Running Head: ABSOLUTE AND COMPARATIVE PERFORMANCE FEEDBACK

Use of Absolute and Comparative Performance Feedback in Absolute and Comparative Judgments and Decisions

Don A. Moore

Carnegie Mellon University

William M. P. Klein

University of Pittsburgh

NOTE: This is not the copy of record.
This paper appears in *Organizational Behavior and Human Decision Processes*.

We thank Mindy Kelly, Lauren DeVito, and Sam Swift for their assistance in collecting the data reported herein, Rebecca Miller and Jennifer Cerully for their assistance in the preparation of this manuscript, Stephanie Land for her useful advice, and three anonymous reviewers for their helpful suggestions in revising the manuscript.
Correspondence may be sent to Don A. Moore, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 (e-mail: don.moore@alumni.carleton.edu) or to William M. P. Klein at Department of Psychology, University of Pittsburgh, 210 South Bouquet Street, Pittsburgh, PA 15260 (e-mail: wmklein@pitt.edu).

Abstract

Which matters more—beliefs about absolute ability or ability relative to others? This study set out to compare the effects of such beliefs on satisfaction with a performance, self-evaluations, and bets on future performance. In Experiment 1, undergraduate participants were told they had answered 20% correct, 80% correct, or were not given their scores on a practice test. Orthogonal to this manipulation, participants learned that their performance placed them in the 23rd percentile or 77th percentile, or they did not receive comparative feedback. Participants were then given a chance to place bets on two games – one in which they needed to get more than 50% right to double their money (absolute bet), and one in which they needed to beat more than 50% of other test-takers (comparative bet). Absolute feedback influenced comparative betting, particularly when no comparative feedback was available. Comparative feedback exerted weaker and inconsistent effects on absolute bets. Absolute feedback also had stronger (and more consistent) effects on satisfaction with performance and state self-esteem. Experiment 2 replicated these effects in a different university sample, and demonstrates that the effects emerge even when bets are placed *after* participants rate their satisfaction with their performance (although these ratings do not mediate the effect of feedback on bets). These findings suggest that information about one's absolute standing on a dimension may be more influential than information about comparative standing, partially supporting a key tenet of Festinger's (1954) social comparison theory.

Key words: Social comparison, absolute evaluation, relative evaluation, absolute feedback, relative feedback, objective standards, behavioral measures, better-than-average, comparative judgment

Use of Absolute and Comparative Performance Feedback in Absolute and Comparative Judgments and Decisions

In 1954, Leon Festinger coined the term "social comparison" and proposed that people compare themselves with others to make sense of their own outcomes, abilities, and opinions. Festinger also argued, however, that people did not look first to comparative social information to understand what they had, what they accomplished, and who they were. Instead, when objective absolute standards were available, Festinger maintained that the use of those objective standards would obviate the need for comparative social information. Festinger's theory has been one of the most long-lived and influential psychological theories. However, studies building on Festinger's work have generally not tested the fundamental assumption that objective absolute information will trump comparative social information when both are available.

The Power of Comparative Information

Research on social comparison has primarily focused on the effects of different kinds of social comparisons (Mussweiler, 2003; Suls, Martin, & Wheeler, 2002; Suls & Wills, 1991). A great deal of evidence emphasizes the tremendous power of social comparisons, and some of it even suggests that Festinger may have underestimated the degree to which people attend to social comparisons, even when objective absolute information is readily available. For example, Blount and Bazerman (1996) offered participants $7 for 40 minutes' work, and participants knew that other participants would also be paid $7. Seventy-two percent of their respondents agreed to participate. However, when people were offered $8 but told that other participants would get $10,

participation dropped to 54%. People have a fairly good sense of what a dollar is worth. Nevertheless, a reduction in satisfaction brought about by greater wealth when others are still wealthier has been replicated in studies in the lab (Bazerman, Schroth, Shah, Diekmann, & Tenbrunsel, 1994; Loewenstein, Thompson, & Bazerman, 1989) and in the field (Hagerty, 2000; Luttmer, 2005). The implications of these sorts of findings are profound. If people decide how happy they are by comparing their prosperity with that of others, the result is what Brickman and Campbell (1971) called a hedonic treadmill, in which everyone can grow materially better off without improving subjective well-being (see also Frank, 1985; Myers, 2000).

Moreover, research has demonstrated that the power of social comparisons extends well beyond subjective well-being. For example, Klein (1997) manipulated whether participants believed that their performance on an aesthetic judgment task was high or low and above or below average. Participants were then asked to choose between engaging in the same judgment task again – in which they needed to achieve a high absolute score to win a prize – or a chance lottery in which they could also win the prize. Those participants who were told their scores were above average were more likely to choose another trial of the same task (over the chance lottery). As Klein predicted, they appeared to have inferred that if they were better than average, their future scores would be high (in absolute terms). Subjective evaluations of performance, likewise, were only influenced by comparative feedback. These findings, like the hedonic treadmill, suggest that comparative feedback influences judgments in contexts where the kind of objective information identified by Festinger is freely available.

Several other studies that did not explicitly manipulate comparative and absolute feedback in a crossed design have obtained conceptually similar results. Many experiments testing the self-evaluation maintenance (SEM) model show that people are more likely to define themselves according to abilities on which they are better than others than for those on which they have high absolute ability (e.g., Campbell, Fairey, & Fehr, 1986; Tesser & Campbell, 1980). Additional work shows that social comparisons (and reactions to comparative feedback) are largely automatic (Gilbert, Giesler, & Morris, 1995; Mussweiler, 2003), suggesting that it is cognitively effortless to apply comparative information to inferences about absolute self-standing. All these results suggest that comparative information exerts a large influence on absolute evaluations, perhaps even in the context of objective information. Unfortunately, very few studies manipulate comparative and objective information orthogonally – and those studies have numerous limitations – making the magnitude of the influence of comparative vis-à-vis objective information entirely unclear.

The Power of Absolute Information

Other research suggests that knowledge of one's absolute performance overwhelms beliefs about one's standing on a dimension relative to peers. For example, Moore and Kim (2003) found that people playing a competitive game were more confident that they would win the game if the task was easy than if it was difficult, suggesting they failed to adequately consider others, and instead focused on their own solo performances even when explicitly asked to compare themselves with others (for similar results, see Chambers, Windschitl, & Suls, 2003; Hoelzl & Rustichini, 2005; Kruger & Burrus, 2004; Windschitl, Kruger, & Simms, 2003). This "solo comparison

effect" describes the tendency for comparative judgments to be excessively influenced by perceptions of absolute standing. Camerer and Lovallo (1999) argued that budding entrepreneurs underestimate their chances of failure because they focus on their own strengths and abilities, while failing to adequately consider the abilities of their competitors (see also Moore & Cain, 2007; Moore, Oesch, & Zietsma, 2007).

Effects of Absolute and Comparative Feedback on Absolute and Comparative Judgments

The goal of this investigation is to assess the relative effects of objective (or what we will call "absolute") feedback and comparative feedback on both absolute and comparative judgments and decisions. This is the first study of its kind. Despite much research on responses to objective absolute feedback (for a review see Kluger & DeNisi, 1996) and social comparison (for reviews, see Buunk & Gibbons, 2007; Suls & Wheeler, 2000), very little research pits absolute and comparative feedback against each other in a way that allows one to compare their influences. Research on the effects of comparative feedback – as in the classic work testing aspects of social comparison theory and its descendant theories (Festinger, 1954; Suls & Wheeler, 2000) – typically does not employ designs that include both comparative and absolute feedback, and especially not both types of feedback in a crossed design. Studies concerned with the effects of absolute feedback are similarly unlikely to also manipulate comparative feedback in the same design, or to have conditions in which participants get only one type of feedback (absolute or comparative).

Previous studies that *have* examined the effects of comparative feedback have often manipulated it by providing participants with feedback about their own absolute performances as well as the absolute performances of others. Such feedback is often

based on the participants' own feedback, such as "the average person scored 20% higher than you did" (French, Sutton, Marteau, & Kinmonth, 2004; Harris, Sparks, & Raats, 2002; Klein, 1997, 2003a) or is based on a fixed value like "the average person scored a 20%" (French et al., 2004). This sort of manipulation confounds absolute and relative feedback because participants either must be given information about their own performance in order for the comparative information to have the intended effect, or they must take the step of comparing their own score with that of the comparison referent. Instead, our manipulation of comparative feedback varies what participants are told about what percentile rank they fall in, thereby providing them with purely relative information that says nothing about absolute performance. This allows us to include a condition where participants receive *only* relative feedback, allowing us to determine the effects of such feedback on both absolute and comparative judgments independent of absolute standing. Given that much feedback in the real world is entirely comparative (e.g., whether or not wins a competition or not), this condition is a crucial one.

It is tempting to think that we could just to compare effect sizes of each type of feedback across studies. Yet such an endeavor is hamstrung by to tremendous variation in the dependent measures employed in these studies. Dependent measures in most social comparison studies include self-report variables such as affect (Taylor & Lobel, 1989) and self-evaluations (Tesser, 1988), with much less attention to behavioral measures. Few studies discriminate between behaviors that should be differentially influenced by absolute and comparative feedback. There is no one standard set of dependent measures of either comparative or absolute evaluation; this is a problem because evidence suggests that the effects of comparative and absolute feedback differ substantially on slightly

different measures of the same construct (Burson & Klayman, 2005; Klein, 2003b; Moore, 2007b). For example, Klein (1997) found that comparative feedback had a stronger effect on subjective ratings, whereas absolute feedback had a stronger effect on behavioral measures (i.e., participants' willingness to bet on their performances).

In short, no previous studies have manipulated absolute and comparative feedback in an unyoked design (i.e., a design where comparative feedback is not predicated on absolute feedback), in a context where individuals are asked to make both absolute and comparative judgments and choices, and with conditions where some conditions only receive one type of feedback. Such a design is critical not only to testing Festinger's original claim about the priority of objective criteria over social comparison information but also the extent to which each type of information influences inherently comparative judgments and choices (both independently and in concert with the other type of feedback). Finally, no studies address the relative impact of absolute and comparative information on a behavioral variable. Klein (1997) asked participants to place bets, but they were entirely hypothetical. The current manuscript reports two experiments that do not possess any of these limitations.

Hypotheses

We report two studies in which participants were given feedback about their performance on a practice test and then asked to evaluate their performance, place bets on how well they would perform on the actual test, and report their feelings about the feedback. We gave them feedback in absolute terms (the number of questions they got right) and in comparative terms (their percentile rank relative to other test-takers). Participants then placed bets (and rated their confidence in winning those bets) on two

different outcomes. The first was an "absolute" outcome, in that the participant had to answer over 50% of the items correctly on the actual test in order to win. The second was a "comparative" outcome; in this case, a participant would win if his or her score was higher than at least 50% of the other participants. By including equivalent decisions that depended on absolute or comparative ability, it was possible to more directly assess the effects of absolute and comparative feedback on each type of decision.

Our first goal was to replicate the "solo comparison effect" observed by Moore and Kim (2003). In other words, we hypothesized that when participants got only absolute feedback, they would bet more on being better than others when the absolute feedback suggested good rather than bad performance (Hypothesis 1). Similarly, we predicted that comparative feedback (in isolation) would affect beliefs about absolute performance, consistent with research showing the effects of comparative information on absolute judgments (Hypothesis 2).

The more interesting question was how each effect might be influenced by variations in the other type of feedback and whether the effects of absolute feedback or comparative feedback tend to differ in their magnitude. By avoiding the array of limitations outlined above, our experiments address this question directly. More specifically, we can measure the magnitude of each effect by contrasting the effect sizes associated with each type of feedback. Prior research does not provide a sufficient basis to have clear expectations regarding which type of feedback will be stronger. Therefore, we offer two alternative hypotheses. Hypothesis 3 is that absolute feedback will have stronger effects than comparative feedback on predictions of future success, satisfaction with performance, and betting on future performance (a behavioral variable).

Alternatively, Hypothesis 4 predicts that comparative feedback may have stronger (or equivalent) effects. Notably, to assess the relative effects of the two types of feedback, we include a wide range of variables such as bets, affective responses, and state self-esteem. In order to test the robustness of the effects on our behavioral variable (betting), we also vary (across the two experiments) whether this variable comes before or after the various self-report variables. The latter feature allows us to test the ancillary question of whether any of these variables mediate the effects of feedback on betting behavior.

EXPERIMENT 1

Method

Participants

Participants were 415 undergraduate students at Carnegie Mellon University who participated in exchange for monetary payment.

Procedure

Upon arrival, participants were given an instruction sheet with the heading "Guessing weights" and the following instructions:

> "*How good are you at figuring out how much other people weigh? In this exercise, you will be shown a series of pictures of other people and your task will be to guess, within 5 pounds, how much they weigh. You will get $4 for participating in this exercise. So that you can get used to how this game works, we will first have a practice round in which you will see 10 pictures and guess the weights of those 10 people. Then you will get feedback about how you did on the practice test. Then you will take the real test.*"

All participants then saw the 10 pictures and were asked to guess the weights of the individuals in the pictures. After having taken this test, participants received feedback about their performances. In actuality, this feedback was randomly determined and was unrelated to participants' actual performances.

Experimental Manipulations

There were three absolute performance conditions. One group was told that they had answered 2 (or 20%) of the 10 items correctly (low absolute), and another group was told they had gotten 8 (or 80%) of the items correct (high absolute). A third group did not receive any absolute feedback (no absolute). Similarly, there were three comparative performance conditions; participants learned that they had scored better than 23 percent (low comparative) or 77 percent of other test-takers (high comparative), or they received no comparative feedback (no comparative). We selected these numbers to be similar to the absolute feedback numbers, but not identical so as to reduce suspiciousness among participants who got both high absolute and comparative feedback (or low). Crossing these two factors created nine cells in a 3 x 3 between-participants design. The order in which participants received absolute and comparative feedback was counterbalanced.

Feasibility of feedback. Clearly, some of the cells created by this design appear more feasible than others; for example, it may be surprising for someone to learn that a low score of "2" places one in the 77th percentile. In reality, of course, there are some difficult tasks on which even a mediocre performance is enough to outperform others. Moreover, students at a competitive university (such as those in our sample) are likely to have experienced cases where they earned a low score on an exam yet outperformed many other students because the test was particularly difficult and yielded a low mean.

Naturally, it was important that participants in the main experiment viewed all possible combinations as feasible. In order to provide such assurance, we collected data from a sample of 78 Carnegie Mellon students who did not participate in the main study. They were handed a questionnaire that said, "*Please imagine the following scenario. You are shown pictures of 10 people, and you must guess (within 5 lbs.) how much each person weighs. After you complete the task, you are given the following feedback about your performance:*" Each participant was then given hypothetical feedback from one cell in our 3 x 3 experimental design (with the exception of the no-feedback cell, yielding 8 possible combinations of feedback). For example, a participant might read the following:

1. Out of the 10 guesses you made, 2 of them were correct.
2. Your performance is at the 77th percentile, meaning that you performed better than 77% of other CMU students (and worse than 23% of them).

Participants were then asked to rate how pleased they would be about the feedback if they received it about their own performance (on a five-point scale running from "*Not at all pleased*" and "*Extremely pleased*"), how likely it would be for someone (not necessarily them) to achieve this outcome (on a five-point scale with endpoints labeled "*Impossible*" and "*Extremely likely*"), and how surprised they would be if they got this particular combination of feedback (on a five-point scale from "*Not at all surprised*" to "*Extremely surprised*"). Each of these three measures was subjected to a 3 (absolute feedback) x 3 (relative feedback) ANOVA.

As would be expected, participants reported they would be more pleased by being in the 77th percentile group (M = 3.21, SD = .86) than in the 23rd percentile group (M = 1.57, SD = .72), F (2, 70) = 37.77, $p < .001$, $\eta^2 = .51$. There was also a main effect of

absolute feedback, because participants reported that they would be more pleased to have gotten 8 right (M = 3.04, SD = 1.29) than to have gotten 2 right (M = 2.13, SD = .87), F (2, 70) = 11.44, $p < .001$, $\eta^2 = .25$.

For reports of likelihood, there were no significant main effects or interaction effects of the different types of feedback. The mean response on this item across all conditions was 2.82 (SD = .92), close to the "moderately likely" scale point. The means across all conditions were within .5 SDs from this grand mean, with a miniscule effect size (suggesting that a larger sample would still fail to reveal significant effects). The lowest mean was for the 8 correct/23rd percentile condition (M = 2.40) but no one in this condition responded with a "1" (signifying "impossible"). Indeed, only 2 people (3% of our sample) reported that the hypothetical feedback they got was so unlikely as to be impossible, and neither of these people were in the cells that we feared participants would find implausible. Finally, a planned comparison which contrasts the two most "unusual" conditions (2 correct/78th percentile and 8 correct/23rd percentile conditions) with a combination of the other conditions is not significant ($p > .05$).[i] Evidently, participants recognized that any of the combinations was feasible, and equally so.

As for reports of the surprisingness of feedback, there was only one effect: a main effect of absolute feedback, F (2, 70) = 5.73, p = .005, η^2 = .14. This is because those who got 8 out of 10 reported that they would be more surprised (M = 3.68, SD = .98) than those who got 2 of 10 (M = 2.81, SD = 1.12). This is probably due to the fact that students may not consider themselves to be particularly able to estimate weights, and would be surprised by a high score. More importantly, though, the effect of absolute feedback did not interact with the comparative feedback variable, F (3, 70) = 1.46, p =

.23. Thus, in addition to finding the 2 correct/78th percentile and 8 correct/23rd percentile conditions reasonably feasible, participants did not necessarily find these combinations particularly surprising. We do not think this is a function of minimal effort or confusion, and have no reason to believe that participants misunderstood the weight-guessing task or the questions we asked of them. Instead, we attribute this pattern of findings to participants' extensive experience receiving grades in many different contexts where absolute performance was only modestly correlated with comparative performance. These results are reassuring, because they suggest that the cells in our experimental design do not vary with respect to their plausibility, and that, therefore, feedback about performance is unlikely to be confounded with suspiciousness of that feedback.

Dependent Measures

Now back to the main experiment. Participants received the combination of absolute and comparative feedback as determined by their assigned experimental condition. We now turn our attention to the principal dependent measures in our main experiment.

Bets. Participants were now given the option of betting any amount of their $4 earnings on their subsequent performance on the actual test. If they bet and won, the amount of the bet would be doubled. Half of the participants were first asked to specify an absolute bet: "To win this bet, you will need to *get more than 5 of the 10 items correct.* That is, your guesses will have to fall within 5 pounds of the person's real weight on more than 5 of the 10 questions. You may risk up to $2 of your $4 on this bet." Note that this score was 3 points higher than that ostensibly achieved by participants in the low absolute condition, and 3 points lower than that in the high absolute condition.

The other half of the participants were first asked to specify a comparative bet: "To win the bet, you will need to *do better than average on the test.* That is, you will need to score better than at least 50% of the other test-takers. You may risk up to $2 of your $4 on this bet." Again, the criterion (50%) was the midpoint between the comparative scores given to participants in the two feedback conditions (23% and 77%). After participants had made this first bet they were then invited to make the other kind of bet. Those who made the comparative bet first made the absolute bet second, and vice versa. After participants had specified both absolute and comparative bets, they were invited to bet again:

> "Now that you know that there are two types of bets, you may want to re-apportion your money between the two types. On this page, you may revise your bets if you like. *What you write on this page will supersede the bets you have written on the previous two pages. Please do not change what you have written on previous pages*. You still have two betting opportunities, but you are not restricted to a maximum of a $2 bet in either one. You may choose to bet any of your $4 on the two bets. You may even bet all $4 on one of the bets if you choose. The only restriction is that the amount of money you bet and the amount of money you keep must add up to $4."

Our rationale for permitting participants to reconsider their bets was that participants may have failed to pay adequate attention to the nature of their first bet (i.e., whether it was absolute or comparative), and thus would have been reminded of this distinction upon seeing the second bet. As it turned out, there were no systematic

differences between initial bets and reconsidered bets, so as noted later, the two set of bets were collapsed to create a more reliable measure of betting behavior.

After recording their bets, participants completed the remaining dependent measures.

First, participants estimated how many of the 10 items they would answer correctly, and guessed the percentage of other participants who would earn scores lower than their own on the upcoming test. These items were used as manipulation checks; we reasoned that if predictions of future performance were consistent with the feedback that participants received, we can infer that the manipulations were successful.

Next, we included items to assess the impact of the feedback provided to participants on affect, confidence, and state self-esteem in order to examine the breadth of the experimental effects. All of the items were completed on 5-point scales ranging from 1 ("not at all") to 5 ("extremely").

Affective response. To measure affective response, we asked participants to indicate how anxious, concerned, proud, satisfied, and confident they felt as a result of their performance. We did not use a standard affect scale such as the PANAS because we were specifically interested in participants' specific emotional reactions to the feedback (e.g., frustration about their scores) rather than more diffuse emotional states such as anger, and because these scales include other unrelated emotional states.

Confidence. We also asked participants how confident they were regarding their chances of winning the bets they placed, on the same scales as above.

State self-esteem. Finally, to measure state self-esteem, we used 10 performance and social items from Heatherton and Polivy's (1991) state self-esteem scale, which

includes items such as “I feel self-conscious” and “I feel concern about the impression I am making.” We excluded items regarding personal appearance given their lack of relevance to the current study. Heatherton and Polivy (1991) showed that responses to the items we used are influenced by personal feedback such as that received here, and that the scale has acceptable reliability and validity. For example, they report a Cronbach’s α of .92 for the entire scale, and a correlation of .72 with the Rosenberg self-esteem scale.

Our expectation was that low absolute and comparative feedback would be associated with a greater negative affective response, lower self-evaluations, lower confidence, and lower state self-esteem. After completing these measures, participants completed the actual test (which was scored by an experimenter), after which they were paid, debriefed, thanked, and dismissed.

Results

We begin by reporting findings on manipulation checks followed by bets, confidence in bets, affective reactions to the feedback (both negative and positive), and state self-esteem. Most analyses involve 3 x 3 factorial ANOVAs with comparative and absolute feedback condition as the independent variables. The order in which participants got feedback (either absolute or comparative first) did not qualify any results and so all analyses collapsed across order.

Manipulation Checks

Participants’ predictions of future performance show that our feedback manipulations had their intended effect. Absolute feedback had a powerful influence on predicted score, $F\,(2, 411) = 170.12$, $p < .0001$, $\eta^2 = .29$ Absolute score estimates were

highest among those ostensibly scoring an 8 (M = 6.75), followed by those among the no-feedback condition (M = 4.93) and then the 2 condition (M = 3.55); all groups differed by post-hoc Tukey tests, ps < .05.

Comparative feedback also had powerful effects on predicted percentile rank, F (2, 411) = 70.12, p < .0001, η^2 = .15. Participants predicted their percentile rank would be highest in the 77th percentile condition (M = 59.5), followed by participants in the no-feedback condition (M = 42.2) and then those in the 23rd percentile condition (M = 35.9); all groups differed significantly by Tukey post-hoc tests, ps < .05.

Unfortunately, we did not ask participants directly to recall their feedback combination, so these manipulations checks are admittedly indirect. We did include a direct manipulation check in Experiment 2 and found that participants in that study had no trouble remembering the feedback they received, suggesting that was probably the case in the current experiment as well.

Bets

The order in which participants made their bets (absolute or comparative first) produced no significant main or interaction effects. Also, time of bet (first or second bet) did not moderate the effects of absolute or comparative feedback, and first bets were highly correlated with second bets (rs = .72 and .73 for absolute and comparative bets, respectively). Thus, we computed the mean absolute bet (that is, the mean of the first and the final absolute bet) and mean comparative bet (mean of the first and the final comparative bet), and conducted analyses with these measures. Importantly, the effects reported below are identical whether we analyze just initial bets, just final bets, or summed bets.

Our first hypothesis was that in the absence of clear information about comparative performance, information about absolute performance would affect beliefs about comparative performance. As Table 1 shows, when participants did not receive comparative feedback, their comparative bets were significantly higher when they were told their score was an 8 (M = \$1.27) than when it was said to be a 2 (M = \$.52), t (92) = 4.01, $p < .001$, $\eta^2 = .15$. When they lacked comparative information, participants used their absolute scores as a cue for comparative performance – consistent with Hypothesis 1 and the solo comparison effect (Moore & Kim, 2003).

Our second hypothesis was that in the absence of clear information about absolute performance, information about comparative performance would affect beliefs about absolute performance. Our results failed to support Hypothesis 2. Among participants who received no absolute feedback, comparative feedback had no significant influence on absolute bets, F (1, 135) < 1, $p = .77$. This finding fails to support our second hypothesis but is consistent with Klein (1997), who found that comparative feedback had no effect on betting behavior when it was objectively irrelevant, although it did exert a stronger effect on satisfaction with performance.

Our third and fourth hypotheses made alternative predictions regarding which of our two feedback manipulations would be more powerful. Our failure to confirm Hypothesis 2 foreshadows support for Hypothesis 3 over Hypothesis 4: the greater influence of absolute vs. comparative feedback to influence participants' optimism about future performance and satisfaction with that performance. In order to test the overall strength of our feedback manipulations on our various dependent measures, we

conducted a series of omnibus 3 x 3 between-subjects ANOVAs examining each of the key dependent variables in turn, beginning with bets.

The bets were included in a 3 (absolute feedback: No feedback, 2 items correct, or 8 items correct) x 3 (comparative feedback: No feedback, 23rd percentile, or 77th percentile) x (2) (type of bet: absolute or comparative) mixed analysis of variance (ANOVA). The first two factors were between-subjects manipulations whereas the third was a within-subjects factor. Mean absolute and comparative bets are reported in Table 1.

The ANOVA revealed a main effect of absolute feedback, $F(2, 405) = 48.46$, $p < .0001$, $\eta^2 = .11$. Bets were highest among those informed they scored an 8, followed by those given no feedback, followed in turn by those learning they scored a 2. Tukey post-hoc tests showed that all differences among the three conditions were significant ($p < .05$). The effect of absolute feedback interacted with type of bet ($F[2, 405] = 45.97$, $p < .0001$, $\eta^2 = .10$), because the effect of absolute feedback was stronger on absolute than comparative bets (see Table 1).

The omnibus ANOVA also yielded a main effect of comparative feedback, $F(2, 405) = 12.92$, $p < .0001$, $\eta^2 = .03$. Consistent with Hypothesis 3, the effect size for this effect ($\eta^2 = .03$) is significantly lower than the effect size of the absolute feedback variable ($\eta^2 = .11$), $p < .05$. The main effect of comparative feedback was qualified by an interaction with type of bet (absolute or comparative), $F(2, 405) = 40.88$, $p < .0001$, $\eta^2 = .09$, because the effect of the comparative feedback was stronger on comparative than absolute bets. Participants placed higher absolute bets when their comparative performance was lower.[ii]

Confidence in Winning Bets

Absolute feedback appears to have trumped comparative feedback when it came to bets on future performance. We tested whether the same pattern held for participants' confidence in winning their bets in a similar 3 x 3 x (2) mixed ANOVA on participants' confidence. The results of these analyses were similar to those on participants' bets; there were significant main effects of both absolute and comparative feedback (*F*s[2, 402] = 45.06 and 16.98, respectively, *p*s < .0001, $\eta^2 s$ = .10 and .04, respectively), with the effect of absolute feedback being significantly larger than the effect for comparative feedback, $p < .05$.[iii]

Affective Reactions to Feedback

Although absolute feedback appears to have trumped comparative feedback regarding people's confidence regarding future performance and their willingness to bet on it, a number of research findings suggest that people have powerful affective responses to being worse than others (Klein, 1997; Tesser, 1988). In order to assess affective reactions to feedback, we created two composite measures. The first reflected negative affective reactions and was composed of the items measuring anxiety and concern resulting from one's performance (α = .73). The second composite represented positive affective reactions and included the items measuring pride, satisfaction, and confidence based on one's performance (α = .81). The latter item was the first item on the state self-esteem scale but was more highly related to the present items than to the items on that scale (although the findings were the same with or without this item). These two composites were unrelated (r = .06, *n.s.*), consistent with other work

suggesting that positive and negative affective variables need to be treated separately (Watson & Tellegen, 1985).

Each of the composites was then included in a 3 x 3 ANOVA including absolute and comparative feedback as independent variables. The analysis on negative affect revealed only a main effect of absolute feedback ($F[2, 405] = 4.03$, $p < .02$, $\eta^2 = .01$; other Fs < 1, *n.s.*) such that participants receiving a score of 2 experienced more anxiety and concern about their performance (by a Tukey post-hoc test, $p < .05$) than those getting a score of 8 and those getting no feedback (which did not differ themselves, $p > .05$). Participants were disturbed if their scores were low, but seemed unfazed if they were told that their scores were below those of other students (as suggested by the lack of an effect of comparative feedback).

Both comparative and absolute feedback influenced positive affective reactions to performance (Fs[2, 405] = 26.33 and 92.27 for comparative and absolute feedback, respectively, ps < .0001). Notably, the effect size for absolute feedback ($\eta^2 = .18$) was again significantly greater than that for comparative feedback ($\eta^2 = .06$), $p < .01$. Tukey post-hoc tests showed that participants receiving an 8 were more pleased with the performance than those receiving no absolute feedback, who in turn were more pleased than those receiving a 2 (ps < .05). Thus, as is true of many of the variables reported here, the no-feedback group fell in between the two feedback groups, suggesting these participants assumed moderate performance. The effects of comparative feedback were similar; those learning their performance placed them in the 77th percentile were more pleased than those receiving no feedback, who in turn were more pleased than those told they placed in the 23rd percentile (ps < .05).[iv]

State Self-Esteem

Twelve of the 13 items on the state self-esteem scale were collapsed ($\alpha = .73$) into a composite measure (the exception was the item measuring confidence in one's ability, which was included in the earlier composite of positive affective responses). All items were coded such that higher values signified higher state self-esteem. This composite was then submitted to the same 3 x 3 ANOVA as above. As Table 1 shows, variability was limited; most means were over 4 on a 5-point scale. Nevertheless, there was a significant effect of absolute feedback ($F[2, 405] = 8.32$, $p < .001$, $\eta^2 = .02$); a Tukey post-hoc test showed that those told they scored an 8 had higher state self-esteem than those told they scored a 2. The mean in the no-feedback condition was in between these two and was not significantly different from either, ps > .05. These findings once again demonstrate the impact of absolute feedback and support Hypothesis 3. On the other hand, there was no effect of comparative feedback on state self-esteem (nor an interaction, Fs < 1.8, *n.s.*).

Discussion

The first experiment yielded three principal findings. First, when participants were provided only with comparative information, they used to make decisions about their absolute bets, and when provided with only absolute information, they used it to make decisions about their comparative bets. Consistent with research on the fair process effect and fairness heuristic (van den Bos, Wilke, Lind, & Vermunt, 1998), participants seemed to use whatever information was available to them in order to make a best guess. These findings are also consistent with research showing effects of comparative feedback

on absolute judgments (Klein, 1997) and effects of absolute feedback on comparative judgments (Moore & Kim 2003).

The second finding – one that has not been observable in previous studies given various methodological weaknesses including no measure of actual behavior – was that absolute feedback exerted more of an effect on behavior (in this case betting behavior) than did comparative feedback. Both types of feedback did have a significant effect, consistent with predictions that might be drawn from research on the solo comparison effect and on effects of social comparison, but the absolute information was more powerful.

Our third finding – perhaps the most surprising one – was that absolute feedback also maintained a stronger influence than comparative feedback on confidence, affective reactions to the feedback, and state self-esteem. In the case of negative affect and state self-esteem, comparative feedback had no significant effect. These findings are contrary to those of Klein (1997), who found strong effects of comparative feedback and negligible effects of absolute feedback on affect and self-evaluation. This discrepancy is likely a function of differences in the manipulations. Klein's (1997) participants were given percentages as absolute scores, rather than being told how many they answered correctly, and their comparative feedback was yoked to their absolute feedback. Thus, the comparative feedback manipulation was less pure than in the current experiment. Also, participants in that study were given feedback on a dimension about which they already had self-enhancing *a priori* beliefs. In short, the results of this experiment are consistent and clear – absolute feedback about performance is more influential than comparative feedback, much as predicted by Festinger's (1954) original theory. Of

course, the importance of these findings to such a crucial theory – particularly given that they diverge from Klein's (1997) findings – demands a replication.

We designed Experiment 2 with this goal in mind, and took the opportunity in this new study to address some of the limitations of Experiment 1. First, Experiment 1 did not include manipulation checks to confirm that participants were conscious of the absolute and relative feedback when they made their bets. The feedback did influence participants' predictions, but we do not know for sure that they processed their feedback accurately. If they remembered their absolute scores more than their percentile rankings, that could explain the stronger effects of absolute feedback over comparative feedback. Second, the experiment was conducted at a highly competitive university with mathematically-minded students who may be particularly attentive to how they compare with others. The proportional influence of comparative and absolute feedback may be quite different in a less competitive setting, and unusual combinations of relative and absolute feedback (e.g., receiving an absolute score of 2 that is ostensibly at the 78th percentile) could be more misunderstood among students with less of a mathematical focus. Third, it is difficult to know whether the effects of feedback on bets would be obtained had they followed rather than preceded the affective and self-evaluative variables. In Klein (1997), where only effects of absolute feedback were observed, bets were placed after most of the self-report variables. Perhaps comparative feedback is more likely to influence betting behavior when the feedback is proximal, making it more ephemeral than absolute feedback. If such were the case, this would explain the divergence from Klein's (1997) findings. In Experiment 2, then, we measure participants' responses to performance feedback *prior* to making bets. Doing so has the

added benefit of permitting an ancillary test of whether these variables mediate effects of feedback on bets. Because this was an entirely exploratory feature of the study, we did not generate explicit meditational hypotheses.

EXPERIMENT 2

In this experiment we followed the same procedure as that of Experiment 1 with the exceptions discussed above and using a sample from a different university. Also, because we were primarily interested in understanding the greater impact of absolute over comparative feedback, we included only conditions in which participants got both types of feedback. The result was a 2 (absolute feedback: 2 vs. 8 out of 10 correct) X 2 (comparative feedback: 23rd vs. 77th percentile) between-subjects design. Participants were 182 undergraduate students at the University of Pittsburgh who participated in exchange for course credit in an introductory psychology course. We measured betting behavior exactly as in Experiment 1. We included the same manipulation checks, as well as a measure asking them to exactly recall both types of feedback. We omitted the confidence and state self-esteem measures. We did include three items to measure positive affective responses to the feedback including “I feel proud of my performance on the practice test,” “I am satisfied with my performance so far,” and “I feel confident in my abilities at this task,” as well as three items measuring negative affective reactions: “I feel frustrated by my performance on the practice test,” “I am concerned that I will not perform well on the test” and “I am disappointed with my ability to assess people's weights on the practice test.” We also included the items “After doing the practice task, I have a better idea how good I am at guessing people’s weights” and “I am determined to do well on the actual test.” All items were accompanied by 5-point scales from “not at

all" (1) to "extremely" (5). They appeared directly after receipt of feedback, and just prior to placing the bets.

Once again, we predicted significant effects of absolute score and percentile rank on responses to feedback and on betting behavior. Based on Experiment 1, we once again expected absolute feedback to have a greater effect than comparative feedback on all variables, with the exception of comparative bets (which should be influenced more by comparative feedback). Because we believed that the effects of absolute feedback would endure regardless of the proximity of bet placement, we did not alter our hypotheses even though bets were placed after the response variables. Finally, for exploratory reasons, we tested whether the response variables mediated the effects of feedback on betting behavior.

Results and Discussion

As in Experiment 1, neither the order in which participants got feedback (either absolute or comparative first) nor the order in which participants made their bets (either absolute or comparative first) produced significant main or interaction effects, so all the results reported below collapse across order. What follows is a series of 2 x 2 between-subjects ANOVAs examining each of the key dependent variables in turn (manipulation checks, responses to feedback, and bets).

Manipulation Checks

Manipulation checks confirm that our feedback manipulations influenced participants' beliefs about their absolute and relative performances as we had intended. Every single participant in the low absolute feedback condition reported that they had received 2 of 10 correct on the pre-test. And every single participant in the high absolute

feedback condition reported that they had received 8 of 10 correct on the pre-test, consistent with the feedback we gave them.

In the low relative feedback condition, every participant reported that his or her percentile rank was 23rd. In the high relative feedback condition, participants reported, on average, that their percentile rank was 77.2. All but two individuals recounted to us exactly the feedback we had given them. Participants' near-perfect recollection of the feedback allays concerns that their responses and bets were based on subjective evaluations rather than actual feedback.

Another confirmation that the manipulations had their intended effect was that they produced parallel effects on participants' predictions of subsequent performance, as in Experiment 1. Those given high absolute feedback expected to score higher on the real test (M = 6.64, SD = .93) than did those who got low absolute feedback (M = 3.34, SD = 1.02), F (1, 177) = 577.23, $p < .001$, $\eta^2 = .77$. And those given high comparative feedback expected to outscore a higher percentage of others (M = 61.21, SD = 12.45) than did those given low comparative feedback (M = 36.39, SD = 14.99), F (1, 177) = 151.55, $p < .001$, $\eta^2 = .46$. Thus, the manipulations were again effective.

Responses to Feedback

The negative affect items were reliable as a scale ($\alpha = .77$) as were the positive affect items ($\alpha = .84$). Unlike Experiment 1, the positive and negative indexes were highly correlated ($r = .49$, $p < .0001$), so we combined them into an overall affective index ($\alpha = .82$). A 2 (absolute feedback) X 2 (comparative feedback) ANOVA on this composite variable yielded a main effect of absolute score ($F(1, 175) = 158.32$, $p < .0001$, $\eta^2 = .48$) as well as percentile rank ($F(1, 175) = 12.52$, $p < .001$, $\eta^2 = .07$), with the

effect size for absolute feedback being significantly higher ($p < .05$). There was no interaction ($p > .05$). Once again, both types of feedback had an effect, but absolute feedback had more of an effect. Means for each item by condition appear in Table 2.

We conducted a similar 2 X 2 analysis on the other two response measures (how determined and informed participants felt after getting the feedback), and there were no significant effects (Fs < 2.9, ps > .05). The lack of an effect on these measures helps allay fears that our manipulations induced differences in motivations to perform on the actual test.

Bets

As in Experiment 1, we averaged initial and final bets in order to produce a single betting measure for each of the two types of bets (rs = .62 and .47 for the absolute and relative bets, respectively, ps < .0001. Again, the results reported below are the same whether we consider initial bets alone, final bets alone, or the combination of the two bets. As Table 2 shows, the results of the 2 (absolute feedback) X 2 (relative feedback) X 2 (type of bet: absolute vs. comparative) mixed ANOVA parallel those of Experiment 1. As before, there are significant main effects of both absolute feedback (F (1, 177) = 37.60, $p < .0001$, $\eta^2 = .18$) and comparative feedback (F (1, 177) = 20.34, $p < .001$, $\eta^2 = .10$. Once again, absolute feedback seemed to exert more of an influence over bets than did comparative feedback, although in this case the difference in effect sizes did not quite achieve statistical significance ($p > .05$). Participants were once again attentive to the type of bet they were placing, as evidenced by significant interactions between absolute feedback and bet type (F (1, 177) = 106.09, $p < .0001$, $\eta^2 = .38$), and between relative feedback and bet type (F (1, 177) = 53.14, $p < .0001$, $\eta^2 = .23$). These interactions show

that absolute feedback had a stronger influence on absolute than comparative bets, and comparative feedback had a stronger influence on comparative than absolute bets. These results are fully consistent with those of Experiment 1, and emerged even despite the fact that participants placed their bets well after they had obtained feedback (due to the inclusion of response to feedback variables in the questionnaire prior to the bets).

Comparative bets were uncorrelated with the affective response variable computed above ($r = .04$, $p > .05$), suggesting that the effects of comparative feedback on responses and bets were independent. However, absolute bets were significantly correlated with the affect variable ($r = .44$, $p < .0001$), such that more positive affect was associated with higher bets. Given that absolute feedback exerted significant effects on both affective response as well as bets, and given that affective responses were correlated with bets, we then tested for mediation by determining whether the effect of absolute feedback on bets was dampened when controlling for the affective response variable. The relationship between affect and bets fell to insignificance in this analysis, thereby failing to satisfy one of the steps necessary to establish mediation, and suggesting that affective responses cannot explain the strong effect of the absolute feedback on betting behavior.

These findings demonstrate the robust (and broad) influence of both absolute and comparative feedback (and the stronger effects of absolute feedback). At the same time, they suggest the need for further research to identify variables that may elucidate the processes underlying the effects of absolute and comparative feedback on betting behavior, given the lack of mediation displayed here.

In sum, the findings of Experiment 2 replicate those of Experiment 1. When making bets, people reasonably pay more attention to comparative feedback when making comparative bets and to absolute feedback when making absolute bets. Overall, however, their bets are more sensitive to their absolute score than to their percentile ranking, as we had found in Experiment 1. Experiment 2 also showed that absolute feedback had a greater impact on responses to the feedback, much like Experiment 1. Experiment 2 replicates these findings with a sample from a different university, and in a design where bets were placed after responses to feedback. Together, the two experiments reinforce the idea that absolute information may overwhelm the effects of comparative information, at least in a performance context.

GENERAL DISCUSSION

We set out in the current set of studies to put a critical assumption of Festinger's (1954) social comparison theory to the test. In particular, we examined the supremacy of objective information over comparative information in self-evaluation as well as behavior. Although there has been a smattering of attempts to address this question, all of these attempts have had a host of limitations ranging from a failure to manipulate objective (or absolute) feedback and comparative feedback independently, the absence of conditions in which only one source of feedback is present, the lack of behavioral variables, and confusing feedback (for similar arguments see Harris et al., 2002). In the current set of studies, we manipulated absolute and comparative feedback independently, and Experiment 2 showed that participants had no trouble processing this feedback. In Experiment 1, we included conditions where participants only received one type of feedback. Both experiments included bets with real monetary outcomes, allowing us to

identify effects of objective and comparative feedback that go beyond Festinger's predominant focus on self-evaluation and group functioning.

Implications

The findings across multiple variables in both studies (with undergraduates at two different types of universities) were highly consistent. As Festinger (1954) predicted – albeit without empirical backing – absolute feedback exerted a stronger effect than comparative feedback on a wide range of variables including affective responses, state self-esteem, confidence, and (of most interest to us) actual behavior. Importantly, comparative feedback still maintained an effect on many of these variables as well, contradicting the strongest form of Festinger's hypothesis—that objective information obviates the need for social comparison. However, this fact could have something to do with whether the performance scores were viewed as truly objective information (Harris et al., 2002). More importantly, our cleaner design questions the conclusions of previous studies that have attempted to manipulate the two types of information in a factorial design (e.g., Klein, 1997). We also showed that one type of feedback is likely to influence judgments and behaviors that are more relevant to the other type of feedback if the latter is unavailable, consistent with the idea that people use whatever information is available to them.

Another advantage of having conditions with one type of feedback is that they serve as *de facto* control conditions that allow us to determine whether our effects are due to participants getting favorable feedback, unfavorable feedback, or both. In Experiment 1, the no-feedback control group was sometimes significantly different than both feedback groups and other times only different than one group, but in all cases this group

fell in between the group that was given a low score and the group that was given a high score. Thus, the locus of these effects seemed to occur in *both* feedback groups, an observation that would not have been possible without the inclusion of this no-feedback group. Evidently, Festinger's assumption is not moderated by level of performance, another crucial finding. The addition of these control groups also permitted us to show that when people are exposed to comparative feedback in a context where they must make an absolute judgment orbehavior – or vice-versa – they use the information available to them rather then rendering it irrelevant.

As a group, these findings are important not only because they test a crucial assumption of a seminal theory in the field, but also because they begin to link together two literatures that have yet to be integrated satisfactorily. Some work shows that absolute standing on a dimension may influence judgments on dimensions tied to comparative standing (e.g., Moore & Kim, 2003), and that comparative standing may influence judgments tied to absolute standing (Klein, 1997). The present study examined the magnitude of each of these effects in the same context, and did so across a range of variables. Thus, it was possible to examine the generalizability of these two effects, and more importantly, compare their relative strengths.

Our results are consistent with recent research showing the power of absolute outcomes to drive social comparisons. For instance, even when they know it will be graded on a forced curve, students' expectations for getting an A on a test go down significantly as the test gets more difficult (Windschitl et al., 2003). When an event is highly unlikely (such as being struck by lightning or living past 100), people assume that they will be below average (Chambers et al., 2003; Kruger & Burrus, 2004). People

believe that they are below average on difficult tasks and above average on simple ones (Moore & Small, 2007). Note that participants in the current study did not use absolute feedback indiscriminately; on the contrary, effects of absolute feedback on comparative bets and inferences about comparative performance were muted, at least relative to absolute bets and inferences about absolute performance.

On the other hand, the effects of comparative feedback in these two studies were weaker than the effects of absolute feedback. Although comparative feedback had predictable effects on the variables it should have influenced – namely comparative bets, perceived likelihood of winning the comparative bet, and estimated percentile (which represents one's comparative standing) – it had weaker and sometimes no effects on other variables. Comparative feedback rarely influenced absolute bets, and had no effects on perceived likelihood of or confidence in winning the absolute bets. These results are consistent with those of Klein (1997), who found that comparative feedback was more likely to influence self-evaluative and affective variables than behavioral variables. However, in the current study, even the latter variables were not clearly affected by comparative feedback. Comparative feedback had no significant effect on negative affect (anxiety and concern) or on the composite self-esteem scale. This type of feedback did influence positive reactions such as satisfaction with performance, but the effect size was significantly lower than that for absolute feedback. We have offered many possible explanations for this discrepancy with the findings of Klein (1997); key among them is the fact that comparative feedback in the Klein (1997) studies was yoked to absolute feedback – possibly masking how much of the effect was truly due to the comparative

portion of the feedback. This is of course an empirical question that deserves further attention.

Underlying Causes

Why did absolute feedback produce much stronger and more consistent effects than comparative feedback? We see three clear possibilities. The first is the fairly rational explanation suggested by Festinger (1954): Objective information more informative, useful, and generalizable than is comparative social information. We purposely chose a task for which participants did not know, after completing the task, how well they had done, in order to maximize the impact of the feedback they received. We expected that participants were unlikely to have prior beliefs one way or the other about their skills in the tested domain. This is exactly the situation that Festinger (1954) believed would instigate the most social comparison, given that such comparison can disambiguate personal ability in an unfamiliar domain. However, later work suggests that much of social comparison is directed at confirming cherished positive views of the self, and that people are persuaded more by favorable comparative feedback than by unfavorable comparative feedback (e.g., Dunning & Hayes, 1996; Weinstein & Klein, 2001; Wood, 1989). If self-enhancement is a primary goal of social comparison, it makes sense that comparative feedback will have less impact in cases where self-relevance is low. In the current study, participants who received no comparative feedback did not believe they were better than average, suggesting that participants did not hold favorable views regarding this skill that they were motivated to protect. By contrast, the task in Klein's (1997) studies concerned esthetic judgment, a skill that participants considered desirable and for which they believed they possessed above average skills. Clearly, it

will be important to replicate the current findings in a context where self-relevance of the ability varies orthogonally with feedback on the associated task.

The second possibility is a less sensible one: That comparative judgments overweight information about the self, perhaps due to egocentrism (Kruger, 1999; Windschitl et al., 2003). When people are led to focus on others, the standard effect of absolute feedback is weakened or eliminated (Moore & Kim, 2003; Windschitl et al., 2003). Purely comparative feedback, such as the percentile rank feedback we provided our participants, weighs self and others appropriately, and is not open to egocentric weighting.

A third possibility is that we "cooked the books" because our manipulation of comparative feedback was inherently weaker than our manipulation of absolute feedback. In particular, the difference between being in the 23rd and 77th percentile may have been perceived to be smaller than the difference between scoring a 2 or an 8 on the practice test. On the other hand, given that the comparative feedback told participants something meaningful about where their performance on an unfamiliar task stood relative to that of others, it is easy to argue that it should still have been more influential. The absolute feedback manipulation was primarily a function of the difficulty of the test, which had been arbitrarily based on a 5-pound difference between a participant's answer and the correct weight. The task would have been substantially more difficult if that range had been 1 pound, making the scores lower, and the task would have been substantially simpler if that range had been 30 pounds. Nevertheless, absolute feedback had a more potent and consistent effect across a range of dependent variables, consistent with Festinger's (1954) original theory.

One important conclusion to be taken from the current data is that people are quite adept at distinguishing between the implications of absolute and comparative feedback. Comparative feedback yielded strong effects on comparative bets and judgments (such as estimated percentile), and little or no effect on absolute bets and judgments. Absolute feedback had much stronger effects on absolute than on comparative bets and judgments. The evidence that people are attuned to this difference suggests that past work may have exaggerated the extent to which people make inappropriate inferences from performance feedback. Such careful distinctions may be less likely in situations characteristic of high cognitive load. On the other hand, they may be more likely in cases where outcomes are highly consequential or where participants are answering questions on vague scales that make it easy to conflate comparative with absolute evaluation (Moore, 2007a). Future research needs to explore these possibilities.

Summary

The current study is the first to address the effects of absolute and comparative feedback on both absolute and comparative judgments, and in the same context. The inclusion of no-feedback groups is also a unique feature of the study, and allowed for an examination of how people respond to only one type of feedback when forced to make judgments and decisions that are associated with the other type of feedback. As a result, the various conditions we created in this experiment represent many of the feedback scenarios people face in everyday life. Sometimes they may know their absolute standing on a dimension (e.g., running ability) and must make inferences about how they will fare in a competitive setting (e.g., a marathon). Other times they may be aware only of their comparative standing – knowing for example, that they were selected above others for a

promotion at work – and must make inferences about the underlying absolute performance that put them ahead of others. In many other situations, both types of feedback are available and must be weighed accordingly. To the extent that people make these inferences carefully and rationally, decision-making will lead to favorable consequences. We hope that the two experiments presented here will not only contribute to extant theories of how people respond to feedback and engage in social comparison but also help in the development of aids to promote rational decision making.

References

Bazerman, M. H., Schroth, H. A., Shah, P. P., Diekmann, K. A., & Tenbrunsel, A. E. (1994). The inconsistent role of comparison others and procedural justice in reactions to hypothetical job descriptions: Implications for job acceptance decisions. *Organizational Behavior and Human Decision Processes, 60*(3), 326-352.

Blount, S., & Bazerman, M. H. (1996). The inconsistent evaluation of absolute versus comparative payoffs in labor supply and bargaining. *Journal of Economic Behavior and Organization, 30*(2), 227-240.

Brickman, P., & Campbell, D. T. (1971). Hedonic relativism and planning the good society. In M. H. Appley (Ed.), *Adaptation level theory: A symposium*. New York: McGraw-Hill.

Burson, K. A., & Klayman, J. (2005). *Judgments of performance: The relative, the absolute, and the in-between*. Unpublished manuscript, Ann Arbor. Available at SSRN: http://ssrn.com/abstract=894129

Buunk, B. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes, 102*(1), 3-21.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review, 89*(1), 306-318.

Campbell, J. D., Fairey, P. J., & Fehr, B. (1986). Better than me or better than thee? Reactions to intrapersonal and interpersonal feedback. *Journal of Personality, 54*, 479-493.

Chambers, J. R., Windschitl, P. D., & Suls, J. (2003). Egocentrism, event frequency, and comparative optimism: When what happens frequently is "more likely to happen to me". *Personality and Social Psychology Bulletin, 29*(11), 1343-1356.

Dunning, D., & Hayes, A. F. (1996). Evidence of egocentric comparison in social judgment. *Journal of Personality and Social Psychology, 71*, 213-229.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140.

Frank, R. H. (1985). *Choosing the right pond : human behavior and the quest for status*. New York: Oxford University Press.

French, D. P., Sutton, S. R., Marteau, T. M., & Kinmonth, A. L. (2004). The impact of personal and social comparison information about health risk. *British Journal of Health Psychology, 9*, 187-200.

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology, 69*(2), 227-236.

Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology, 78*(4), 764-771.

Harris, P., Sparks, P., & Raats, M. (2002). Theoretical and applied issues in the provision of absolute and comparative risk information. *Risk Decision and Policy, 7*, 153-163.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*(6), 895-910.

Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do you put your money on it? *Economic Journal, 115*(503), 305-318.

Klein, W. M. P. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology, 72*(4), 763-774.

Klein, W. M. P. (2003a). Effects of objective feedback and 'single other' or 'average other' social comparison feedback on performance judgments and helping behavior. *Personality and Social Psychology Bulletin, 29*(3), 418-429.

Klein, W. M. P. (2003b). Self-prescriptive, perceived, and actual attention to comparative risk information. *Psychology and Health, 18*, 625-643.

Kluger, A. N., & DeNisi, A. S. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 199*(2), 254-284.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77*(2), 221-232.

Kruger, J., & Burrus, J. (2004). Egocentrism and focalism in unrealistic optimism (and pessimism). *Journal of Experimental Social Psychology, 40*(3), 332-340.

Loewenstein, G., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology, 57*, 426-441.

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics, 120*(3), 963-1002.

Moore, D. A. (2007a). Not so above average after all: When people believe they are worse than average and its implications for theories of bias in social comparison. *Organizational Behavior and Human Decision Processes, 102*(1), 42-58.

Moore, D. A. (2007b). When good = better than average. *Judgment and Decision Making, 2*(5), 277-291.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior & Human Decision Processes, 103*, 197-213.

Moore, D. A., & Kim, T. G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology, 85*(6), 1121-1135.

Moore, D. A., Oesch, J. M., & Zietsma, C. (2007). What competition? Myopic self-focus in market entry decisions. *Organization Science, 18*(3), 440-454.

Moore, D. A., & Small, D. A. (2007). Error and bias in comparative social judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology, 92*(6), 972-989.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*(3), 472-489.

Myers, D. G. (2000). The funds, friends, and faith of happy people. *American Psychologist, 55*(1), 56-67.

Suls, J. M., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science, 11*(5), 159-163.

Suls, J. M., & Wheeler, L. (Eds.). (2000). *Handbook of social comparison: Theory and research*. New York: Kluwer Academic/Plenum Publishers.

Suls, J. M., & Wills, T. A. (Eds.). (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review, 96*(4), 569-575.

Tesser, A. (1988). Toward a self evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). New York: Guilford.

Tesser, A., & Campbell, J. (1980). Self-definition: The impact of the relative performance and similarity of others. *Social Psychology Quarterly, 43*(3), 341-346.

van den Bos, K., Wilke, H. A. M., Lind, E. A., & Vermunt, R. (1998). Evaluating outcomes by means of the fair process effect: Evidence for different processes in fairness and satisfaction judgments. *Journal of Personality and Social Psychology, 74*(6), 1493-1503.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Weinstein, N. D., & Klein, W. M. P. (2001). Resistance of personal risk perceptions to debiasing interventions. In T. Gilovich, D. W. Griffin & D. Kahneman (Eds.), *Heurisics and biases: The psychology of intuitive judgment* (pp. 313-323). Cambridge, England: Cambridge University Press.

Windschitl, P. D., Kruger, J., & Simms, E. (2003). The influence of egocentrism and focalism on people's optimism in competitions: When what affects us equally affects me more. *Journal of Personality and Social Psychology, 85*(3), 389-408.

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin, 106*(2), 231-248.

Footnotes

Table 1

Exepriment 1: Bets, confidence in each bet, affective responses, and state self-esteem by condition

	Absolute feedback (score out of 10)								
	2			None			8		
Comparative Feedback (percentile rank):	23rd (n = 45)	None (n = 48)	77th (n = 46)	23rd (n = 43)	None (n = 49)	77th (n = 46)	23rd (n = 46)	None (n = 46)	77th (n = 45)
Absolute bets (out of $4)	$0.48 (0.69)	$0.41 (0.70)	$0.36 (0.52)	$0.78 (0.73)	$0.86 (0.72)	$0.76 (0.65)	$1.82 (0.84)	$1.74 (0.89)	$1.30 (0.82)
Comparative bets (out of $4)	$0.69 (0.76)	$0.57 (0.69)	$1.65 (0.91)	$0.94 (0.94)	$1.15 (0.87)	$1.97 (0.87)	$0.70 (0.65)	$1.20 (0.83)	$1.38 (0.83)
Absolute bet confidence (0-5)	1.95 (1.24)	1.71 (1.03)	1.74 (0.93)	2.36 (1.03)	2.53 (1.16)	2.78 (1.19)	3.61 (0.95)	3.50 (0.98)	3.33 (1.00)
Comparative bet confidence (0-5)	2.53 (1.10)	2.33 (0.98)	3.50 (0.78)	2.37 (1.07)	2.96 (1.08)	3.78 (0.84)	2.46 (1.26)	3.07 (1.18)	3.49 (0.66)
Negative affect	2.20 (1.10)	2.26 (1.04)	2.16 (0.91)	2.40 (0.99)	2.60 (1.08)	2.50 (0.98)	2.71 (1.01)	2.43 (1.07)	2.42 (0.92)
Positive affect	2.07 (0.88)	2.03 (0.62)	2.56 (0.83)	2.25 (0.82)	2.80 (0.76)	3.49 (0.76)	3.32 (0.76)	3.53 (0.72)	3.59 (0.71)
State self-esteem	4.16 (0.56)	4.16 (0.48)	4.33 (0.57)	4.32 (0.49)	4.36 (0.52)	4.40 (0.48)	4.39 (0.59)	4.52 (0.41)	4.49 (0.43)

Note: Standard deviations appear in parentheses.

Table 2

Experiment 2: Bets and response variables by condition

	Absolute feedback (score out of 10)			
	2		8	
Comparative feedback (percentile rank)	23rd (n = 44)	77th (n = 45)	23rd (n = 45)	77th (n = 47)
Absolute bets (out of $4)	$0.66 (0.73)	$0.52 (0.47)	$1.92 (0.81)	$1.71 (0.72)
Comparative bets (out of $4)	$0.85 (0.63)	$2.00 (0.82)	$0.84 (0.77)	$1.46 (0.72)
I feel proud of my performance on the practice test.	1.73 (0.9)	2.33 (0.88)	3.42 (0.94)	3.43 (0.97)
I am disappointed with my ability to assess people's weights on the practice test.	2.66 (1.10)	2.29 (1.18)	1.51 (0.82)	1.26 (0.53)
I am satisfied with my performance so far.	2.20 (0.85)	2.73 (0.72)	3.67 (0.80)	3.89 (0.70)
I feel confident in my abilities at this task.	1.73 (0.76)	2.27 (0.89)	3.00 (0.77)	3.17 (0.73)
After doing the practice task, I have a better idea how good I am at guessing people's weights.	3.36 (0.92)	3.02 (1.20)	3.38 (0.83)	3.30 (0.98)
I feel frustrated by my performance on the practice test.	2.16 (1.22)	1.69 (0.97)	1.24 (0.61)	1.13 (0.49)
I am determined to do well on the actual test.	3.30 (0.95)	3.16 (0.98)	3.56 (0.94)	3.45 (1.23)
I am concerned that I will not perform well on the test.	2.39 (1.24)	2.24 (1.07)	1.93 (1.03)	1.81 (0.85)

[i] We thank an anonymous reviewer for suggesting this analysis.

[ii] There was also an interaction between absolute and comparative feedback (F[2, 405] = 4.55, $p < .001$, $\eta^2 = .01$), although it was not qualified by bet type ($F = 1.1$, *n.s.*). As shown in Table 1, when absolute feedback was high (8), there was no effect of comparative feedback ($p > .10$). When absolute feedback was absent or low (2), bets were higher when comparative feedback was high (77th percentile) than when it was absent or low (23rd percentile), ps < .05.

[iii] We also obtained a significant interaction between the two (F[4, 402] = 2.99, $p < .02$, $\eta^2 = .01$), and an interaction between each type of feedback and type of bet (Fs[2, 402] > 31, ps < .0001, $\eta^2 s > .07$). Table 1 shows the nature of these interactions. As they are not central to the hypotheses, we do not discuss them further.

[iv] This analysis also yielded an interaction between absolute and comparative feedback (F[4, 405] = 5.75, $p < .001$, $\eta^2 = .01$), such that the effects of comparative feedback were strongest when no absolute feedback was available, and absent when participants learned they had scored an 8 on the practice test (in which case their reactions were equivalently high regardless of comparative feedback). Again, this interaction was not relevant to our main hypothesis.