as 'Y2K' were computers expected to crash due to the 'Millennium bug'?	*2000*
6	How many World Wars have there been?"	*2*
7	The Italian village of Pompeii was destroyed in 79 AD by what type of natural disaster?	*Volcano*
8	What was the title of George Lucas's original 1977 science fiction film about Luke Skywalker and Darth Vader, set 'Long ago, in a galaxy far, far away'?	*Star Wars*
9	Singer Celine Dion sang the hit song My heart will go on for the soundtrack of what 1997 film about the sinking of a famous ship?	*Titanic*
10	What Chicago Bulls guard wore number 23 and led his team to 6 championships in the 1990's?	*Michael Jordan*

Hard 2		
No.	Question	Answer
1	Baghdad is the capital of what middle-eastern country?	*Iraq*
2	On what continent is France located?	*Europe*
3	What country is comprised of over 17,000 known islands?	*Indonesia*
4	What is the capital of Turkey?	*Ankara*
5	The _____ is the star at the center of our solar system.	*Sun*
6	Which bone in the human body is not attached to any other bone?	*Hyoid*
7	What leader is associated with the death of more than 1.5 million Cambodians?	*Pol Pot*
8	Who was the first unseeded man to win the Wimbledon singles title?	*Boris Becker*
9	What was the final battle that Napoleon fought in?	*Waterloo*
10	What was the name of the person who shot and killed Marvin Gaye one day before his 45th birthday?	*Marvin Gaye, Senior*

Appendix E Robustness Checks

E.1 Checking for Hedging

Under our initial payment scheme, subjects were paid for both of their reported beliefs in the same randomly chosen sequence. As mentioned above, this introduces a hedging opportunity where subjects could potentially report beliefs that contradicted each other to gain a higher payoff. For example, a subject could report her first belief as above 50% and report her second belief to be below her first belief. While this behavior makes sense if the subject receives a signal contradicting her initial belief, it does not if she receive a signal that confirms her belief. We conducted some sessions under a revised payment scheme in which subjects were paid for one randomly chosen belief. This design should, in theory, mitigate any hedging opportunities that might arise.

To see if our data provided evidence for hedging, we conducted the following analysis. We first documented any observations that could potentially be counted as hedging. We classified an observation to be a "potential hedge" if the subject reported a first belief above (below) 50%, received a Red (Blue) signal and reported a second belief *below (above)* her first belief. Table 11 below shows the frequencies of observations classified as potential hedging under both payment schemes. We find that at most 5% of the observations can be classified as potential hedging (out of all observations). Furthermore, there is no significant difference in the distributions across payment schemes (χ^2 test $p = 0.595$). The frequency of potential hedging increases to 11-12% when restricting the analysis to observations in which subjects receive signals confirming their initial belief. Again, we find no significant differences in the distributions across both payment schemes (χ^2 test $p = 0.636$). We thus conclude that there is no significant evidence of hedging in our data.

Table 11: Checking for Hedging

All Observations

	Both Beliefs	One Belief
Not Hedging	94.9%	95.2%
Potential Hedging	5.1%	4.8%

Confirming Signals

	Both Beliefs	One Belief
Not Hedging	88.3%	88.9%
Potential Hedging	11.7%	11.1%

E.2 Multinomial Logit for Welfare Change and Belief Updating

Table 12: Welfare

	(1)		(2)		(3)		(4)	
	-1	1	-1	1	-1	1	-1	1
Hard	-0.425	-0.423*	3.828**	-2.914*	2.863*	-0.727	2.812*	-0.741
	(0.289)	(0.237)	(1.409)	(1.230)	(1.553)	(1.363)	(1.548)	(1.335)
Score Known	0.323	-0.0969	0.277	-0.0656	0.282	-0.265	0.267	-0.280
	(0.321)	(0.268)	(0.320)	(0.269)	(0.322)	(0.293)	(0.321)	(0.301)
Hard x Score Known	-0.449	-0.102	-0.413	-0.131	-0.460	0.166	-0.376	0.250
	(0.420)	(0.358)	(0.420)	(0.359)	(0.429)	(0.384)	(0.433)	(0.389)
Ave Others' Score			0.731**	-0.435*	0.663**	-0.250	0.654**	-0.265
			(0.232)	(0.209)	(0.229)	(0.210)	(0.231)	(0.209)
Relative Confidence					-0.0893	0.341**		
					(0.103)	(0.100)		
E_i(other score)					-0.140	0.211*		
					(0.106)	(0.0988)		
Absolute Confidence							-0.000248	-0.441**
							(0.144)	(0.113)
E_i(own score)							-0.131	0.250**
							(0.0958)	(0.0878)
Constant	-0.355	0.256	-6.536**	3.854*	-5.001*	0.604	-4.978*	0.513
	(0.458)	(0.394)	(2.062)	(1.785)	(2.315)	(2.051)	(2.280)	(1.997)
Observations	916		916		916		916	

The table looks at the effects of several variables of interest on welfare change using a multinomial logit with base group of no switch in beliefs. Models 1 and 2 present the total effect of quiz difficulty; model 3 presents the effect of relative confidence; and model 4 present the effect of absolute confidence.

Standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$