Organizational Behavior and Human Decision Processes
Vol. 86, No. 2, November, pp. 168–196, 2001
doi:10.1006/obhd.2000.2950, available online at http://www.idealibrary.com on IDEAL®

The Effect of Information Strength and Weight on Behavior in Financial Markets

Mark W. Nelson, Robert Bloomfield, Jeffrey W. Hales, and Robert Libby

Johnson Graduate School of Management, Cornell University

Griffin and Tversky (1992) explain evidence of individual over- and underconfidence as resulting from attending too much to the strength (i.e., extremity) of information and not enough to the weight (i.e., statistical reliability) of information. We report two experiments that demonstrate how information strength and weight affect confidence, trading, prices, and wealth in laboratory markets. Our results indicate that information strength and weight affect individual over- and underconfidence and that market participants lack sufficient self-insight to avoid trading when they are biased. As a consequence, market prices are biased, and market participants with high-strength, low-weight information systematically transfer wealth to participants with low-strength, high-weight information.

Key Words: strength; weight; confidence; overconfidence; financial markets; calibration; aggressiveness; investor welfare.

INTRODUCTION

People appear underconfident in many circumstances, systematically underestimating the probability that their judgments are correct (see, e.g., Edwards, 1968; Keren, 1987; Lichtenstein, Fischhoff, & Phillips, 1982). However, people appear overconfident in other circumstances, systematically overestimating the probability that their judgments are correct (e.g., Dunning, Griffin, Milojkovic, & Ross, 1990; Koehler, Brenner, Liberman, & Tversky, 1996; Lichtenstein & Fischhoff, 1977; Tversky & Kahneman, 1971). These errors in confidence are often referred to as *miscalibration*.

We thank Cornell's Johnson Graduate School of Management for financial support and Susan Krische and Steve Smith for comments.

Address correspondence and reprint requests to Mark W. Nelson, Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853-6201. E-mail: mwn2@cornell.edu; fax: 607-254-4590.

0749-5978/01 $35.00

Griffin and Tversky (1992) argue that miscalibration arises because people attend too much to information strength (i.e., salience) and attend too little to information weight (i.e., statistical reliability). For example, Griffin and Tversky (1992) have subjects estimate the probability that a coin drawn from a sample will be heads-biased (as opposed to tails-biased), given a 50% prior probability that the coin will be heads-biased and the outcome of a sample of coin flips. Griffin and Tversky predict that subjects will attend too much to the proportion of heads in the sample, as that aspect of the sample will be very salient, and will attend too little to the size of the sample and thus to the reliability of an inference that can be drawn from the sample. Results support their predictions. Probability estimates are insufficiently extreme when large samples have proportions that are not extreme, but are too extreme when small samples have extreme proportions. Griffin and Tversky note that an estimate of the probability that a coin is heads-biased is similar to an estimate of the probability that a judgment is correct in other studies of confidence and suggest that strength and weight can predict miscalibration in a variety of contexts. Results from other experiments in their study support this generalization.

We examine how information strength and weight affect the behavior of participants (hereafter, “investors”) in laboratory financial markets. In our experiments, investors trade artificial securities whose values are determined by the correct posterior probabilities in coin-flipping problems adapted from Griffin and Tversky (1992). With this approach, we can assess effects of strength and weight on (1) individual miscalibration, by examining the extremity of estimates of security value; (2) trading behavior, by examining the number of shares that investors trade; (3) aggregate market-level miscalibration, by examining the extremity of the market prices that result from investors’ trading decisions; and (4) investors’ wealth, by examining trading profits. Each of these dependent variables (individual value estimate, number of shares traded, market price, and trading profit) allows us to make a different contribution.

Examining investors’ individual value estimates allows us to assess whether the systematic effects of strength and weight on miscalibration that Griffin and Tversky (1992) observed persist in the presence of market incentives. Prior studies indicate limited effectiveness of incentives in debiasing judgment in general (for reviews, see Bonner, Hastie, Sprinkle, & Young, 2000; Camerer & Hogarth, 1999; Jenkins, Mitra, Gupta, & Shaw, 1998; and Smith & Walker, 1993) and under- and overconfidence in particular (see, e.g., Fischer 1982). Therefore, we anticipated that we would still find under- and overconfidence in our setting.

Examining the number of shares investors choose to trade, given their estimates of share value, allows us to measure the aggressiveness with which investors are willing to bet on their beliefs being correct (and therefore on their trades being profitable). Because an estimate of share value in the coin-flipping paradigm constitutes a confidence judgment, aggressiveness reflects investors’ confidence in their confidence judgments. Prior research indicates that people generally dislike ambiguity in the uncertainties they face. For example, people generally prefer uncertain alternatives for which probabilities are less ambiguous (e.g., Einhorn & Hogarth, 1985; Ellsberg, 1961) and for which people believe they have relatively more knowledge or competence (e.g., Curley, Yates, & Abrams, 1986; Heath & Tversky, 1991; Keynes, 1943).[1] Thus, all else equal, as confidence in judgment increases, investors will risk trading more shares of a security. If strength and weight affect aggressiveness in the same way they affect confidence, adverse wealth effects could be compounded (e.g., overconfident investors could overvalue securities and buy many shares). On the other hand, if investors have self-insight that they are vulnerable to miscalibration, they may compensate by being less aggressive and suffer less of a wealth effect than they would otherwise.

Examining market prices allows us to determine the extent to which the effects of strength and weight on individual miscalibration are diminished via market aggregation of individual beliefs to form a security price. Accurate market prices result when individual investors' judgments are accurately weighted through the process of prices moving in response to supply and demand (Camerer, 1992). To the extent that all individual estimates of security value exhibit under- or overconfidence, a market will likewise exhibit under- or overconfident market prices. Consistent with this, prior research has found evidence of overconfidence in market prices. For example, market participants have been shown to make trading decisions that indicate they overweight relatively unreliable information such as rumors (DiFonzo & Bordia, 1997) and media reports that lack predictive ability but enable the construction of a causal story (Andreassen, 1987, 1990). However, the factors that affect extent of miscalibration of market prices are largely unexplored. Therefore, we test whether information strength and weight predict under- and overconfidence in market prices.

Examining trading profits allows us to determine the extent to which strength and weight actually affect investor wealth. In general, market participants' trading profits depend on the accuracy of their valuation judgments and trading decisions relative to other market participants. If all investors possess the same information, all may be affected to the same extent, and wealth transfers among investors should be relatively low. However, if investors differ in their information, and therefore in the under- or overconfidence that strength and weight induce, some investors may systematically transfer wealth to others. The market context allows us to examine a practical consequence of any effect of strength and weight on security valuations, trading decisions, and market prices.

The remainder of this paper proceeds as follows: the Background section

[1] Ambiguity avoidance has a long history of support. Keynes (1943) uses the now classic example of a known versus unknown proportion of white balls in an urn to illustrate that probability estimates can be the same while the "weight of evidence" differs. He asks, "If two probabilities are equal in degree, ought we, in choosing our course of action, to prefer that one which is based on a greater body of knowledge?" In answer, he states the "degree of completeness of the information upon which a probability is based does seem to be relevant, as well as the actual magnitude of the probability, in making practical decisions" (p. 313).

discusses research relevant to information strength, weight, and market behavior. The following two sections report Experiment 1 (parts of which are also reported in Bloomfield, Libby, & Nelson, 2000) and Experiment 2. The Discussion section summarizes our findings and describes possible implications.

BACKGROUND

Griffin and Tversky (1992) distinguish information according to two characteristics: strength and weight. In their terminology, the strength of evidence is the degree to which it is favorable or unfavorable. The weight of evidence is its statistical reliability. Griffin and Tversky (1992) provide evidence that people tend to pay too much attention to strength and not enough attention to weight. As a result, overconfidence is observed when evidence is of high strength but low weight, and underconfidence is observed when evidence is of low strength but high weight. More generally, overconfidence decreases (or underconfidence increases) as information strength decreases and/or weight increases.

Example

To better see how information strength and weight interact to affect calibration, consider the following coin-flipping exercise adapted from Griffin and Tversky (1992) and used in Bloomfield *et al.* (2000).

> Assume there are two types of coins: heads-biased and tails-biased. When flipped, heads-biased coins have a 60% chance of coming up heads and a 40% chance of coming up tails; tails-biased coins have a 40% chance of coming up heads and a 60% chance of coming up tails.
>
> Assume we draw coins from a pile that has an equal number of heads-biased and tails-biased coins. We flip each coin some number of times. We toss each coin into a bucket depending on how many heads and tails came up.
>
> For a particular bucket (e.g., 3 flips, 2 heads), what is the proportion of those coins (in percent) that are heads-biased?

In the context of this example, the strength of evidence is the sample proportion, while the weight of evidence is the sample size (i.e., the number of flips).

According to Bayes' rule, the factor that actually determines the proportion of coins that are heads-biased is the difference between the number of heads and the number of tails. For example, when a sample of flips results in three heads and no tails, Heads − Tails = 3, and Bayes' theorem indicates a normative posterior probability of 77. The sample proportion (3/3 = 100%) appears to provide very strong evidence in favor of the coin being heads-biased, but the sample size is relatively small. Thus, the sample outcome should be given relatively little weight. Similarly, when a sample of flips results in ten heads and seven tails, Heads − Tails = 3, and the normative posterior probability is again 77.[2] While this less extreme sample proportion (10/17 = 58.8%) does not appear to provide as strong evidence of a heads-biased coin as the previous sample, it does have a sample size of 17. Thus, the sample outcome is more reliable, higher-weight evidence and should be given relatively more weight. Griffin and Tversky's (1992) evidence thus suggests that decision makers attend too much to strength and not enough to weight, such that the higher-weight 17-flip sample encourages underconfidence and the lower-weight 3-flip sample encourages overconfidence.[3]

Strength, Weight, and Financial Markets

Do biases persist in financial markets? In general, a substantial literature examines whether individual biases are also observed in markets. Some of the biases that have been studied include: base rate neglect and representativeness (Anderson & Sunder, 1995; Camerer, 1987; Ganguly, Kagel, & Moser, 1994), speculative bubbles and information mirages (Camerer & Weigelt, 1991; Smith, Suchanek, & Williams, 1988), the curse of knowledge (Camerer, Loewenstein, & Weber, 1989), preference reversals (Cox & Grether, 1996), framing effects (Weber, Keppe, & Meyer-Delius, 2000), the winner's curse (Kagel & Levin, 1986), the endowment effect (Kahneman, Knetsch, & Thaler, 1990), the ambiguity effect (Sarin & Weber, 1993), the disposition effect (Weber & Camerer, 1998), and order effects (Tuttle, Coller, & Burton, 1997).

Factors determining whether or not an individual bias is observed in a market setting are still not well understood. Sometimes individual biases have been shown to be significantly reduced in markets because learning takes place at the individual level as feedback is received (Kagel & Levin, 1986; Lind & Plott, 1991; Waller, Shapiro, & Sevcik, 1999; Williams, 1987). Other studies have found that markets can also exacerbate biases (Camerer, 1987; Ganguly *et al.*, 1994) and may introduce biases that did not exist at the individual level (e.g., by providing an incentive to speculate, Smith *et al.*, 1988). Finally, some studies show little or no difference in market-level biases (Kahneman *et al.*, 1990; Sarin & Weber, 1993; Weber & Camerer, 1998). In general, even if individual

[2] To see intuitively how it is only the difference between the number of heads and tails which determines security value, regardless of sample size, decompose the 17 flips into the following: (1) a group of three flips with three heads and no tails and (2) a group of 14 flips with seven heads and seven tails. The information conveyed by the first group is identical to the information conveyed by all three flips in the low-weight, high-strength setting. The information conveyed by the second group is completely uninformative because each head outcome is completely offset by a tail outcome. As a result, the posterior is the same in both cases (77). More formally, let h and t be the number of heads and tails, where $h - t = d > 0$, and let H and T denote heads-biased and tails-biased coins. Bayes' rule dictates that $\Pr[H|h, t] = \Pr[h,t|H]/(\Pr[h,t|H] + \Pr[h,t|L]) = 0.6^{t+d}0.4^{t}/(0.6^{t+d}0.4^{t} + 0.4^{t+d}0.6^{t}) = 0.6^{d}/(0.6^{d} + 0.4^{d})$. Thus, d is sufficient to calculate the posterior.

[3] Bloomfield *et al.* (2000) generalize this effect by modeling individuals as Bayesians who have uncertain information about the reliability of their information. This uncertainty causes individuals to be overconfident on average when they have data of low reliability and to be underconfident on average when they have data of high reliability (a phenomenon the authors refer to as "moderated confidence").

biases are reduced in markets, they still tend to persist at statistically significant levels.

The focus of this paper is on miscalibration. Studies have investigated miscalibration and market behavior in different ways.

Tracking studies. Under one approach, participants determine the trading decisions they would make in response to a sequence of price movements adapted from real-world financial markets (e.g., Andreassen, 1987, 1990; DiFonzo & Bordia, 1997). Assuming that market prices are accurate, participants' wealth is enhanced by "tracking" price movements (i.e., buying when prices are low and selling when prices are high). Since participants possess less information than does the rest of the market, they are at an informational disadvantage compared to the rest of the market, so departures from tracking indicate overconfidence.[4] It is important to note that, in these studies, market participants' judgments do not affect market prices because they are trading in response to a historical price sequence. Hence, these studies do not reveal the effect of miscalibration on market price, but do provide evidence that participants' responses to rumors or news stories cause departures from tracking and reduce their wealth.

Of particular interest in DiFonzo and Bordia (1997) is that knowledge of the reliability of information has little effect on reducing this behavior. Even though participants rate unpublished rumors as less credible than published rumors, and rate both types of rumor as less credible than news, information from each of these sources leads to similar departures from tracking. This result is consistent with information weight (credibility) having less of an impact than does information strength (e.g., a lurid story contained in a rumor), leading to overconfidence and lower participant welfare than would have resulted if better tracking had occurred.

Laboratory market studies. Another approach to investigating the effects of miscalibration on market behavior allows participants to trade securities with each other in simulated financial markets. Many of these studies provide indirect evidence on miscalibration. For example, Bloomfield, Libby, and Nelson (1996) show that traders who have more accurate answers to almanac questions trade more aggressively, and thereby influence prices more, than do traders who have less accurate answers. Papers examining whether biases persist in financial markets (as cited above) indicate that traders who are less susceptible to biases do not necessarily trade any more actively and thus have limited ability to debias prices.

Another study by Bloomfield, Libby, and Nelson (1999) provides more direct evidence on the role of miscalibration in markets. In that study, some more-informed participants have access to all available information, and trade with

[4] This interpretation assumes that price changes will tend to reverse themselves in the near future, making it profitable on average to buy after price decreases and sell after price increases. If markets are weak-form efficient (Fama, 1970), price reversals are unpredictable and deviations from a tracking strategy need not indicate overconfidence.

other less-informed participants who have only a strict subset of the information held by more-informed participants. Because less-informed investors only have information that is also held by more-informed investors, less-informed investors should not trade actively and should have no effect on prices. However, market prices are strongly biased in the direction of less-informed participants' information. As a consequence, less-informed participants buy high and sell low, systematically transferring wealth to more informed participants.

Relation to current experiments. The results of the tracking and laboratory-market studies described above are consistent with the idea that over-attending to information strength and under-attending to information weight could bias prices and lead to wealth transfers between market participants. The results of Bloomfield *et al.* (1999) also suggest that the strength and weight of evidence may affect not only the extremity of participants' value estimates, but also the aggressiveness with which they communicate their estimates to others (through trading). However, neither approach is designed to examine the effects of both strength and weight. The tracking studies always use information of zero weight, since the information available to participants is unrelated to prices. The Bloomfield *et al.* (1999) study operationalizes weight as R^2 and manipulates the amount of information such that the information held by less-informed participants yields a substantially lower R^2 than does that of the more-informed participants. Thus, weight is manipulated, strength is not, and weight is confounded with the amount of value-relevant information.

In this paper, we use a methodology that varies both information strength and weight simultaneously, which allows us to examine directly the effects of these variables on individual and market judgments. In Experiment 1, all investors possess either high-strength, low-weight information or low-strength, high-weight information. In Experiment 2, we vary the strength and weight of information between investors in the same markets, which allows us to investigate whether relatively overconfident investors (i.e., those that have information that is of high strength but low weight) trade too aggressively and suffer adverse consequences as a result of their overconfidence. Thus, in both experiments we examine how strength and weight affect individual miscalibration, aggressiveness of trading decisions, and miscalibration in market prices, and in Experiment 2 we also examine how strength and weight affect transfers of wealth between market participants.

EXPERIMENT 1

In Experiment 1 (parts of which are also reported in Bloomfield *et al.*, 2000), investors trade artificial securities in laboratory financial markets. Security values are determined by the correct posterior likelihood for coin-flipping problems adapted from Griffin and Tversky (1992). All of the investors in a given market are identically informed with value-relevant information. In some markets, all investors are given a high-strength, low-weight 3-flip signal; in others, all investors are given a low-strength, high-weight 17-flip signal. Between the

3-flip and the 17-flip settings, we compare the accuracy of individual investors' estimates, investors' aggressiveness (as measured by the number of shares they want to trade), and the accuracy of ending market price. We anticipate that individuals in the 3-flip setting will be relatively more overconfident (or, equivalently, less underconfident) than will individuals in the 17-flip setting and that this miscalibration will be reflected in market prices.

The following sections describe in detail the design of our laboratory financial markets and the procedures used in conducting these markets. For discussion of the standard procedures used in experimental economics, see Davis and Holt (1993, chap. 1).

Method

Participants. Participants were 27 MBA students (nine cohorts of three) at Cornell's Johnson Graduate School of Management. The participants had all traded in similar laboratory markets at least once before. They volunteered in response to email announcing the scheduled laboratory sessions.

Securities. The securities used in our markets are closely modeled after the individual decision-making problems presented in Study 1 of Griffin and Tversky (1992) and described in the example in the Background section. Investors were told that we used computer simulations to generate large numbers of two types of coins: "heads-biased" coins, which have a 60% chance of coming up heads, and "tails-biased" coins, which have a 60% chance of coming up tails. We then flipped each coin some number of times (3 or 17) and placed it in a bucket with all of the coins that achieved the same result. For each security, we selected a bucket and set the value of the security in francs (the experimental currency) equal to the proportion of coins (in percent) that are heads-biased in the bucket. Thus, the value of a share of the security (*share value*) is equal to the Bayesian posterior likelihood indicated by a sample of coin flips.

The five securities are shown in Table 1. *Signal weight* is the number of coin

TABLE 1

Securities Used in Experiment 1

# of flips = Signal weight	|#heads − #tails| = Signal difference	Signal strength	Value extremity
17 = High	11 − 6 = 5	14.7	38
17 = High	10 − 7 = 3	8.8	27
17 = High	9 − 8 = 1	2.9	10
3 = Low	3 − 0 = 3	50.0	27
3 = Low	2 − 1 = 1	16.7	10

Note. Signal strength is the absolute difference between the proportion of heads observed and the prior expected proportion of 50%. Value extremity is the absolute distance of the value implied by signal difference from the prior expected value of 50 francs. The value implied by the signal difference is equal to the posterior probability that the coin lands heads 60% of the time, given the result of the coin flips and an equal prior probability of favoring heads 60% of the time and favoring tails 60% of the time.

flips in the sample—more flips provide higher weight. *Signal difference* is the absolute difference between the number of heads and the number of tails. As noted previously, signal difference determines the value of the security. *Signal strength* is the absolute difference between the proportion of heads observed and the prior expected proportion of 50%—more extreme proportions are stronger signals. *Value extremity* is the absolute difference between the security value and the prior expected value of 50 francs.[5]

Experimental design. The experiment uses nine cohorts of three investors each. Each cohort trades the same five securities shown in Table 1. Information strength and weight are manipulated between securities, with three of the securities based on the results of the low-strength, high-weight 17-flip samples and two based on the results of high-strength, low-weight 3-flip samples. To control for order effects, we vary whether the 17-flip or the 3-flip securities occur first and randomize between cohorts the order in which various securities appeared and whether information of a given strength and weight is favorable (more heads than tails) or unfavorable (more tails than heads). About half the securities traded by each cohort have favorable information, and the other half have unfavorable information.

Market structure. The market price of a security is determined by the buying and selling behavior of the investors (buying drives prices up; selling drives prices down). Investors trade securities in a clearinghouse market.[6] In each round of trading, each investor estimates the value of the security and chooses a linear demand schedule by choosing a reservation price and a slope. The reservation price is the price below (above) which the investor wishes to buy (sell) securities. The slope of the demand schedule determines how many shares the investor will buy (sell) for a price that is a given distance below (above) his or her reservation price. Trades are limited to 50 shares per investor per trading round.

Once all three investors have entered a demand schedule, a computer determines a *market-clearing* price at which supply (the number of shares that investors want to sell) equals demand (the number of shares that investors want to buy). If there is a range of such prices, the computer chooses the midpoint of that range as the market-clearing price. Each investor then learns the market-clearing price, the number of shares that he or she traded at that price, the total trading volume, and his or her current balance in shares and cash.

The clearinghouse auction is approximately as efficient as the more common double-auction market (Friedman, 1993; Smith, Williams, Bratton, & Vannoni,

[5] For example, 10 heads and 7 tails results in a signal difference of 3, a proportion of 58.8%, a signal extremity of 8.8, a security value of 77, and a value extremity of 27. Three heads and no tails results in a signal difference of 3, a proportion of 100%, a signal extremity of 50, a security value of 77, and a value extremity of 27.

[6] Other studies using clearinghouse market mechanisms include Friedman and Ostroy (1995), Gillette, Stevens, Watts, and Williams (1999), Van Boening, Williams, and LaMaster (1993), and Bloomfield and Wilks (2000).

1982) and provides similar incentives to estimate value accurately and trade appropriately. However, it has several advantages for this study. The clearinghouse market elicits investors' estimates of value (in the form of a reservation price) as an integral part of trading. Estimates of value can be inferred only indirectly in double-auction markets by observing investors' orders to trade. The clearinghouse market also elicits investors' trading aggressiveness (in the form of the slopes of their demand functions). Steeper slopes indicate that investors are willing to trade more shares for prices that differ from their reservation prices. Aggressiveness is difficult to infer in double-auction markets, because investors can enter many offers to trade but still not be aggressive (if they believe others will not accept their offers); alternatively, they can be quite aggressive without entering many orders (if many of their orders are accepted).[7]

Trading. Investors start trading of each security with 0 shares and 0 francs. To ensure that investors can trade as aggressively as they wish, investors can borrow money at zero interest and can sell shares they do not own.[8] At the end of each round, investors learn the market-clearing price for that round and the number of shares they traded at that price. Trading for each security begins with one round of trading (round 0) before investors receive information about security value (i.e., the sample outcome for the bucket of coins tied to the security being traded).[9] Investors trade for three more rounds after receiving this information, at which point investors move to round 0 of the next security (if any).

When trading of a particular security concludes, investors who hold shares of the security receive the value of the security in an experimental currency which is converted to cash at the end of the experiment. Thus, investors make money by selling shares at prices that are higher than share value and buying shares at prices that are lower than share value. Investors lose money by selling shares at prices that are lower than share value and buying shares at prices that are higher than share value.

Instructions. The instructions are shown in the Appendix. The first 25 minutes of each session are used to review the logistics of the laboratory market (the determination of gains and losses, the trading process, the conversion of trading profits and losses into actual dollars, etc.) and to explain how the security values were determined. To facilitate understanding, two practice

[7] Camerer *et al.* (1989) provide a useful discussion of these problems, which make it difficult for them to test whether less-biased traders are "more active" in a double-auction market.

[8] This is similar to real-world markets wherein investors can borrow cash from banks and/or sell securities they do not own by "short selling" with a promise to buy them in the future.

[9] Trading in round 0 allows subjects to allocate share positions to reflect their risk preferences. For example, risk-seeking subjects can buy shares at prices above 50 from risk-averse subjects, shifting risk from the latter to the former. Including this round of trading should reduce the prevalence of risk-sharing trades in round 2, so that trades in that round more cleanly reflect responses to information.

securities[10] are traded for three rounds each. While the practice securities were traded, an experimenter respond to questions relating to the market mechanism and security valuation. After the two practice securities are traded, additional instructions are distributed which explain the manner in which the actual security values are determined. Upon completing these instructions, the investors begin actual trading and are told that all results will now affect their earnings.

Incentives. Subjects gain or lose francs by trading shares of each security. Gain or loss from trading each round is calculated by multiplying the number of shares a subject purchased by the difference between true security value and that round's current market price. Subjects who purchase (sell) securities earn francs if value is greater (less) than price and lose francs if value is less (greater) than price. Gains from trading are not revealed until trading in all securities is completed, so as not to reveal the true value of the security.

At the end of the experiment, a constant is added to the total number of francs gained or lost, and the sum is multiplied by a conversion rate. A minimum payment of $10 is guaranteed. To avoid risk-seeking behavior by investors whose earnings might leave them near or below the minimum payment, we do not tell subjects either the constant or the conversion rate. We tell them only that average winnings would be approximately $30/session.

Results

Individual judgments (estimates). The fourth column of Table 2 shows errors in investors' estimates of security values for each of the five securities, computed by subtracting security value from investors' estimates. We reverse the sign of the error for securities with values below 50, so that negative errors indicate underreactions (estimates too close to 50) and positive errors indicate overreactions (estimates too far from 50).

Mean estimate error is −15.36 ($t = -16.4$, $p < .0001$) in the 17-flip setting, indicating an underreaction to the low-strength, high-weight information. Investors' estimates are insufficiently extreme in the 17-flip setting. In contrast, mean estimate error is 4.72 ($t = 2.5$, $p = .0182$) in the 3-flip setting, indicating an overreaction to high-strength, low-weight information. Investors' estimates are too extreme in the 3-flip setting. The difference between estimate error in the two settings is also statistically significant ($t = 9.6$, $p < .0001$), indicating that the extent to which investors underreacted increased as information weight increased (and as information strength decreased).

Investors overreacted to their information in the high-strength, low-weight 3-flip setting and underreacted to their information in the low-strength, high-weight 17-flip setting. These results are very consistent with those of Griffin and Tversky (1992), who report median differences between confidence ratings

[10] The practice securities' values are determined by the correct answers (in percent) to almanac-style, business-related questions used in Bloomfield *et al.* (1996).

TABLE 2

Results of Experiment 1

# of flips = Signal weight	|#heads − #tails| = Signal difference	Signal strength	Individual estimate errors	Market price errors
17 = High	11 − 6 = 5	14.7	−22.26	−25.64
17 = High	10 − 7 = 3	8.8	−15.81	−13.69
17 = High	9 − 8 = 1	2.9	− 8.00	− 5.96
Mean high			−15.36	−15.10
			($p < .0001$)	($p < .0001$)
3 = Low	3 − 0 = 3	50.0	3.70	− 2.88
3 = Low	2 − 1 = 1	16.7	5.74	4.01
Mean low			4.72	0.56
			($p = .0182$)	($p = .4294$)

Note. Signal strength is the absolute difference between the proportion of heads observed and the prior expected proportion of 50%. Individual estimate error is equal to (individual estimate of security value) − (security value). Market price error is equal to (market price) − (security value). All p-values are the result of one-sided t-tests. $N = 9$ (each cohort supplied one observation, based on either the average ending individual estimate or the ending market price).

and correct probabilities of −15.5 (compared to −15.36 in our study) in the low-strength, high-weight setting and 5.5 (compared to 4.72 in our study) in the high-strength, low-weight setting.

Aggressiveness. Next, we examine the slope of the investors' individual demand schedules. These slopes reflect investors' aggressiveness, measured by the number of shares the investors are willing to trade for any given deviation of market price from their value estimates. More aggressive traders will tend to choose steeper slopes. In the first (last) trading round of the 3-flip markets, this slope averages −6.2 (−17.1) shares/franc, implying that an investor could buy 6.2 shares at 1 franc above the closing market price, or sell 6.2 shares at 1 franc below that price. The slope in the 17-flip markets is −9.0 (−8.4) shares/franc. These slopes suggest that, in later rounds, investors in the low-strength, high-weight 17-flip markets were trading somewhat less aggressively than were those in the high-strength, low-weight 3-flip markets, again consistent with 17-flip investors being relatively underconfident. However, in no round did slopes differ significantly between 3-flip and 17-flip markets at conventional significance levels. Also, the correlation between the slope of individuals' demand schedules and the accuracy of their estimates is not significant, indicating that more accurate investors did not attempt to trade significantly more shares. While we cannot conclude that less-informed investors were more aggressive than more-informed investors, it is clear that they are not less aggressive.[11]

[11] We also examined trading volume, which reflects the number of interactions between investors that occur during trading. Investors trade an average total volume of 72.8 shares with 3-flip information, and trade an average of 41.3 shares with 17-flip information (the difference in volumes is marginally significant at $p = 0.0644$). This volume data indicates that investors in the high-

Market prices. Our previous analyses replicate Griffin and Tversky's (1992) results for individual judgments. We now conduct market-level analyses to test our prediction that individual biases aggregate to bias market outcomes. The fifth column of Table 2 reports errors in market prices, computed by subtracting security value from the market price obtained in the last round of trading. As with individual estimates, we reverse the sign of the error for securities with values below 50, so that negative errors indicate underreactions (prices too close to 50) and positive errors indicate overreactions (prices too far from 50). Mean price error is −15.10 ($t = -17.6$, $p < .0001$) in the 17-flip setting, indicating an underreaction to the low-strength, high-weight information. Mean price error is 0.56 ($t = 0.2$, $p = .4298$) in the 3-flip setting, indicating an insignificant overreaction to high-strength, low-weight information. The difference between price errors in the two settings is statistically significant ($t = 4.9$, $p = .0004$), indicating that the extent to which market prices underreacted increased as information weight increased (and information strength decreased).

With the exception of an insignificant overreaction in the high-strength, low-weight setting, these results parallel the results for individual estimates. It could be that an even higher-strength, lower-weight signal would produce an overreaction. We believe the inability of overreactions to extend to market prices in our setting may reflect a general tendency of laboratory market prices to react less strongly to information than do individual estimates. Gillette *et al.* (1999) show such a result when all investors share the same information, although the difference between individual estimates and market prices diminishes over time. Bloomfield (1996) shows a similar result when (as in our setting) investors have different information. Bloomfield argues in that paper that traders have difficulty inferring other traders' information by observing their trades, which are noisy indicators of their true judgments. As a result, information that is contained in individual estimates is lost in the process that aggregates those estimates into market prices, even when trading continues for a long time.[12] However, for our purposes, the key finding is that the extent of overreaction increases as strength increases and weight decreases, which we see clearly in the comparison of estimate errors and price errors between the 3-flip and 17-flip settings.

Overall, these results indicate that markets did not correct for individual

---

weight/low-strength 17-flip markets were interacting somewhat less than were those in the low-weight/high-strength 3-flip markets, consistent with 17-flip investors being relatively underconfident or having less initial disagreement in estimates.

[12] Another potential explanation for a lack of overreaction in market prices is that our ability to detect overreactions was limited by ceiling or floor effects or other aspects of the datacollection interface. An overreaction requires that prices be more extreme than value. Given that values were relatively high in the 3-flip treatment (e.g., 77 when signal difference = 3), and that subjects might be reluctant to use the endpoints of the response scales, there may be little room to detect an overreaction which occurs. However, the significant overreactions evident in investors' individual estimates of security value on the same scale indicate that such overreactions are detectable using our methodology.

errors in the trading process by which prices were formed. Thus, as is the case with individual judgments, prices in markets whose investors possessed low-strength, high-weight information tended to reflect relatively more underconfidence than did prices in markets whose investors possessed high-strength, low-weight information.

EXPERIMENT 2

Experiment 1 manipulates information strength and weight between markets in which all investors are identically informed and provides evidence that such identically informed markets respond relatively less to information when it is of low strength but high weight. However, since all of the investors in those markets are identically informed, Experiment 1 does not address the extent to which differences between investors' information strength and weight result in differing levels of confidence and, thus, differing effects on market prices and trading profits.

To address this question, we vary the strength and weight of information within markets in Experiment 2. We examine very simple markets that contain two investors: one investor who has information of high strength but low weight (which we predict will lead to relative overconfidence) and one who has information of low strength but high weight (which we predict will lead to relative underconfidence). In contrast to the tracking and laboratory-markets approaches discussed previously, the information held by each of these investors is independent of the other's information and, therefore, is incrementally useful for determining the value of the security. Thus, it is rational for investors to trade based on their information, but also to temper their trading activity with the knowledge that the other investor also possesses an equal amount of unique value-relevant information.

We hypothesize that, when trading with investors whose information is of low strength but high weight, investors whose information is of high strength but low weight will: (1) exhibit relative overconfidence, (2) trade aggressively enough to bias prices in the direction of their information, creating overconfident market prices, and (3) lose wealth to relatively less-overconfident investors who possess information of low strength but high weight.

Method

Except as stated otherwise, the method in Experiment 2 was the same as that in Experiment 1. As in Experiment 1, immediately after going through instructions investors traded two practice securities (whose values were determined in a manner analogous to the way in which the actual security values were determined).[13] Then they proceeded with trading the securities on which

[13] The practice securities were similar to those used in Experiment 1. The main difference was that each of these securities used two almanac-style questions, instead of one. One investor would see one question, the other investor would see the other question, and the security's actual value was the average of the correct answers to the two questions.

the data analysis is based, with one round of trading (round 0) before information is released and three rounds of trading after.

Participants. Participants were 24 MBA students (12 dyads) at Cornell's Johnson Graduate School of Management. The participants had all traded in similar laboratory markets at least once before. They volunteered in response to email announcing the scheduled laboratory sessions.

Securities. The securities used in Experiment 2 were very similar to those used in Experiment 1, except that the value of each security was based on the sample proportions in two buckets, one of coins flipped three times and one of coins flipped 17 times. The value of each security (in francs) was given by averaging the proportion of coins (in percent) that are heads-biased for the two buckets.

The investors were told that, for each security, we chose two buckets, that they would only see the results of the coin flipping for one of the two buckets for each security, and that the investor they were trading with would only see the results of the coin flipping for the other bucket. The point was emphasized that each investor had some information about each security's value that the other investor did not have and that the other investor had some information which he or she did not have.

As examples, we told investors that, if one bucket has 30% heads-biased coins and the other has 50% heads-biased coins, then the average proportion is 40%, because (30% + 50%)/2 = 40%, and the value of this security would be 40 francs. Similarly, if one bucket has 70% heads-biased coins, and the other has 60% heads-biased coins, then the average proportion is 65%, because (70% + 60%)/2 = 65%, and the value of this security would be 65 francs. Investors were also told that, because the pile from which the coins were drawn had an equal number of heads-biased and tails-biased coins, before they saw how many times the coins in a bucket came up heads and tails, they should expect that bucket to have 50% heads-biased coins. However, once they saw the results of coin flipping, they might want to alter their estimate of this probability for that specific bucket of coins.

Each market traded 12 securities, consisting of two versions of each of the six securities shown in Table 3. In the table, 3-flip difference and 17-flip difference refer to the absolute values of the signal differences (and thus the signal strength) associated with the 3-flip and 17-flip signals, respectively. Value extremity 3-flip difference and value extremity 17-flip difference refer to the distance from 50 of the value implied by the absolute signal difference. The value of a given security is equal to the average of the two values indicated by its signals.

There are four versions of each security possible, depending on the signs of the 3-flip signal and the 17-flip signal. Versions were assigned randomly between markets, with the provision that each signal be positive approximately half the time within each market. The order in which each security appeared was randomized within each market.

TABLE 3

Securities Used in Experiment 2

	High-strength, low-weight signal			Low-strength, high-weight signal		
Security	3-flip difference	Strength	Value extremity | 3-flip difference	17-flip difference	Strength	Value extremity | 17-flip difference
1	3	50.0	27	5	14.7	38
2	3	50.0	27	3	8.8	27
3	3	50.0	27	1	2.9	10
4	1	33.3	10	5	14.7	38
5	1	33.3	10	3	8.8	27
6	1	33.3	10	1	2.9	10

Note. 3-flip difference and 17-flip difference are the absolute values of the signal differences associated with the 3-flip and 17-flip signals, respectively. Strength is the absolute distance of the proportion of heads in the signal from the prior expected proportion of 50. Value extremity | *x*-flip difference is the absolute distance of the value implied by signal difference *x* from the prior expected value of 50 francs. The value implied by the signal difference is equal to the posterior probability that the coin lands heads 60% of the time, given the result of the coin flips and an equal prior probability of favoring heads 60% of the time and favoring tails 60% of the time.

Results

Overview. As discussed previously, we hypothesize that, when trading with investors whose information is of low strength but high weight, investors whose information is of high strength but low weight will (1) exhibit relatively more overconfidence, (2) trade aggressively enough to bias prices, and (3) transfer wealth, harming themselves but benefiting the relatively less overconfident investor. In this section, we address each of these predictions in turn.

Individual judgments (estimates). Since the 3-flip investors have information of high strength but low weight, we expect them to be overconfident relative to 17-flip investors. Overconfidence will encourage 3-flip investors to revise their value estimates more in response to their information than do 17-flip investors.

Testing this requires that we compare the accuracy of investors who hold 3-flip information with that of investors who hold 17-flip information prior to security trading. Therefore, for each security and each investor, we compute an estimate error by subtracting the security's value from the investor's initial estimate of security value.[14] If an investor has a sample that contains more

[14] We focus on initial estimates to insure that we are analyzing independent observations (post-initial estimates could be affected by trading activity). Results are very similar when based on final-round estimates. Results are unchanged if estimate errors are computed using the security value implied by only the information possessed by the investor, rather than actual security value (which is equal to the average of the security values indicated by the information possessed by both members of the cohort). However, when estimating security value, investors would rationally base their estimate not only on their own information, but also on their expectation of the value implied by the information possessed by the other member of their cohort. Therefore, we present estimate errors based on actual security value.

TABLE 4

Means of Individual Estimate Errors, Experiment 2

	Signal weight	
Absolute signal strength	High (17-flip)	Low (3-flip)
1	−2.9	1.4
3	−7.9	3.9
5	−9.9	
Mean	−6.6	2.6
	(p = .0002)	(p = .1823)

Note. Estimate error equals the initial value estimate minus the security value. If the difference between the number of heads and tails given by the participants' signal is negative, the estimate error reported has been multiplied by −1. Thus, negative errors indicate underreactions, and positive errors indicate overreactions. Absolute signal strength is the absolute difference between the number of heads and tails given by a particular signal for a security. Signal weight refers to the number of coin flips, and thus is greater for the 17-flip signal then for the 3-flip signal. All p-values are the result of one-sided t-tests. $N = 12$ (each participant supplied one observation).

heads than tails (indicating a security value of less than 50), we change the sign of the investor's estimate error to ensure that negative estimate errors always indicate underreactions and positive estimate errors always indicate overreactions. For each level of signal strength and each investor, we compute the mean of his or her estimate errors, ensuring that analyses are based on one observation per subject at each level of signal strength.

Results are shown in Table 4. Estimate errors are significantly negative for 17-flip investors ($p = 0.0002$), indicating underreactions, and insignificantly positive for 3-flip investors ($p = 0.1823$). The difference between estimate errors in the two settings is statistically significant ($p = .0037$), indicating that the extent to which investors underreacted increased as their information weight increased (and information strength decreased). Thus, 17-flip investors were underconfident relative to 3-flip investors or, equivalently, 3-flip investors were overconfident relative to 17-flip investors.

Aggressiveness. As in Experiment 1, we examine the slope of the aggregate market-demand schedule, which is analogous to how aggressively on average investors were trading. In the first (last) trading round, the 3-flip investors' slope averages −2.9 (−5.3) shares/franc. The slope of the 17-flip investors is −2.2 (−6.0) shares/franc. In no round did the slopes of the 3-flip and 17-flip investors differ from each other at conventional levels of significance. As in Experiment 1, this result implies that strength and weight affect the extremity of estimates, but do not create differences in how aggressively investors trade on their estimates.[15] Thus, we would not expect biases apparent at the individual level to be eliminated (or exacerbated) at the market level.

[15] We also examine trading volume, which reflects the number of interactions between market participants during trading. Over the three rounds of trading in each security, the two investors in each market trade an average of 16.2 shares. This is somewhat less volume than we observed

Market prices. Given that 3-flip investors are relatively more overconfident than 17-flip investors, and trade just as aggressively, we should see market prices biased in the direction of the value implied by 3-flip investors' information. We compute market price errors by subtracting security value from market price for that security.

Our ability to detect price errors depends in part on whether the 3-flip and 17-flip investors had signals of the same sign. For example, consider a security where investors hold signals of different signs (a *different-sign security*): the signed signal extremity for the 17-flip signal equals -1 and the signed signal extremity for the 3-flip signal equals 1. Underreacting to the negative 17-flip signal would bias price upward, as would overreacting to the positive 3-flip signal. Price would also be biased upward if investors overreacted to both signals, but more to the 3-flip signal, or if investors underreacted to both signals, but more to the 17-flip signal. Consequently, different-sign securities provide a powerful test of whether 3-flip investors are relatively more overconfident than 17-flip investors. More generally, this reasoning suggests that differences between investors in individual over- or underconfidence are likely to affect prices the most when investors are most likely to disagree with each other.

On the other hand, consider a security where the investors hold signals of the same sign (a *same-sign security*): the signed signal extremity for the 17-flip and 3-flip signals equals 1. Overreactions by 3-flip investors would bias price upward, but that would be offset by underreactions by 17-flip investors that bias price downward. Thus, same-sign securities provide a less powerful net effect test, with positive price errors suggesting that 3-flip investors' overconfidence exceeds 17-flip investors' underconfidence, and negative price errors suggesting the opposite. Also, given that same-sign securities will tend to have relatively extreme share values (since both signals are positive or negative), price errors may be biased toward underreaction.

Consequently, we base our analyses on different-sign securities. If the 3-flip (17-flip) signal was negative (positive), the price error reported has been multiplied by -1 so that positive price errors can always be interpreted as overreactions and negative price errors can always be interpreted as underreactions.

Results for the three trading rounds are shown in Table 5. Beginning in round 1 and continuing throughout trading, prices are systematically biased in the direction of 3-flip traders' low-weight information. The overall mean price error is 6.9 in round 1 ($t = 3.04$, one-sided $p = .0056$), 7.0 in round 2 ($t = 2.94$, one-sided $p = .0057$), and 7.3 in round 3 ($t = 3.52$, one-sided $p = .0024$). Prices are biased in the direction of the relatively overconfident 3-flip investors' information.[16]

---

in Experiment 1, consistent with investors being somewhat more cautious when trading with people they know have different information.

[16] Separate analyses of same-sign securities reveal persistently negative price errors. This could be viewed as a net under- and overreaction that provides evidence that 17-flip investors underreact more than 3-flip investors overreact. However, as mentioned above, it also could be driven by the

TABLE 5

Means of Market Price Errors, Experiment 2

	3-flip signal strength	
17-flip signal strength	1	3
	Round 1	
−1	3.7	− 1.9
−3	8.5	3.7
−5	13.6	13.5
	Round 2	
−1	4.4	− 0.2
−3	7.2	6.7
−5	12.7	11.1
	Round 3	
−1	5.1	0.8
−3	7.5	7.3
−5	13.1	10.2

Note. This table includes only those securities for which 3-flip signals and 17-flip signals were of different signs. For those securities where the 3-flip signal was negative (and the 17-flip signal was positive), the sign of both signals was changed to allow a more parsimonious presentation of results. Market price errors equal the market price minus the security value. If the 3-flip (17-flip) signal was negative (positive), the price error reported has been multiplied by −1. One way to interpret price error is that it is the return to a portfolio that sells securities when the 3-flip signal is positive and buys securities when the 3-flip signal is negative. That strategy will produce positive returns if market prices tend to underreact to 17-flip participants' information and/or overreact to 3-flip participants' information. 3-flip signal strength is the absolute difference between the number of heads and tails given by the 3-flip signal. 17-flip signal strength is the negative of the absolute difference between the number of heads and tails given by the 17-flip signal.

Individual welfare. Results previously reported indicate that, relative to the 17-flip investors, the 3-flip investors are overconfident and bias prices in the direction of their information. As a result, we anticipate that 3-flip investors will buy high and sell low to 17-flip investors, systematically transferring wealth.

As with our analyses of price errors, powerful analyses of wealth transfers are possible only with different-sign securities. For example, consider a case where a 3-flip investor has a signal of 1 (implying share value of 60) and a 17-flip investor has a signal of −1 (implying share value of 40). Share value will equal $(60 + 40)/2 = 50$, but prices will be biased in the direction of the 3-flip investor's information. The 3-flip investor will believe the security is underpriced and buy shares, transferring wealth to the 17-flip investor. In general, we predict systematic transfers of wealth by 3-flip investors to 17-flip investors when trading different-sign securities.

On the other hand, with same-sign securities the prediction is less clear. For

relative extremity of the values of same-sign securities, so we consider this evidence to be of very "low weight".

example, consider a case where a 3-flip investor has a signal of 1 (implying share value of 60) and a 17-flip investor has a signal of 1 (implying share value of 60). Share value will equal (60 + 60)/2 = 60. Depending on whether 3-flip investors overreactions are greater or less than 17-flip investors' underreactions, price could be biased upward or downward, and wealth transfers could go in either direction.

Consequently, we base our analyses on different-sign securities. Mean wealth transfers from 3-flip to 17-flip traders are shown in Table 6 for each of the three trading rounds and in total. Wealth transfers are almost always positive, indicating that 3-flip traders are transferring wealth to 17-flip traders. The overall mean wealth transfer is 78.7 in round 1 (t = 1.97, one-sided p = .038), 46.2 in round 2 (t = 1.10, one-sided p = .1498), 60.4 in round 3 (t = 1.49, one-sided p = .0821), and 185.4 overall (t = 1.58, one-sided p = .0713). Beginning in round 1 and continuing throughout trading, 3-flip traders transfer wealth to 17-flip traders. At a conversion rate of $0.01 for each franc, the total wealth transfers shown in Table 6 represent approximately a $10 transfer of wealth, a substantial portion of the average experimental winnings of $30/subject.[17]

TABLE 6

Mean Wealth Transfer from 3-Flip to 17-Flip Participants, Experiment 2

	3-flip signal strength	
17-flip signal strength	1	3
	Round 1	
−1	40.8	48.5
−3	23.0	65.5
−5	86.3	207.9
	Round 2	
−1	10.9	−2.6
−3	11.4	36.6
−5	57.4	163.7
	Round 3	
−1	13.2	59.4
−3	19.5	89.9
−5	42.8	137.8
	Total	
−1	65.0	105.3
−3	53.9	192.1
−5	186.5	509.4

Note. Positive numbers indicate wealth transfers from 3-flip to 17-flip participants. This table includes only those securities for which 3-flip signals and 17-flip signals were of different signs. For those securities where the 3-flip signal was negative (and the 17-flip signal was positive), the sign of both was changed to allow a more parsimonious presentation of results.

[17] Consistent with concerns indicated earlier, separate analyses of same-sign securities reveal generally small wealth transfers with signs varying between rounds.

DISCUSSION

Summary. We report the results of two experiments that examine whether and how the effects of evidential strength and weight identified by Griffin and Tversky (1992) affect behavior in financial markets. Experiment 1 replicates portions of Griffin and Tversky's Study 1 in a market setting. As in Griffin and Tversky's study, we find that information with high strength and low weight leads investors to make more extreme estimates of value than does information of low strength and high weight. However, individuals with more extreme estimates trade just as aggressively as do individuals with less extreme estimates. As a result, the effects of strength and weight on miscalibration in market prices parallel those observed at the individual level.

Experiment 2 extends these results by examining the performance of two-person markets where investors were differentially informed. The results strongly support our findings from Experiment 1. Investors with high-strength and low-weight information had more extreme estimates (i.e., were relatively overconfident), but were as aggressive as were investors with low-strength and high-weight information. Thus, market price was biased toward information with high strength and low weight. In addition, investors with high-strength and low-weight information consistently bought at prices that were too high (when their information was favorable) and sold at prices that were too low (when their information was unfavorable). As a result of this systematic buying high and selling low, investors with high-strength and low-weight information consistently lost money to investors with low-strength and high-weight information.

Implications. Our experiments show that variations in strength and weight can greatly influence individual over- and underconfidence (measured by value extremity), but have no apparent effect on aggressiveness (measured by willingness to trade). Thus, trading decisions do not exacerbate individual bias, but neither do they mitigate it, leaving market prices roughly as biased as are individual judgments.

We believe our results have implications for the performance of cooperative interacting groups as well as markets. Davis (1992, p. 4) notes a number of similarities between cooperating groups and markets where participants interact for purposes of individual profit. For example, the accuracy of group consensus judgments and market prices both depend on accurately weighting individual participants' judgments (Bloomfield *et al.*, 1996; Sniezek & Henry, 1990). Groups also tend to focus on commonly held information (Kim, 1997), as do markets (Bloomfield & Libby, 1996; Bloomfield *et al.*, 1999). Similar to markets, groups are often miscalibrated (see, e.g., Plous, 1995; Sniezek, 1990, 1992). Thus, similar to markets, groups' miscalibration may be predictable from the strength and weight of the information that individual group members possess.

Our paper also contributes to the markets literature. The market setting allows investors to express their aggressiveness unambiguously through their trading decisions, and all investors know that there are explicit monetary incentives that encourage accuracy. Often these characteristics are viewed as encouraging such accurate market prices that individual investors are protected from their own biases. Nonetheless, we find that miscalibration persists in market prices. Our investors did not possess the self-insight necessary to drive away the effects of information strength and weight, so market prices were biased and predictable transfers of wealth occurred between investors. Warning real-world investors of this tendency toward miscalibration and its wealth effects could help reduce unintended wealth transfers, particularly from less-informed investors who lack access to high-weight information. Of course, laboratory markets differ from larger markets (such as the NYSE or AMEX) in many ways. Future studies could examine how information strength and weight predict security mispricing and wealth transfers among investors in such markets.

An interesting direction for future research would be to replicate this study using different market mechanisms. Our clearinghouse markets allow investors to choose a reservation price equal to their estimate of share value and to set a demand schedule that determines how many shares they would trade at various prices. Then, the market pools reservation prices to determine the prices at which shares are traded. We found significant effects of strength and weight on those reservation prices, but no effects on investors' aggressiveness. Other markets present investors with a preset price, execute whatever trades can occur at that price, and move price in response to imbalances in supply and demand. Although traders in such markets presumably still have a reservation price and demand schedule, the only way they can express an effect of strength and weight is via their choice of trade size, so effects on aggressiveness may be more observable in those markets.

Similarly, our results might have been affected by not providing market participants with the enhanced opportunities they would have if they received immediate feedback. Prior research indicates that the benefits of incentives increase with feedback (see, e.g., Sprinkle, 2000). However, we are not sure that the feedback investors receive in real markets is of sufficiently high quality that it would help debias investors. Given the noise in market prices, large price changes are likely to be of high strength but low weight and thus may lead investors to make erroneous adjustments to their behavior.

Future research could also examine trading decisions with respect to more realistic stimuli. We employed the coin-flipping paradigm to enhance comparability with prior work, but it may have affected investors' responses. However, the significance of our treatment effects provides comfort that investors did not disregard them, and the provision of financial incentives helps ensure that our results reflect the economic behavior that would be observed in market settings. Given that miscalibration in market settings has been identified using more realistic stimuli (e.g., Bloomfield *et al.*, 1999), we believe it likely that the effect of strength and weight might generalize to those stimuli as well.

Our results might also be extended to other groups in which the welfare of each investor is tied directly to how well-calibrated that investor is relative to the other investors. For example, sell-side analysts interact through a process similar to the *delphi technique*, repeatedly reporting their individual estimates and revising them in light of new information (including past reports). Recent research has suggested that analysts care more about relative accuracy than absolute accuracy (Mikhail, Walther, & Willis, 1999). Future research could further consider the effects of this incentive structure on individual analyst judgments (for example on herding behavior) and also how the strength and weight of evidence affects these experts in their domain.

APPENDIX

"Clearinghouse" Market: Instructions to Investors

Overview. During this session, you will trade shares of many securities. Each security pays a single final dividend of between 0 and 100 "francs," which are a laboratory currency. This sum is referred to as the security's "true value." You will not be told the true value of a security until trading in that security is over. Whenever you sell a share for more than true value or buy a share for less than true value, you gain francs. Whenever you sell a share for less than true value or buy a share for more than true value, you lose francs. At the end of the session, we convert the number of francs you gain or lose into U.S. dollars.

Entering a demand schedule. In each round of trading, you will choose a demand schedule that states for each possible price how many shares you wish to buy or sell (up to a maximum of 50 shares). You set this demand schedule by choosing two numbers:

1. The price at which you don't want to trade any shares (your *reservation price*);
2. The change in the number of shares you want to trade for every 1-franc decrease in price (the *angle* of your demand schedule);

Choosing these two numbers results in a complete demand schedule.

How trade occurs. Once everyone has entered a demand schedule, a central computer determines the "market-clearing price" at which the total number of shares bought equals the total number sold. If there is a range of market-clearing prices, the computer chooses the middle one. The computer then uses your demand schedule to determine how many shares you bought or sold at that price. You will then see the following information:

(1) the market-clearing price;
(2) the total number of shares traded at the market clearing price ("volume")
(3) the number of shares you personally traded in that period; and
(4) your cumulative holdings in shares and cash through the last round.

At the beginning of trading in each security, you are shown as having no cash and no shares. However, you can still buy or sell up to 50 shares in each round, regardless of the amount of shares and cash that you have. At the end of trading in each security, each share you hold is converted into from 0 to 100 francs, depending on the share's true value.

Estimating the value of the security. In every round of trading, you will be asked to estimate the true value of the security. Please consider this estimate carefully. It is very useful to our research, and will also help you trade more wisely.

Determining your winnings. At the end of trading in a security, you will learn its true value. Every share you trade changes your wealth in francs according to the formulas:

BUYING: Change in Wealth per share = (Value − Price)

SELLING: Change in Wealth per share = (Price − Value)

Your cash winnings will be determined by the number of francs you gained or lost over the course of the session. Francs are converted into cash according to the formula

Dollar winnings = (Gain or Loss in francs + "Floor") × Conversion Rate .

You will not learn the floor or conversion rate. However, these parameters are set so that the average winnings will be approximately $30 per session. If you earn fewer francs than the "Floor" you are given a minimum payment of $10. Note that because the floor is positive, you can win money even if you lose francs. Thus, you can increase your winnings by either winning more francs or losing fewer francs.

Some General Rules

1. *Once the experiment begins, please do not talk with other subjects or look at their computer screens.*
2. *You are not permitted to use paper and pencil, textbooks or calculators at any time during the experiment.*
3. *Please do not discuss any aspect of the experiment with other students until you receive e-mail telling you that you may do so (in about 2 weeks). Such discussions could contaminate our results in future markets.*

I consent to participate in this experiment, and agree to abide by all of the rules determined by the experiment coordinator throughout my participation. I recognize that: (1) if I breach any of the rules governing the market, I forfeit my right to any money I might have earned by participating; (2) I have the right to leave the experiment at any time, without penalty, but that in doing so I forfeit my right to any money I might have earned by trading; (3) this experiment has been approved by the Cornell University Committee on Human Subjects as research that uses no deception of any kind.

Signature ______________________ Date ______________________

Information for Securities 1–2 (Practice Securities) in Experiment 1

To determine the value of each security, we use a question from a general knowledge test that has a correct answer between 0% and 100%. The value of the security (in francs) is equal to the correct answer (in percent). Thus, if the correct answer is 34%, the security will have a value of 34 francs. Everyone will always see the same question just before the second round of trading.

Information for Securities Used in Experiment 1

To determine the value of each security, we used a computer to simulate the random choice of coins from an enormous pile containing two types of coins. One type of coin is called “heads-biased.” When flipped, heads-biased coins have a 60% chance of coming up heads,” and a 40% chance of coming up “tails.” The other type of coin is called “tails-biased.” When flipped, tails-biased coins have a 40% chance of coming up “heads,” and a 60% chance of coming up “tails.” The pile has an equal number of heads-biased and tails-biased coins.

We flipped each coin some number of times. We tossed each coin into a bucket depending on how many heads and tails came up. *The value of each security (in francs) is equal to the proportion of those coins (in percent) that are heads-biased in that particular bucket.*

For example, you could be presented with a bucket of coins containing those that were flipped two times and came up heads once and tails once. If exactly 50% of the coins in this bucket were heads-biased, the value of the security would be 50 francs. If 45% of the coins in this bucket were heads-biased, the value of the security would be 45 francs, and so on.

Because the pile had an equal number of heads-biased and tails-biased coins, before you know how many times the coins in a bucket came up heads and tails, you would expect that bucket to have 50% heads-biased coins. However, once you see the results of coin flipping, you may want to alter your estimate of this probability for that specific bucket of coins: a bucket of coins that come up heads more often than tails is likely to include more heads-biased coins than tails-biased coins. Similarly, a bucket of coins that comes up tails more often than heads is likely to include more tails-biased coins than heads-biased coins.

For each security, we choose a new bucket. All three of the investors in your market will see the same information about the bucket on which that security is based, just before the second round of trading. *Whatever information you see about the bucket is also seen by all of the traders in your market.*

Information for Securities 1–2 (Practice Securities) in Experiment 2

To determine the value of each practice security, we use two practice questions from a general knowledge test that have correct answers between 0% and 100%. e value of the security (in francs) is equal to the average of the two correct answers (in percent). Thus, if the correct answers are 34% and 50%, the security will have a value of 42 francs because (34 + 50)/2 = 42%.

You will only see one of the two questions for each practice security. The investor you are trading with in your market will only see the other question. *Therefore, you have some information about each security's value that the other trader does not have. Similarly, the other trader has some information about each security's value that you do not have.*

Information for Securities Used in Experiment 2

To determine the value of each security, we used a computer to simulate the random choice of coins from an enormous pile containing two types of coins. One type of coin is called "heads-biased." When flipped, heads-biased coins have a 60% chance of coming up "heads," and a 40% chance of coming up "tails." The other type of coin is called "tails-biased." When flipped, tails-biased coins have a 40% chance of coming up "heads," and a 60% chance of coming up "tails." The pile has an equal number of heads-biased and tails-biased coins.

We flipped some coins 3 times, and we flipped other coins 17 times. We tossed each coin into a bucket depending on how many heads and tails came up. Then, for each security, we chose two buckets—one bucket of coins flipped 3 times and one bucket of coins flipped 17 times. *The value of each security (in francs) is equal to the average proportion of coins (in percent) that are heads-biased in its two buckets.*

For example, if one bucket has 30% heads-biased coins, and the other has 50% heads-biased coins, then the AVERAGE proportion is 40%, because (30% + 50%)/2 = 40%. The value of this security would be 40 francs.

Similarly, if one bucket has 70% heads-biased coins, and the other has 60% heads-biased coins, then the AVERAGE proportion is 65%, because (70% + 60%)/2 = 65%. The value of this security would be 65 francs.

Because the pile had an equal number of heads-biased and tails-biased coins, before you know how many times the coins in a bucket came up heads and tails, you would expect that bucket to have 50% heads-biased coins. However, once you see the results of coin flipping, you may want to alter your estimate of this probability for that specific bucket of coins: a bucket of coins that come up "heads" more often than tails is likely to include more heads-biased coins than tails-biased coins. Similarly, a bucket of coins that comes up "tails" more often than heads is likely to include more tails-biased coins than heads-biased coins.

For each security, we choose two new buckets. You will only see the results of the coin flipping for one of the two buckets for each security. The investor you are trading with in your market will only see the coin flipping results for the other bucket. *Therefore, you have some information about each security's value that the other trader does not have. Similarly, the other trader has some information about each security's value that you do not have.*

REFERENCES

Anderson, M. J., & Sunder, S. (1995). Professional traders as intuitive Bayesians. *Organizational Behavior and Human Decision Processes*, 64, 185–202.

Andreassen, P. (1987). On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction. *Journal of Personality and Social Psychology,* 53, 490–496.

Andreassen, P. (1990). Judgmental extrapolation and market overreaction: On the use and disuse of news. *Journal of Behavioral Decision Making,* 3, 153–174.

Bloomfield, R. (1996). Quotes, prices, and estimates in a laboratory market. *Journal of Finance,* 51, 1791–1808.

Bloomfield, R., & Libby, R. (1996). Market reactions to differentially available information in the laboratory. *Journal of Accounting Research,* 34, 183–207.

Bloomfield, R., Libby, R., & Nelson, M. (1996). Communication of confidence as a determinant of group judgment accuracy. *Organizational Behavior and Human Decision Processes,* 68, 287–300.

Bloomfield, R., Libby, R., & Nelson, M. (1999). Confidence and the welfare of less-informed investors. *Accounting, Organizations and Society,* 24, 623–648.

Bloomfield, R. J., R. Libby, & Nelson, M. W. (2000). Underreactions, overreactions and moderated confidence. *Journal of Financial Markets,* 3, 113–137.

Bloomfield, R., & Wilks, T. J. (2000). Disclosure effects in the laboratory: Liquidity, depth, and the cost of capital. *The Accounting Review,* 75, 13–42.

Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Managerial Accounting Research,* forthcoming.

Camerer, C. (1987). Do biases in probability judgments matter in markets-Experimental evidence. *American Economic Review,* 77, 981–997.

Camerer, C. (1992). The rationality of prices and volume in experimental markets. *Organizational Behavior and Human Decision Processes,* 51, 237–272.

Camerer, C., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty,* 19, 7–42.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy,* 97, 1232–1254.

Camerer, C., & Weigelt, K. (1991). Information mirages in experimental asset markets. *Journal of Business,* 64, 463–493.

Cox, J. C., & Grether, D. M. (1996). The preference reversal phenomenon: Response mode, markets and incentives. *Economic Theory,* 7, 381–405.

Curley, S. P., Yates, J. F., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes,* 38, 230–256.

Davis, J. H. (1992). Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples, 1950-1990. *Organizational Behavior and Human Decision Processes,* 52, 3–38.

DiFonzo, N., & Bordia, P. (1997). Rumor and prediction: Making sense (but losing dollars) in the stock market. *Organizational Behavior and Human Decision Processes,* 71, 329–353.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology,* 58, 568–581.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York: Wiley.

Einhorn, H. J., & Hogarth, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review,* 92, 433–461.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics,* 75, 643–669.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance,* 25, 383–417.

Fischer, G. W. (1982). Scoring-rule feedback and the overconfidence syndrome in subjective probability forecasting. *Organizational Behavior and Human Performance,* 29, 352–369.

Friedman, D. (1993). How trading institutions affect financial market performance-Some laboratory evidence. *Economic Inquiry*, 31, 410–435.

Friedman, D., & Ostroy, J. (1995). Competitivity in auction markets-An experimental and theoretical investigation. *Economic Journal*, 105, 22–53.

Ganguly, A. R., Kagel, J. H., & Moser, D. V. (1994). The effects of biases in probability judgments on market prices. *Accounting, Organizations and Society*, 19, 675–700.

Gillette, A. B., Stevens, D. E., Watts, S. G., & Williams, A. W. (1999). Price and volume reaction to public information releases: An experimental approach incorporating traders' subjective beliefs. *Contemporary Accounting Review*, 16, 437–479.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24, 411–435.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4, 5–28.

Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology*, 83, 777–787.

Kagel, J. H., & Levin, D. (1986). The winner's curse and public information in common value auctions. *American Economic Review*, 76, 894–920.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy*, 98, 1325–1348.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 39, 98–114.

Keynes, J. M. (1943). *A treatise on probability*. Great Britain: R. & R. Clark.

Kim, P. H. (1997). When what you know can hurt you: A study of experiential effects on group discussion and performance. *Organizational Behavior and Human Decision Processes*, 69, 165–177.

Koehler, D. J., Brenner, L. A., Liberman, V., & Tversky, A. (1996). Confidence and accuracy in trait inference: Judgment by similarity. *Acta Psychologica*, 92, 33–57.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance*, 20, 159–183.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, UK: Cambridge University Press.

Lind, B., & Plott, C. R. (1991). The winner's curse: experiments with buyers and with sellers. *American Economic Review*, 81, 335–346.

Mikhail, M. B., Walther, B. R., & Willis, R. H. (1999). Does forecast accuracy matter to analysts? *The Accounting Review*, 74, 1–39.

Plous, S. (1995). A comparison of strategies for reducing interval overconfidence in group judgments. *Journal of Applied Psychology*, 80, 443–454.

Sarin, R. K., & Weber, M. (1993). Effects of ambiguity in market experiments. *Management Science*, 39, 602–615.

Smith, V., Suchanek, G., & Williams, A. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, 56, 1119–1151.

Smith, V., & Walker, J. M. (1993). Rewards, experience and decision costs in first price auctions. *Economic Inquiry*, 31, 237–244.

Smith, V., Williams, A., Bratton, W. K., & Vannoni, M. G. (1982). Competitive market institutions:-Double auctions versus sealed bid-offer auctions. *American Economic Review*, 72, 58–77.

Sniezek, J. A. (1990). A comparison of techniques for judgmental forecasting by groups with common information. *Group & Organizational Studies*, 15, 5–19.

Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidences in group decision making. *Organizational Behavior and Human Decision Processes*, 52, 124–155.

Sniezek, J. A., & Henry, R. A. (1990). Revision, weighting, and commitment in consensus group judgment. *Organizational Behavior and Human Decision Processes*, 45, 66–84.

Sprinkle, G. B. (2000). The effect of incentive contracts on learning and performance. *The Accounting Review*, 75, 299–326.

Tuttle, B., Coller, M., & Burton, F. G. (1997). An examination of market efficiency: Information order effects in a laboratory market. *Accounting, Organizations and Society*, 22, 89–103.

Tversky, A., & Kahneman, D. (1971). The belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110.

Van Boening, M. V., Williams, A. W., & LaMasters, S. (1993). Price bubbles and crashes in experimental call markets. *Economic Letters*, 41, 179–185.

Waller, W. S., Shapiro, B., & Sevcik, G. (1999). Do cost-based pricing biases persist in laboratory markets? *Accounting, Organizations and Society*, 24, 717–739.

Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization*, 33, 167–184.

Weber, M., Keppe, H., & Meyer-Delius, G. (2000) The impact of endowment framing on market prices-an experimental analysis. *Journal of Economic Behavior & Organization*, 41, 159–176.

Williams, A. W. (1987). The formation of price forecasts in experimental markets. *Journal of Money, Credit, and Banking*, 19, 1–18.

Received March 24, 2000; published online June 19, 2001