

Abstract
People with low self-esteem (LSE) seem to focus on self-protection; rather than trying to achieve gains for their self-esteem, they try to avoid losses. This research examined, in a social comparison context, the hypothesis that LSE Ss seek self-enhancement when they have an opportunity that is "safe," that is, carrying little risk of humiliation. Exps 1 and 2 indicated that LSE Ss sought the most social comparisons after receiving success feedback, whereas high-self-esteem (HSE) Ss sought the most comparisons after failure. Further results suggested that LSE Ss who succeeded were seizing a safe means of self-enhancement and that HSE Ss who failed were seeking to compensate for the failure. Also supporting this interpretation for LSE Ss was Exp 3, in which LSE Ss who succeeded sought the most comparisons when such comparisons promised to be favorable. All 3 studies illustrate the value of a new measure of social comparison selection. (PsycINFO Database Record (c) 2010 APA, all rights reserved)

Full Text

Contents
• Abstract
• Experiment 1
• Method
• Pretest Phase
• Experimental Phase
• Subjects and design
• Procedure
• Success–failure manipulation
• Ratings of self and other
• Comparison-selection measure
• Jealousy, envy, and mood
• Probing for suspicion and debriefing
• Results
• Manipulation Checks
• Success in future career
• Jealousy, envy, and mood
• Avoiding Versus Selecting Comparisons
• Subjects’ Ratings of Self and Other on the Test-Selection Dimensions
• Discussion
• Self-Protection and Self-Enhancement Among Low-Self-Esteem Subjects
• Comparison Choices Among High-Self-Esteem Subjects
• Mood Effects After Social Comparisons
• Does Our Comparison Measure Truly Reflect Social Comparison?
• Interview Study
• Experiment 2
• Method
• Pretest Phase
• Experimental Phase
• Subjects and design
• Procedure
• Success–failure manipulation
• Comparison-selection measure
• Expectations measure
• Mood measure
• Beliefs about the validity of the test
• Results
• Manipulation Checks
• Avoiding Versus Seeking Comparisons
• A Different Way of Measuring Degree of Comparison-Seeking
• Were Subjects Who Sought Comparisons Puzzled or Skeptical?
• Exploring Motivations Underlying Comparisons
• Expectations for self and other
• Relation between expectations and selection of matches
• Relation between mood and selection of matches
• Discussion
• Avoiding Versus Seeking Comparisons
• Motivations Underlying Comparison Choices
• High-self-esteem failure
Low-self-esteem success
Exploration of motivations
Self-Protection Among Low-Self-Esteem Subjects?
Experiment 3
Method
Pretest Phase
Subjects and Design
Procedure
Comparison-Selection Measure
Results and Discussion
Manipulation Checks
Avoiding Versus Selecting Comparisons
General Discussion
A New Comparison-Selection Measure
Avoiding Versus Selecting Comparisons
Comparisons After Failure: Self-Protection and Compensation
Comparisons After Success: Lack of Interest and Revealing
Beyond Self-Protection Among People With Low Self-Esteem

Figures and Tables
Table 1
Table 2
Table 3
Table 4
Table 5
Table 6
Table 7

Abstract
People with low self-esteem (LSE) seem to focus on self-protection; rather than trying to achieve gains for their self-esteem, they try to avoid losses. This research examined, in a social comparison context, the hypothesis that LSEs seek self-enhancement when they have an opportunity that is “safe,” that is, carrying little risk of humiliation. Experiments 1 and 2 indicated that LSEs sought the most social comparisons after receiving success feedback, whereas high-self-esteem Ss (HSEs) sought the most comparisons after failure. Further results suggested that LSEs who succeeded were seizing a safe means of self-enhancement and that HSEs who failed were seeking to compensate for the failure. Also supporting this interpretation for LSEs was Experiment 3, in which LSEs who succeeded sought the most comparisons when such comparisons promised to be favorable. All three studies illustrate
the value of a new measure of social comparison selection.

Who is more likely to engage in strategies to benefit their self-esteem—people with dispositionally low self-esteem or high self-esteem? Either possibility is plausible: People with low self-esteem may use more esteem-bolstering strategies, because they need them more; people with high self-esteem may do so, and that may be precisely why they are high in self-esteem. Research supports each position, although the weight of evidence presently favors the view that high-self-esteem people are more likely to engage in self-enhancement (see Taylor & Brown, 1988, for a review).

Baumeister, Tice, and Hutton (1989) offered a third view. They proposed that both high- and low-self-esteem people attempt to benefit their self-esteem, but that their specific goals and strategies differ (cf. Brown, Collins, & Schmidt, 1988). Specifically, people with high self-esteem seek self-enhancement; they seek ways of drawing attention to their skills and talents. They are bold people, willing to take risks to achieve gains for their self-esteem. In contrast, people with low self-esteem aim for self-protection; they “focus not on their good points but on trying to minimize their weaknesses” (Schlenker, Weigold, & Hallam, 1990, p. 856). Because their main goal is to avoid exposing their unfavorable characteristics, they steer clear of challenges that may bring rewards, but that also carry the risk of revealing their flaws. This view of the cautious, self-protective, low-self-esteem person is growing (Arkin, 1981; Raynor & McFarlin, 1986), as is supportive evidence (Tice, 1991; Wolfe, Lennox, & Cutler, 1986; see Baumeister, 1993; Baumeister et al., 1989; for reviews).

In the present research, we examined self-protection and self-enhancement in a social comparison context. Low- and high-self-esteem subjects experienced a failure or a success, and then had an opportunity to compare themselves with the very person against whom they had just failed or succeeded. We predicted that low-self-esteem people would avoid comparisons after failing. That would be the safe, self-protective course, because continuing to compare themselves with a superior person may further wound their self-esteem. The possibility that people may avoid social comparisons has received more theoretical than empirical attention (Brickman & Bulman, 1977; Wood, 1989). However, there is some evidence that when people experience a threat to self-esteem or when they fear public exposure of their inferiority, they do avoid upward comparisons—comparisons with others who are superior (e.g., Pyszczynski, Greenberg, & LaPrelle, 1985). This is especially true of people with low self-esteem (see Wood & Taylor, 1991, for a review). Some research is consistent, then, with our prediction that after failing, low-self-esteem subjects will avoid comparisons with a person who had succeeded.

Although low-self-esteem people may be oriented mainly toward self-protection, we reasoned that they may sometimes venture beyond self-protection and seek self-enhancement, but only when it is safe to do so. “Safe” opportunities for self-enhancement are those in which a favorable outcome is virtually guaranteed. Some research supports this idea. People who are low in self-esteem or high in fear of failure seek further information about themselves when it promises to be favorable—such as in areas that they regard as their strengths (Swann, Pelham, & Krull, 1989), or after already receiving favorable feedback in the same domain (
Willerman, Lewit, & Tellegen, 1960). In addition, although they usually give up more readily than do high-self-esteem people, low-self-esteem people exert special effort when the task is easy, or when they have received success feedback (Brockner, 1979; Shrauger & Sorman, 1977).

Such evidence supports our view that people with low self-esteem will seek self-enhancement when there is little risk of disappointment. Similarly, Brown et al. (1988) have found that low-self-esteem people engage in strategies of self-enhancement that are indirect—instead of claiming superiority for things they are directly associated with, as high-self-esteem people do, they claim superiority for things that are associated with groups to which they belong. Such circuitous routes to self-enhancement may be safer for low-self-esteem people because they are less likely to lead to disconfirmation and personal humiliation (Baumeister, 1993).

In three studies, we examined the possibility that low-self-esteem people would not only protect themselves by avoiding comparisons after failure, but that they would seize a safe opportunity for self-enhancement—specifically, they would seek comparisons after success with a person who failed against them. Such comparisons promise success and hence should appeal to the low-self-esteem person’s penchant for safe strategies.

To examine these predictions, we used a comparison context and a comparison measure that are new to the social comparison literature. In the typical comparison study, subjects are provided with information about how they performed relative to normative standards or to others in a small group of strangers participating at the same time. Subjects then have the opportunity to receive more information about either norms or about an anonymous individual in their small group. In the present research, subjects received feedback not only about how they performed relative to norms, but also, as in Tesser’s (1986) paradigm, how they performed relative to a specific person. Like Tesser, we are interested in the strategies that people use after they receive feedback about how they performed relative to another individual—strategies seemingly designed to boost their self-esteem. However, the specific strategy that we are investigating is social comparison selection, that is, whether subjects want to make further comparisons with the other person.

Most previous measures of comparison selection have required subjects to select comparison targets (Wood, 1994). Indeed, most of the studies that have shown avoidance of upward comparisons have done so by showing that threatened subjects seek less superior comparison others than do other subjects, not that they avoid comparisons altogether. In the present research, we used a new measure that permitted subjects to avoid comparisons completely. In addition, whereas the information available in previous selection measures concerns one dimension—the same dimension on which subjects have already received feedback—the present measure concerns multiple dimensions. Whether subjects compare themselves, and on which dimensions they compare, is up to them. We believe that this measure is less constraining and is more realistic than most previous measures of comparison seeking (Wood, 1994).
Experiment 1

Method

Pretest Phase
Between 4 and 12 weeks before the experiment, 574 introductory psychology students received credit toward a class requirement for completing several measures in large group testing sessions. Among these measures was the 77-item “short” version of the Multidimensional Self-Esteem Inventory (MSEI; O’Brien & Epstein, 1988), which consists of a Global Self-Esteem subscale, as well as several subscales that tap into such components of self-esteem as competence and body appearance. Each subscale of the short form correlates at least .95 with its long form counterpart (O’Brien & Epstein, 1987), and the long form of the MSEI has considerable discriminant and convergent validity data to support it (O’Brien & Epstein, 1988). For example, the global subscale correlates .81 with the Rosenberg Self-Esteem Scale (Rosenberg, 1965) and .69 with the Beck Depression Inventory (Beck, 1967). Overall self-esteem can be measured by either the global subscale or a total score (56 items) across all of the items, excluding the identity and defensive self-enhancement subscales (O’Brien & Epstein, 1987). In our sample, the global subscale and the total scale correlated at .81 (p < .01). In the analyses, we used the total score, which has been used in previous published research (e.g., Campbell & Fairey, 1985; Winters & Neale, 1985).

Experimental Phase

Subjects and design
Subjects who had participated in the pretest were recruited over the telephone. Of the 104 students who eventually participated (M age = 18.7 years), 11 were not included in the analyses: Four were suspicious of the false feedback, and seven said they did not notice their own feedback (to be described). The remaining 93 subjects (45 men and 48 women) were randomly assigned to the success or failure condition, with the restriction that an equal number of high- and low-self-esteem people (the MSEI was divided at the median) and an approximately equal number of men and women were assigned to each condition. The resulting design was a 2 (high vs. low self-esteem) × 2 (success vs. failure) × 2 (sex), with 10–13 subjects per cell. For their participation, subjects received either an additional experimental credit (n = 75) or $5 (n = 18). Each cell included between one and four paid participants.

Procedure
The procedure was modeled after a study by Salovey and Rodin (1984) and ostensibly concerned “personality characteristics and career interests.” The cover story began during the initial phone call, when subjects were led to believe that they would participate with
another same-sex subject who shared their career interest, which had been assessed in the pretest materials. When subjects arrived for the experiment, the experimenter told them that they would not meet the “other subject,” because it could “bias their responses.” The experimenter pretended to speak briefly with the other subject, who was supposedly in another room, through a bogus intercom that actually was hooked up to a tape recording of the supposed other subject’s voice.

The experimenter explained that the study involved an “important aspect of personality: what sort of first impression people make,” and that one subject would be the “rater,” forming an impression of the other subject. In fact, all subjects were ostensibly “randomly assigned” to the role of rater, which set the stage for the comparison measures described below. [2]

Success–failure manipulation

Success–failure feedback was presented as a comparison of the subject’s own and the “other subject’s” potential for professional success. The experimenter introduced the other subject’s scores as being the first piece of information that subjects would use to form their impression of the other subject. The scores supposedly were from personality tests taken during the pretesting session, and they represented the “Ackerman Personality Match Index, a test that has been shown to predict how well a person’s personality is matched with his or her career choice.” The experimenter inconspicuously left the subject’s own scores within reach (saying, “we won’t need these since you are the rater”).

The feedback included a histogram made up of subscales supposedly related to the subject’s career interest. In the success condition, the subject’s own scores were considerably higher than the other subject’s on all but one subscale. In the failure condition, the subject’s scores were considerably lower than the other subject’s on all but one subscale. The experimenter left the room for 5 min to allow the subject time to compare the discrepant test scores. When she returned, the experimenter presented a “Self-Descriptive Essay” that the other subject had supposedly just completed, allegedly to help subjects form an impression of him or her. The handwritten essay was constructed so that both likable (e.g., a range of interests) and unlikable (e.g., boastfulness) aspects of the other subject were evident (cf. Salovey & Rodin, 1984).

Ratings of self and other

Subjects then rated the other subject’s ability according to “the impression they had formed thus far” on 14 dimensions, which included those used in the comparison measure to be described shortly, as well as the dimension of “potential for success in future career,” which served as a manipulation check. Subjects made their ratings on 7-point scales (no ability to excellent ability). Subjects also were asked to rate themselves, as well as the “typical Stony Brook student,” ostensibly so that these ratings could be “controlled.” We examined the ratings other than the manipulation check to see whether subjects rated themselves higher or lower than the other, and hence, appeared to expect favorable or unfavorable comparisons, respectively. (We do not report the subjects’ ratings of the “Stony Brook student,” which
yielded nothing meaningful.)

Comparison-selection measure
As our measure of comparison selection, we asked each subject to select 3 tests for the other subject to complete and 3 tests for himself or herself to complete from a list of 13 tests: social popularity, appreciation of the fine arts, intelligence, ability in school, physical attractiveness, creativity, ability to cope, overall competence, leadership ability, sensitivity to the feelings of others, political awareness, problem solving, and athletic ability. Subjects knew that they would receive both their own and the other’s test scores, so by selecting the same tests for both self and other, they could expect to make further comparisons with the other; by choosing different tests, they could avoid further comparisons. The ostensible purpose of this measure was in keeping with the cover story. Subjects were instructed to choose tests for the other subject that would assist them in forming more complete impressions of him or her. The experimenter explained that the tests were either self-administered or administered by the experimenter. Subjects also were asked whether they would mind selecting three tests for themselves to complete while they were waiting for the other subject’s scores, ostensibly because of the researchers’ needs for more normative data on the tests. Both the oral and written instructions made it clear that subjects were free to select the same or entirely different tests from those they selected for the other subject. The experimenter told subjects that they would score their own tests, so that they would think that their scores would be known only to themselves.

Jealousy, envy, and mood
Finally, subjects rated how “jealous” and “envious” they were. We added these items to the 65 items of the Profile of Mood States (POMS; McNair & Lorr, 1964). All items were rated on a 5-point scale (not at all to extremely). Subjects were told that a “person’s mood can affect how he/she forms impressions of others ... [so] we need to control for your mood.”

Probing for suspicion and debriefing
The experimenter extensively probed for suspicion before subjects realized that the experiment was over, following procedures suggested by Aronson and Carlsmith (1968). Then the experimenter revealed the deception in a gradual manner. The debriefing session also included a “process debriefing” (Ross, Lepper, & Hubbard, 1975), in which the experimenter explained that subjects sometimes believe the false feedback provided in experiments even after it has been discredited and explained the importance of realizing that it was untrue. Finally, subjects agreed not to discuss the experiment with other potential subjects.

Results
Manipulation Checks

Success in future career
Subjects’ ratings of self and other on the dimension, “potential for success in future career,” were analyzed as a repeated-measures factor of self–other in an analysis of variance (ANOVA) with the between-subjects factors of success–failure, self-esteem (divided at the median), and sex. (We will be using ANOVAs of this design throughout this Results section, except where indicated. The degrees of freedom vary somewhat because subjects occasionally did not complete a measure.) The critical effect, a Success–Failure ×Self–Other interaction, emerged, $F(1, 85) = 55.54, p < .001$, which indicates that our manipulation was successful: After failure, subjects rated the other ($M = 5.65$) higher than the self ($M = 5.04$), after success, subjects rated the other ($M = 4.29$) lower than the self ($M = 5.58$). There was one exception, however: A Success–Failure ×Self-Esteem ×Sex ×Self–Other interaction, $F(1, 85) = 5.61, p = .020$, indicated that high-self-esteem men who failed rated the self and other equally. In addition, as can be seen in Table 1, although all success subjects rated the self higher than the other, high-self-esteem women did so to a greater degree than other subjects, and low-self-esteem women did so to a lesser degree. [3]

Enlarge this Image.

Jealousy, envy, and mood
The jealous and envious items were combined ($= .75$) and examined in ANOVAs. The difference between conditions was marginally significant, $F(1, 83) = 2.93, p = .091$. Failure subjects tended to be more jealous and envious ($M = 1.15$) than were success subjects ($M = 0.60$). There was also a main effect for self-esteem, $F(1, 83) = 5.77, p = .019$, such that people with low self-esteem were higher in jealousy–envy ($M = 1.27$) than were people high in self-esteem ($M = 0.53$).

In summary, our threat manipulation appears to have been successful. Subjects’ ratings on the “success in future career” dimension differed as a function of the manipulation, and failure subjects were marginally significantly more jealous and envious than were success subjects.

It is noteworthy, however, that on the POMS, no differences emerged between success and failure conditions ($M_s = .66$ and $1.13$, respectively, $F < 1.0$). (The POMS yields a total score, which involves subtracting one positive subscale, Vigor, from several negative subscales.) We were not surprised by this finding, because the study ours was modeled after, Salovey and Rodin’s (1984), found condition differences only on a mood measure that required subjects to imagine actually meeting the supposed other subject, and not on a measure of mood that did not.

Avoiding Versus Selecting Comparisons
What do people do when given a choice of whether to compare themselves with someone who succeeded or failed against them? For each subject, we counted the number of tests chosen for the self that matched those chosen for the other and considered this to be the number of social comparisons sought. The frequencies of selections of matching tests are presented in the left side of Table 2. Also presented are the frequencies one would expect on the basis of chance, because even if subjects made their choices completely randomly, they

<table>
<thead>
<tr>
<th>Group</th>
<th>Self</th>
<th>Other</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Success</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High self-esteem</td>
<td>6.00</td>
<td>4.67</td>
<td>1.33</td>
</tr>
<tr>
<td>SD</td>
<td>0.43</td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td>Low self-esteem</td>
<td>5.10</td>
<td>3.90</td>
<td>1.20</td>
</tr>
<tr>
<td>SD</td>
<td>0.74</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High self-esteem</td>
<td>5.92</td>
<td>3.75</td>
<td>2.17</td>
</tr>
<tr>
<td>SD</td>
<td>0.67</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Low self-esteem</td>
<td>5.18</td>
<td>4.82</td>
<td>0.36</td>
</tr>
<tr>
<td>SD</td>
<td>0.75</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td><strong>Failure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High self-esteem</td>
<td>5.50</td>
<td>5.50</td>
<td>0.00</td>
</tr>
<tr>
<td>SD</td>
<td>0.80</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td>Low self-esteem</td>
<td>4.64</td>
<td>5.18</td>
<td>−0.54</td>
</tr>
<tr>
<td>SD</td>
<td>1.03</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High self-esteem</td>
<td>5.33</td>
<td>6.33</td>
<td>−1.00</td>
</tr>
<tr>
<td>SD</td>
<td>1.16</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Low self-esteem</td>
<td>4.69</td>
<td>5.54</td>
<td>−0.85</td>
</tr>
<tr>
<td>SD</td>
<td>1.11</td>
<td>1.13</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Ratings were made on 7-point scales, with higher ratings indicating judgments of higher ability.
might sometimes select the same tests for self and other. Accordingly, we calculated expected frequencies for each of the possible outcomes—0, 1, 2, and 3 matches—under a random model in which 3 tests were chosen independently for self and for other from 13 tests.

Table 2

<table>
<thead>
<tr>
<th>No. of matches</th>
<th>Expected frequency (%)</th>
<th>Observed frequency (%)</th>
<th>Raw frequency</th>
<th>Expected M no. of matches</th>
<th>Observed M no. of matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1 (n = 93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>41.96</td>
<td>10.75</td>
<td>10</td>
<td>0.692</td>
<td>1.731</td>
</tr>
<tr>
<td>1</td>
<td>47.20</td>
<td>30.11</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10.49</td>
<td>34.41</td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.35</td>
<td>24.73</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2 (n = 72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>19.37</td>
<td>15.28</td>
<td>11</td>
<td>1.250</td>
<td>2.208</td>
</tr>
<tr>
<td>1</td>
<td>44.02</td>
<td>27.78</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>29.35</td>
<td>20.83</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.77</td>
<td>12.50</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.48</td>
<td>4.17</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01</td>
<td>19.44</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. In Experiment 1, subjects selected 3 tests for self and 3 tests for other out of 13 tests; in Experiment 2, subjects selected 3 tests for self and 5 for other out of 20 tests.

The mean number of matching tests selected across the whole sample was 1.73, out of a possible 3 matches. If subjects had selected tests completely randomly, they should have selected an average of only 0.692 matches. Almost 25% chose to match on all 3 tests, whereas less than one half of one percent would be expected to. [4]

What determined the degree to which subjects sought comparisons? We entered each subject’s number of matching dimensions into the usual ANOVA design. An interaction between success–failure and self-esteem emerged, $F (1, 85) = 7.57, p = .007$, which is presented in Table 3. Contrast analyses indicated, as predicted, that low-self-esteem persons who succeeded chose more matches with the other than did low-self-esteem persons who failed. They also indicated that low-self-esteem success subjects chose more matches than did high-self-esteem success subjects. High-self-esteem persons who failed were in between; they did not differ significantly from either the low-self-esteem persons who succeeded, who matched the most, or from the other two groups, who matched the least. [5]

Next, we compared each group’s matches with chance expectations. Using $t$ tests, we compared the group’s actual mean with the expected mean of 0.692, using the error term from the ANOVA of matches reported above. All four $t$ tests exceeded this expected mean at the $p < .001$ level (the ts[85] ranged between 4.01 and 7.07). Thus, all groups appear to have
sought comparisons at an above-chance level. However, chance may not be the appropriate baseline for determining whether subjects avoided or sought comparisons. If subjects were interested in some dimensions but not in others, this would effectively narrow the number of dimensions from which they chose. If subjects then chose randomly from that restricted number, the number of expected matches would be higher than our calculation based on all 13 dimensions. A second problem with our chance baseline is that it assumes that if subjects approach the task mindlessly, they may choose matches “accidentally.” However, if subjects are truly uninvolved in the test-selection task, they may select more matches, not fewer, because they would not bother to pick different tests for self and other. They could pick the first three tests for both, for example. Both of these arguments suggest that our finding that all groups compared at rates above chance may be a product of a chance estimate that is too low. Hence, subjects actually may have used the strategy of avoiding comparisons more than we can claim on the basis of our chance estimate. It may be better, then, to focus on relative rates of comparison between groups, rather than on comparisons with chance.

Subjects’ Ratings of Self and Other on the Test-Selection Dimensions
Before subjects selected tests, they rated their own and the other subject’s abilities on all 13 dimensions used in the test-selection measure. We examined subjects’ ratings on the dimensions they selected that matched those chosen for the other to see whether subjects appeared to expect favorable or unfavorable comparisons. Contrast analyses, which compared the average self-rating to the average other-rating on the matching tests within each cell (see footnote 5), indicated that low-self-esteem success subjects rated the self higher than the other ($M_s = 5.18$ and $4.55$, respectively), $F(1, 79) = 9.38, p < .004$; as did high-self-esteem failure subjects ($M_s = 5.60$ and $5.20$, respectively), $F(1, 79) = 3.91, p = \ldots$
The two groups who sought the most comparisons, then, appeared to expect that their comparisons would be favorable to the self. Although we conducted parallel analyses for the two groups of subjects who sought the fewest matching tests, one should keep in mind that a few subjects were lost in these two groups because not all subjects chose matches, and the ratings of subjects who remained may not be representative of those who did not seek matches. The low-self-esteem failure subjects had rated their own ability higher than the other’s, but not significantly so ($M_s = 5.17$ and $4.99$, respectively, $F < 1.00$), which suggests that they expected neither unfavorable nor favorable comparisons. High-self-esteem success subjects appeared to expect favorable comparisons ($M_s$ for self and other = 5.75 and 4.52, respectively), $F (1, 79) = 32.98$, $p < .001$, which suggests that, although these subjects sought very few comparisons, when they did so, they expected them to yield favorable outcomes.

Discussion

Self-Protection and Self-Enhancement Among Low-Self-Esteem Subjects

As predicted, low-self-esteem persons made comparison selections that would benefit their self-esteem: They chose to compare more with the other when the outcome of such a comparison was likely to be favorable (when they had just succeeded relative to the other) than when the outcome was not likely to be favorable (when they had just failed). These results are consistent with past evidence that people—especially people who are low in self-esteem—avoid upward comparisons after experiencing threat (e.g., Smith & Insko, 1987). Although low-self-esteem subjects may not have truly avoided comparisons when they failed (because they compared more than would be expected by our chance baseline), they did compare less when they failed than when they succeeded, which is consistent with the view that low-self-esteem people are self-protective (Baumeister et al., 1989).

Our finding that low-self-esteem success subjects sought more comparisons than most other subjects is consistent with Wills’s (1981) prediction that people with low self-esteem are especially likely to select downward comparisons. However, there is little previous evidence for this prediction; most of the evidence cited in its favor (e.g., Friend & Gilbert, 1973) actually demonstrates avoidance of upward comparisons, rather than selection of downward comparisons (cf. Wheeler & Miyake, 1992). In addition, most of the supportive evidence ( DeVellis et al., 1990; Gibbons, 1986) has concerned low-self-esteem (or depressed) people who have undergone a recent threat or a deflation in their mood. Our results, in contrast, suggest that low-self-esteem people seek downward comparisons even after experiencing success.

We interpret this finding to mean that at times, people with low self-esteem will reach beyond their self-protective confines and actually seek self-enhancement. When they had the option of comparing themselves further with someone who was inferior to them on an important dimension, which should have been a relatively safe opportunity for self-enhancement, low-self-esteem people seized that opportunity. We suspect that these subjects were seeking to
revel in their success by making further favorable comparisons. Supporting this possibility is their ratings of self and other, which indicated that they were expecting favorable comparison outcomes.

Comparison Choices Among High-Self-Esteem Subjects
In contrast to low-self-esteem subjects, high-self-esteem subjects were not likely to seek comparisons after they succeeded. We suspect that following success, they simply were not interested in making comparisons. A previous study has indicated that people have little interest in making comparisons after they have succeeded (Pyszczynski et al., 1985). They also typically prefer upward comparisons to downward comparisons (see Wood, 1989, for references), and they especially lack interest in competing against others who are clearly inferior (e.g., Gastorf, Suls, &Lawhon, 1978). We believe that the difference between high- and low-self-esteem people following success is that high-self-esteem subjects would have little need to boost their self-esteem after succeeding, whereas low-self-esteem people must take advantage of any opportunity to do so.

When they failed, high-self-esteem subjects did not avoid comparisons—in fact, they compared somewhat more (but not significantly more) than they did when they succeeded. This result is inconsistent with past evidence that people avoid upward comparisons when threatened. However, these high-self-esteem subjects did not appear to expect upward comparisons; they rated themselves higher than the other on the matching dimensions. We defer discussion of these results until Experiment 2.

With the exception of the results for this last group, our results are consistent with theory or past research in social comparison. This is so despite our use of a comparison context that differs from past selection studies—namely, involving a specific person with whom one has already compared on an important dimension—and despite our use of a novel measure of comparison. These results argue for the validity of our test-selection measure.

Mood Effects After Social Comparisons
Although the present research focuses on the selection of future comparisons, the success–failure manipulation involved a comparison that subjects actually experienced: Failure subjects were told not only that they were unlikely to be successful in their future careers, but also that the other subject would be highly successful; success subjects were told that the other subject would not be successful. Some research has indicated that people's moods worsen after making upward comparisons and improve after making downward comparisons (e.g., Hakmiller, 1966). However, our failure subjects, who made an upward comparison, did not differ in their moods from the success subjects, who made a downward comparison. Similarly, our results did not replicate some studies that have indicated that low-self-esteem people in particular benefit in mood after making downward comparisons (e.g., Aspinwall &Taylor, 1993, Study 1; Gibbons &Boney-McCoy, 1991). However, several other studies also have failed to find this effect (Aspinwall &Taylor, 1993, Study 2; Reis, Gerrard, &Gibbons, 1993; Tesser, Millar, &Moore, 1988, Studies 1 and 2).
Recently, researchers have concluded that low-self-esteem must be coupled with a negative mood induction or a threat experience to trigger mood improvement after downward comparisons (Aspinwall & Taylor, 1993; Reis, Gerrard, & Gibbons, 1993). Because the present study did not include such a manipulation—indeed, these subjects had just been told that they succeeded—its failure to find that low-self-esteem subjects benefit from downward comparisons is not surprising.

Nonetheless, the absence of an improvement in mood after downward comparison may seem to undermine our interpretation that the low-self-esteem people were seeking comparisons to revel in their success. However, our belief is that they sought to revel, not that they have reveled already, and that their interest in comparisons was based more on their desire for positive reactions in the future than on the positive reactions they already experienced. Moreover, it is possible that these subjects did experience positive reactions, but that the mood measure failed to capture them. As mentioned earlier, Salovey and Rodin (1984) found effects using a different type of mood measure than we used. Similarly, Tesser et al. (1988) detected affective responses to social comparisons on nonverbal measures (e.g., facial expressions), but not on self-report measures. In our data, there were indications other than mood that low-self-esteem success subjects may have started to revel or that they sought to revel. Their ratings of self and other indicated that they extended their feelings of success beyond the dimension on which they received feedback to other dimensions and that they expected favorable comparisons. These results resemble those of Aspinwall and Taylor (1993, Study 2) and Reis et al. (1993), who found that low-self-esteem subjects improved in self-evaluations or expectations of future success after downward comparisons, but did not improve in mood.

Does Our Comparison Measure Truly Reflect Social Comparison?
Before going further, we should address possible questions about our measure of comparison. First, might motives other than social comparison have guided test selections? For example, subjects may have chosen certain tests simply because they were interested in the domains those tests tapped. One could argue that matches reflect subjects’ beliefs that the chosen dimensions are most meaningful or diagnostic in evaluating the self or the other, and hence, subjects may choose matching tests even if they are not seeking comparisons. Although interest in domains probably influences subjects’ choices, interest cannot account for the differences that emerged between groups in choosing matching tests. A plausible alternative explanation must account for these group differences.

It is important to note that although we can only infer that subjects wished to compare when they selected matching tests, all comparison-selection measures require similar inferences (Dakin & Arrowood, 1981; Wood, 1994). To determine whether our inference would be shared by others, we conducted an interview study.

Interview Study
We asked naive respondents why people would choose matching tests. Our aim was to find out whether people recognize the comparison potential of the measure. Sixteen respondents (6 men and 10 women), none of whom had taken more than one introductory psychology course, participated individually for either course credit or pay ($5). First, respondents received a one-half-page description of our typical procedure, which briefly described the following: how one subject succeeded and one failed at a test, the impression-formation instructions, and the request for subjects to select tests for the self to complete. Respondents then received the following two questions, one at a time:

Participants were free to select the same or different tests for self and other. They could choose tests for the other that all matched those chosen for self, no matching tests at all, or any combination of matches and nonmatches. When participants selected matching tests for the self and the other, why do you think they did so? We found that certain groups of people chose matching tests more than other groups. Can you guess why groups differed in this respect?

Respondents wrote down their answers and then described them to an interviewer, who queried in a nonleading fashion only about any points that were unclear. These interviews were audiotaped, transcribed, and coded by two independent coders. For each of the categories described below, interrater agreement was computed as the percentage of agreements on occurrences only.

Seven respondents mentioned that selecting matches may assist subjects in forming impressions of the other subject (agreement = .88). For example, one person said, “If you have the same tests to work with ... it gives you a basis to study the other person on. ’Cause you know yourself and you can compare.” However, all 16 respondents, including these seven, described a social comparison reason for choosing matching tests. Six respondents mentioned self-evaluative or self-improvement reasons for making comparisons (agreement = .67), and 15 respondents (including these 6) described subjects wanting to compete with the other subject or to choose matches for self-enhancement purposes (agreement = 1.00). The following response was typical:

To feed their ego (laughs). People like to think that life’s a big test and you always want to come out with the high score and if you have someone else to compare to, it kind of feeds your ego.

Several respondents mentioned that people who had failed might want to compare to compensate for the failure. An example was:

Well, for the low scorer I think they’d choose the same tests in sort of an attempt to redeem themselves. As sort of a second chance to prove that they maybe are equal to the other person.

Although these interviews cannot reveal the motivations of subjects who actually participated in our test-selection studies, they do indicate that the test-selection measure is easily recognized as a means of making social comparisons. Moreover, it is easily recognized as offering social comparisons that may be self-enhancing.
Experiment 2
Our second experiment had four aims. The primary aim was to replicate, using different operationalizations, the interaction between condition and self-esteem in comparison selections. We used a different manipulation of threat, a different measure of self-esteem, and we made other changes in procedure. Our second aim was to clarify the relation between comparison frequencies and baseline. In addition to using chance as a baseline, we included a no-feedback condition. Doing so enabled us to ask such questions as, Do low-self-esteem subjects seek more comparisons when they succeed than they normally would, or is it that they avoid comparisons more than usual when they have failed? The third aim was to examine an alternative explanation for subjects’ comparison selections, namely, that they sought comparisons because they were surprised by the feedback. Low-self-esteem people who had just been told that they would be highly successful in their future careers and high-self-esteem people who had just been told that they would be unsuccessful may have been confused or skeptical about their feedback and may have sought further comparisons to help them understand it. The finding that these subjects were expecting favorable comparisons argues against this interpretation in the case of low-self-esteem subjects; if they disbelieved the social accuracy results, they should have expected unfavorable comparisons. However, the favorable expectations of the high-self-esteem failure subjects do not argue against this interpretation. Hence, we addressed this alternative explanation in the second experiment. We included measures of subjects’ puzzlement and skepticism as well as their beliefs about the credibility of the feedback. Our fourth aim was to replicate the self- and other-rating results, in the hopes of finding additional support for our interpretation of subjects’ motivations for comparison. In Experiment 1, subjects rated the self and other before they knew they would have an opportunity to compare the other with the self. This time, we asked subjects about their expectations for performance more directly, after they selected tests.

Method
Pretest Phase
Between 4 and 10 weeks before the experiment, 375 introductory psychology students completed several measures in large group-testing sessions, for which they received one credit toward a class requirement. These measures included the Rosenberg Self-Esteem Inventory, which is a widely used 10-item measure of global evaluations of the self (Rosenberg, 1965). Responses to items such as, “On the whole, I am satisfied with myself” are made on 5-point scales (from strongly disagree to strongly agree). A total score across all of the items was our index of self-esteem.

Experimental Phase
Subjects and design
Subjects who had participated in the pretest were recruited over the telephone. The 72 subjects (35 women and 37 men; \( M \) age = 20.6 years) who eventually participated in the experiment were randomly assigned to the success, no-feedback, or failure condition, with the restriction that an equal number of high- and low-self-esteem subjects (split at the median; \( M \) high-self-esteem score = 36.81; \( M \) low-self-esteem score = 25.78) and a nearly equal number of men and women were assigned to each condition. The resulting design was a 2 (high vs. low self-esteem) \( \times \) 3 (success vs. no-feedback vs. failure condition) \( \times \) 2 (sex) factorial, with six subjects per cell. For their participation, subjects received an additional experimental credit.

Procedure
The experimental procedure was modeled after one used by McFarland and Ross (1982). Subjects were informed during the initial phone conversation that the purpose of the experiment was to investigate the relationship between “social accuracy” and the ability to form accurate first impressions of other people. When subjects arrived for the experiment, they were introduced to another subject of the same sex. Subjects were told that both would be administered a test of social accuracy, that one of them would then be taken to a different room, and that one subject would form an impression of the other. (Once they were separated, all subjects were ostensibly “randomly assigned” to this role by means of the same bogus procedure as was used in Experiment 1.)

Success–failure manipulation
A second experimenter observed each subject through a partially covered one-way mirror as he or she took the social accuracy test. This experimenter created an answer key for each subject such that in the success condition, 12 or 13 of the supposedly correct 15 answers corresponded with the subjects’ answers. In the failure condition, subjects’ answers corresponded with only 3 or 4 of the 15 answers. The second experimenter unobtrusively delivered these answer keys to the first experimenter, who then asked subjects to mark each other’s test. The answer keys indicated which scores constituted an above average (12–15), average (7–11), or below average (0–6) grade. In the no-feedback condition, subjects’ tests were simply collected without being scored.

Next, the experimenter administered a questionnaire that McFarland and Ross (1982) used to manipulate attributions. Because McFarland and Ross (1982) found that when subjects made internal attributions for a success or failure their moods were intensified, we included this questionnaire to enhance the success–failure manipulation. Subjects in the success condition were asked to choose from a list of abilities those that they believed were related to their own social accuracy. Subjects in the failure condition were asked to choose abilities that, if they were lacking, would be related to their low social accuracy ability.
As manipulation checks, subjects were asked two questions about their test result: “How well did you perform on the social accuracy task?” and “How satisfied were you with your performance on the social accuracy task?” Subjects responded on 8-point scales (not at all well [satisfied ] to extremely well [satisfied ]).

Comparison-selection measure
Once the subjects were separated, the experimenter provided the same rationale for the test-selection measure as was used in Experiment 1. With the hope of observing a greater range on the comparison measure this time, we increased the number of possible choices and matches. Subjects were asked to select 5 tests for the other subject and 5 tests for themselves from a list of 20 tests: adaptability in different social settings, social sensitivity, ability to empathize with others, interpersonal skills, insight into others’ problems, accuracy of first impressions, social popularity, compatibility with different types of people, ability to maintain close relationships, social intelligence (defined as knowledge about how to handle various social situations), appreciation of the fine arts, intelligence, ability in school, creativity, overall competence, leadership ability, political awareness, problem solving ability, athletic ability, and emotional adjustment.

Expectations measure
After completing the test-selection measure, subjects were once again given the list of 20 tests and were asked to indicate on a 7-point scale (not at all well to extremely well ) how they expected to perform and how they expected the other subject to perform on each test, regardless of whether they had selected the test.

Mood measure
Subjects completed the 20 adjectives of the Positive and Negative Affective Schedule (PANAS; Watson, Clark, &Tellegen, 1988) to describe how they felt “at the moment,” supposedly so that their moods could be “controlled” in our analyses. According to Watson et al. (1988), the PANAS internal consistencies range from .86 to .90 for Positive Affect and from .84 to .87 for Negative Affect. To the PANAS we added two items—“puzzled” and “skeptical”—to address the possibility that subjects were taken aback by their results. We also added items that we thought may better capture subjects’ experiences after success or failure: humiliated, pleased, stupid, disappointed, confident, worthless, smart, happy, competent, angry, and inadequate. All items were rated on 5-point scales (very slightly or not at all to extremely ).

Beliefs about the validity of the test
Subjects were then asked four questions about their social accuracy test result: “To what extent do you think the score you received from the social accuracy test is a result of your social accuracy ability and the effort you displayed during the test?” “How much could a stranger learn about your social accuracy from seeing the score you received?” “To what extent
extent do you think the score you received on the social accuracy test is not a result of your ability, but is a result of the characteristics of the test itself (e.g., ease or difficulty, or ambiguity)?” and “How much do you agree with the result?” The answers were based on a 7-point rating scale (very little to very much). These questions were adapted from a study by Swann, Griffin, Predmore, and Gaines (1987).

Finally, the experimenter debriefed the subjects using the same procedures as in Experiment 1. No subject expressed any suspicion about the feedback.

Results

Manipulation Checks

Subjects’ responses to the questions concerning performance and satisfaction with performance on the social accuracy task were averaged (the composite’s internal consistency was .98) and entered into an ANOVA with condition (success, failure, or no-feedback), self-esteem (low or high), and sex as between-subjects factors. (We will be using ANOVAs of this design throughout the Results section, except where indicated.) A main effect for condition, $F(2, 60) = 478.42, p <.001$, indicated that success subjects thought they had done better and were more satisfied ($M = 7.33$) than were no-feedback ($M = 4.15$) and failure subjects ($M = 1.40$). Also, as one might expect, subjects with high self-esteem ($M = 4.65$) thought they had done better than did subjects with low self-esteem ($M = 3.93$), $F(1, 60) = 21.02, p <.001$. These effects were qualified by a significant interaction between condition and self-esteem, $F(2, 60) = 16.08, p <.001$. As can be seen in Table 4, contrast analyses indicated that low-self-esteem and high-self-esteem subjects differed significantly only when they were not given feedback, when high-self-esteem subjects were more pleased than were low-self-esteem subjects.

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manipulation Check: Mean Ratings of Performance and Satisfaction With the Social Accuracy Test Result in Experiment 2</strong></td>
</tr>
<tr>
<td>Self-esteem group</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Low $M$</td>
</tr>
<tr>
<td>$SD$</td>
</tr>
<tr>
<td>$n$</td>
</tr>
<tr>
<td>High $M$</td>
</tr>
<tr>
<td>$SD$</td>
</tr>
<tr>
<td>$n$</td>
</tr>
</tbody>
</table>

Note. Means with different subscripts differed significantly at $p < .05$ or less in contrast analyses.
Next we examined subjects’ reports of mood. We thought that mood effects may emerge in this study when they had not in Experiment 1 because the success–failure manipulation was modeled after that of McFarland and Ross (1982), who did find mood differences between conditions. The positive items on the PANAS, the negative items on the PANAS, and the positive and negative mood items that we added to better capture feelings of success and failure were analyzed in a multivariate analysis of variance (MANOVA) with condition, self-esteem, and sex as independent variables. The condition effect was significant, Wilks’s $F (8, 114) = 4.64$, $p < .001$. Univariate tests yielded significant condition effects on the positive PANAS items, $F (2, 60) = 3.35$, $p = .042$, and on the additional negative items, $F (2, 60) = 13.47$, $p < .001$. Success subjects reported more positive mood ($M = 2.98$) than failure or no-feedback subjects (both $M$s = 2.58), and failure subjects reported feeling the most humiliated, worthless, and so forth ($M = 1.71$; $M$s for no-feedback and success subjects = 1.19 and 1.23, respectively). Also, as one might expect, the overall MANOVA yielded a main effect for self-esteem, Wilks’s $F (4, 57) = 4.80$, $p = .002$, such that high-self-esteem people were more positive in their mood than were low-self-esteem people.

In summary, both the manipulation check and the mood results indicate that we successfully manipulated success and failure.

Avoiding Versus Seeking Comparisons
Subjects chose 5 tests for themselves and for the other to take from a selection of 20 tests. To facilitate comparison with Experiment 1, Table 2 presents the expected frequencies of matches under a random model. One would expect an average of 1.25 matches per subject if subjects had selected tests completely randomly; the actual average was 2.21. Over 19% chose matches for all 5 tests, whereas far fewer than 1% would be expected to do so. However, the overall rate of comparison appears to be somewhat lower than it was in Experiment 1. A possible reason is that in Experiment 1, subjects were reassured repeatedly of their anonymity, whereas in Experiment 2, subjects probably assumed that the experimenter would observe their performances, which may have inhibited them from choosing direct comparisons. (When we conducted Experiment 1, we wanted subjects’ choices to be free of self-presentational concerns. We have since been persuaded of the extreme difficulty of disentangling self-presentational and self-enhancement concerns [Tetlock & Manstead, 1985] and have dropped our emphasis on anonymity.)

Did self-esteem and success–failure feedback again determine whether subjects sought or avoided comparisons? An ANOVA replicated the interaction we obtained in Experiment 1: As Table 5 indicates, people with low self-esteem matched most when they had succeeded, and people with high self-esteem matched most when they had failed, $F (2, 60) = 60.88$, $p < .001$. [8] The inclusion of the no-feedback condition in this study enables us to examine whether subjects are seeking or avoiding comparisons in a sense other than relative to chance, namely, relative to a “baseline” condition. Contrast analyses revealed that low-self-
esteem subjects did seek more matches when they had succeeded than when they had failed. They did not avoid matches after failure, however, in that they did not differ in that condition from the no-feedback subjects. High-self-esteem subjects sought significantly more matches when they failed and avoided matches when they succeeded, relative to when they received no feedback.

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Number of Tests Chosen for Self That Matched Those Chosen for Other as a Function of Self-Esteem and Feedback Condition in Experiment 2</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td><strong>Self-esteem Group</strong></td>
</tr>
<tr>
<td><strong>Low</strong></td>
</tr>
<tr>
<td><strong>M</strong></td>
</tr>
<tr>
<td><strong>SD</strong></td>
</tr>
<tr>
<td><strong>High</strong></td>
</tr>
<tr>
<td><strong>M</strong></td>
</tr>
<tr>
<td><strong>SD</strong></td>
</tr>
<tr>
<td><strong>Note.</strong> Means with different subscripts differ significantly at <em>p</em> &lt; .05 or less in contrast analyses.</td>
</tr>
</tbody>
</table>

A Different Way of Measuring Degree of Comparison-Seeking

Next, we considered the possibility that subjects could seek comparisons by selecting tests that did not match those chosen for the other. If subjects believed, for instance, that Dimension A was correlated with Dimension B, they could select A for the self and B for the other and still expect to receive comparative information. We examined this possibility by asking 20 introductory psychology subjects to rate the degree of overlap between the dimensions used in the test-selection measure in Experiment 2. Subjects were told that “two abilities can be said to overlap if a person with one of the abilities would tend also to have the other ability.” They rated the overlap between each dimension and every other dimension, as
well as each dimension with itself, on 9-point scales (from 1 = no overlap to 9 = total overlap). Results indicated that subjects believed that matches were more overlapping than were nonmatches, \( F(1, 18) = 382.75, p < .001 \).

Next, we used these overlap ratings to assign each subject in Experiment 2 a new comparison score based on the degree of overlap between dimensions, rather than on the number of matches. These scores could be high for someone with few or no exact matches, provided that the dimensions selected for self and for other were rated as highly overlapping. Actually, two comparison scores were calculated, each based on a different algorithm. For the “best 5” score, we examined all possible pairings of the 5 dimensions chosen for self and 5 dimensions chosen for other, selected the pair of dimension-chosen-for-self and dimension-chosen-for-other that yielded the highest overlap rating, selected from the 8 remaining dimensions the pair that yielded the highest overlap rating, then selected from the 6 remaining dimensions the pair that yielded the highest overlap rating, and so forth. The sum of the overlap ratings for these five optimal pairings was the “best 5” score. The “all 25” score simply averaged the overlap ratings for all 25 possible pairings of the 5 selections for self and for other.

These new comparison scores were then examined in the same ANOVA design that we had used for matches. The Condition × Self-Esteem interaction emerged again, such that the low-self-esteem success subjects and high-self-esteem failure subjects sought the most highly overlapping dimensions, for “best 5,” \( F(2, 60) = 51.92, p < .001 \), for “all 25,” \( F(2, 60) = 11.87, p < .001 \). The fact that these new comparison scores yielded the same results as our method of simply counting matching dimensions indicates that our method is not failing to capture important comparisons. We conclude that simply counting matches is appropriate.

Were Subjects Who Sought Comparisons Puzzled or Skeptical?

Did high-self-esteem subjects who failed and low-self-esteem subjects who succeeded choose to compare with the other person because they were confused? First, we averaged the “puzzled” and “skeptical” items that we had included in the mood questionnaire ( = .82) and we correlated this composite with the number of matches subjects made. They were not correlated significantly (\( r = .13 \)). The failure to correlate does not appear to be due to any inadequacy on the part of the puzzled–skeptical composite, because we did find effects for it when we analyzed it in the usual ANOVA. There were main effects for condition, \( F(2, 60) = 43.85, p < .001 \), and self-esteem, \( F(1, 60) = 5.33, p = .024 \). Subjects who failed were more skeptical and puzzled (\( M = 3.58 \)) than subjects who succeeded (\( M = 1.71 \)) or who received no feedback (\( M = 1.88 \)), and subjects with low self-esteem were more skeptical and puzzled (\( M = 2.60 \)) than were subjects with high self-esteem (\( M = 2.18 \)). However, there was no interaction (\( F < 1.0 \)), which would have been obtained if the high-self-esteem failure and low-self-esteem success subjects were especially puzzled by their results.

Next, subjects’ responses were averaged across the four questions concerning the perceived validity of the social accuracy test ( = .97). This composite was not correlated with the number of matches that subjects made (\( r = -.06 \)). We then analyzed this perceived
validity composite in an ANOVA, where again, the puzzled–skeptical hypothesis would predict an interaction, such that low-self-esteem success and high-self-esteem failure subjects should be least likely to believe the test was valid. An interaction was obtained, $F(2, 60) = 7.00, p = .002$, but it was of an entirely different form. As can be seen in Table 6, low-self-esteem success subjects did not appear to disbelieve the test results. Although high-self-esteem subjects who failed did doubt the validity of the test, so did low-self-esteem subjects who failed. Where the low- and high-self-esteem subjects differed was in the no-feedback condition, in which high-self-esteem subjects were more inclined to believe the results of the test (although they never saw those results). 

Table 6
Mean Ratings of Beliefs About the Validity of the Social Accuracy Test in Experiment 2

<table>
<thead>
<tr>
<th>Self-esteem group</th>
<th>Success</th>
<th>Failure</th>
<th>No-feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>5.88</td>
<td>2.25</td>
<td>3.81</td>
</tr>
<tr>
<td>$M$</td>
<td>5.88</td>
<td>2.25</td>
<td>3.81</td>
</tr>
<tr>
<td>SD</td>
<td>0.48</td>
<td>0.81</td>
<td>0.50</td>
</tr>
<tr>
<td>n</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>High</td>
<td>6.23</td>
<td>1.96</td>
<td>4.79</td>
</tr>
<tr>
<td>$M$</td>
<td>6.23</td>
<td>1.96</td>
<td>4.79</td>
</tr>
<tr>
<td>SD</td>
<td>0.69</td>
<td>0.66</td>
<td>0.69</td>
</tr>
<tr>
<td>n</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Higher numbers indicate stronger confidence in the validity of the test. Means with different subscripts differ significantly at $p < .05$ or less in contrast analyses.

In sum, these results fail to support the hypothesis that low-self-esteem success subjects and high-self-esteem failure subjects sought comparisons because they were puzzled or doubtful about the feedback they had received.

Exploring Motivations Underlying Comparisons
Why, then, did subjects seek comparisons? For a preliminary look at this question, we examined subjects’ expectations for comparisons as well as their moods.

Expectations for self and other
After subjects selected tests for the self and the other, they rated how they would expect themselves and the other to perform on all 20 dimensions used in the test-selection measure. Table 7 presents their ratings on the specific dimensions they selected for comparison. We examined these ratings to address the question, Did low-self-esteem success and high-self-esteem failure subjects expect favorable comparison outcomes? Contrast analyses, which
compared expectations for self and other within each cell, indicated that high-self-esteem failure subjects expected to perform better than the other, $F (1, 55) = 7.13, p < .010$. Although the low-self-esteem success subjects rated themselves higher than for the other, the difference of .41 was not significant, $F (1, 55) = 1.22, ns$, which suggests that they were expecting their comparisons to be neither favorable nor unfavorable. However, these subjects’ expectations for matching dimensions were favorable relative to those of the other low-self-esteem groups; a contrast comparing their self–other ratings with those of the low-self-esteem failure and no-feedback conditions was significant, $F (1, 55) = 5.68, p = .019$.

Table 7
Mean Expectation Ratings of Self and Other on Matching Tests as a Function of Self-Esteem and Feedback Condition in Experiment 2

<table>
<thead>
<tr>
<th>Self-esteem group</th>
<th>Success</th>
<th>Failure</th>
<th>No-feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>4.74</td>
<td>3.25</td>
<td>4.55</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.99</td>
<td>1.17</td>
<td>1.38</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>4.33</td>
<td>4.89</td>
<td>4.65</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.69</td>
<td>0.80</td>
<td>1.06</td>
</tr>
<tr>
<td>Difference</td>
<td>0.41</td>
<td>-1.64</td>
<td>-0.10</td>
</tr>
<tr>
<td>$n$</td>
<td>12</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>5.44</td>
<td>5.54</td>
<td>5.70</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.68</td>
<td>0.38</td>
<td>0.51</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>3.67</td>
<td>4.55</td>
<td>4.39</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.00</td>
<td>0.77</td>
<td>1.14</td>
</tr>
<tr>
<td>Difference</td>
<td>1.77</td>
<td>0.99</td>
<td>1.31</td>
</tr>
<tr>
<td>$n$</td>
<td>9</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Note. Ratings were made on 7-point scales, with higher ratings indicating higher expectations.

One may question, however, whether this analysis is appropriate. As mentioned with regard to Experiment 1, subjects who did not seek comparisons are lost in this analysis, and subjects who remain may not be representative. The discrepancies in the cell sizes are even more pronounced here. Hence, it may be useful to consider subjects’ ratings across all 20 dimensions, rather than matching dimensions, because no subjects are lost. The ratings across all 20 dimensions reveal what impact the success–failure manipulation had on
subjects’ general beliefs about how they measured up to the other, and hence, what they expected from potential comparisons. Across all 20 dimensions, low-self-esteem subjects typically rated the other higher than the self ($M_{other}$-ratings for failure and no-feedback groups = 5.23 and 4.79, respectively; $M_{self}$-ratings $= 3.98$ and 4.44, respectively). However, this was reversed when they succeeded ($M$s for other and self $= 4.39$ and 4.68, respectively). In fact, these expectations of the low-self-esteem success subjects were significantly more favorable than those of the low-self-esteem no-feedback group, $F(1, 66) = 3.85$, $p = .051$. These figures shed light on low-self-esteem success subjects’ ratings on matching dimensions; a positive self-minus-other difference, no matter how small, may represent a triumph, because low-self-esteem people in this sample typically expected the other to perform better than the self.

Relation between expectations and selection of matches
Subjects’ expectations for self and other on the matching dimensions provide clues about subjects’ motivations for comparison, because they suggest what subjects’ expected from the specific dimensions they selected for comparison. For further clues about motivation, we examined the correlations between the average self-minus-other expectation across all 20 dimensions and the number of matches selected. Correlations between these overall ratings and matches speak to the possible role of expectations for potential comparisons in leading people to seek actual comparisons. For example, did the prospect of making favorable comparisons entice people to select comparisons? These correlations point to a striking difference between low- and high-self-esteem people. In the case of low-self-esteem people, the more favorably they saw themselves relative to the other, the more likely they were to seek matches ($r = .38$, $p < .05$). For high-self-esteem people, there was a nonsignificant tendency in the opposite direction; the more favorably they viewed themselves relative to the other, the less likely they were to seek matches ($r = -.23$, $ns$). These correlations differ significantly, $Zr = 2.47$, $p = .014$. These differences are even more pronounced when subjects who received no feedback are excluded from the analysis; for low-self-esteem subjects, $r = .48$ ($p < .02$), and for high-self-esteem subjects, $r = .32$, $ns$ ($Zr = 2.77$, $p < .006$).

Although high-self-esteem subjects tended to seek more comparisons the less favorable their expectations, it is important to note that high-self-esteem subjects had higher expectations for themselves ($M = 5.60$) than for the other ($M = 4.25$). This suggests that the (nonsignificant) negative correlation does not indicate that high-self-esteem subjects sought comparisons that they expected to be unfavorable; they apparently expected that if they compared themselves with the other, they would perform better than the other.

We did not obtain the same correlations in Experiment 1; they were all near zero. We believe that there are three reasons for this: First, the number of matches variable had a smaller range in Experiment 1 than in Experiment 2 ($SD$s = 0.957 and 1.711, respectively), probably in part because subjects could select up to five matches in Experiment 2. Second, the internal consistency of the self-minus-other differences variable was much higher in
Experiment 2 ( = .94) than it was in Experiment 1 ( = .59), probably in part because it included 20 dimensions rather than 13. Third, the ratings varied more widely in Experiment 2; the standard deviations of the self-minus-other difference scores was 1.28 in Experiment 2, but only 0.539 in Experiment 1. The larger variance in the ratings in Experiment 2 was probably due to several factors: (a) the success–failure manipulation, which appears to have been somewhat more powerful in Experiment 2; (b) the fact that subjects in Experiment 2 were told that half of the dimensions were related to the success–failure feedback dimension; hence, their ratings should have been much more affected by that feedback; and (c) subjects in Experiment 2 may have been more careful to rate themselves and the other in accordance with the feedback, because they were not, as subjects in Experiment 1 were, assured of confidentiality.

Regardless of the reasons behind the more restricted ranges for both variables involved in the correlation and behind the lower reliability for one of the variables, these factors certainly attenuated any correlation between ratings and number of matches that may have been present in Experiment 1.

Relation between mood and selection of matches
For further clues about subjects' motivations for seeking comparisons, we examined the correlations between mood and the number of matches selected. Again, a difference emerged between people with low and with high self-esteem. For people with low self-esteem, positive mood (as indexed by a composite that included all positive mood items) was correlated with comparison seeking ($r = .36, p < .05$), but negative mood (as indexed by a composite that included all negative mood items) was not ($r = -.14, ns$). For people with high self-esteem, negative mood was correlated with comparison seeking ($r = .50, p < .01$), but positive mood was not ($r = -.08, ns$). In other words, low-self-esteem subjects sought comparisons the better they felt; high-self-esteem subjects sought comparisons the worse they felt.

Discussion

Avoiding Versus Seeking Comparisons
In Experiment 2 we replicated the interaction between success–failure and self-esteem that we had obtained in Experiment 1. In fact, our results were stronger than they were in Experiment 1, in that the high-self-esteem failure subjects differed significantly from most other groups, whereas they had not in Experiment 1. In part, this difference may be due to a stronger manipulation of success and failure in Experiment 2. The mood effects in this study suggest that the manipulation may have been more potent (although the mood measure differed from that used in Experiment 1, making comparisons hazardous). It is difficult to know what accounts for the apparently stronger manipulation, but one difference between studies that we believe to be important concerns the attributional questionnaire used in Experiment 2. McFarland and Ross (1982) found that subjects who completed this
questionnaire, which was designed to encourage internal attributions, were more affected by a success–failure manipulation than were subjects who were exposed to a questionnaire designed to encourage external attributions. It is also likely that the social comparative aspect of the feedback was more salient in Experiment 2; whereas in Experiment 1 subjects received information on paper about someone they had never seen, in Experiment 2 they sat right next to their rival, and their relative performance was either painfully or joyfully obvious. This powerful social comparison context may have been responsible for the finding that, in contrast to most previous studies (Gibbons & Gerrard, 1991), both low- and high-self-esteem subjects showed changes in mood in response to both upward and downward comparisons. Differences between Experiments 1 and 2 other than the success–failure manipulation also may have contributed to the stronger results in Experiment 2. The measures of self-esteem differed. Also, the overall rate of comparison was higher in Experiment 1, perhaps making it harder to discern group differences. In addition, the test-selection measure allowed a wider range of comparison making in Experiment 2 (0–5) than in Experiment 1 (0–3), and the resulting larger variance may have allowed group differences to emerge more clearly. In any case, Experiment 2 found that low-self-esteem success subjects and high-self-esteem failure subjects were most interested in comparisons. We also obtained evidence that they did not make more comparisons because they were puzzled or skeptical about their test results or because they doubted the validity of the feedback.

Motivations Underlying Comparison Choices

High-self-esteem failure
What, then, were the motives behind comparison selections? Consider first the high-self-esteem subjects who failed, who sought comparisons with the very person who was just shown to be superior to them. Is this evidence of self-destructiveness on the part of high-self-esteem people? We think not. Rather, we suspect that they thought that they could compensate for the failure. Two studies have suggested that people with high self-esteem are more likely than people with low self-esteem to seek what Baumeister and Jones (1978) have called “compensatory self-enhancement” (Baumeister, 1982; Josephs, Markus, & Tafarodi, 1992). Compensation strategies can be direct, such as overturning one’s failure in a second attempt (McFarlin, Baumeister, & Blascovich, 1984) or behaving in ways that disconfirm unfavorable information about oneself (Steele & Liu, 1981). One also may compensate indirectly by emphasizing one’s talents or values in other areas (e.g., Baumeister, 1982; Baumeister & Tice, 1985; Steele & Liu, 1981).

Although trying to compensate by making further comparisons with someone who is superior to oneself in an important domain may seem risky, high-self-esteem people seem to be willing to take such risks. Even after they have failed, they still expect that they will perform successfully on a second try (McFarlin & Blascovich, 1981). They exhibit more effort and perform better on subsequent tasks after failure than do low-self-esteem people (McFarlin et al., 1984; Shrauger & Sorman, 1977). They sometimes even have higher expectations after
failure than after success (see Brockner, 1983, for references). Thus, unlike people with low self-esteem, who shrink from further evaluation after failure, people with high self-esteem seem to be energized by failure (Raynor & McFarlin, 1986).

Recall that the high-self-esteem failure subjects were the only subjects whose comparison choices contradicted those found in past social comparison studies. We believe that this discrepancy is due to the comparison context used in the present research, which is the only one that allows this compensatory tendency of high-self-esteem people to express itself. The present subjects had the opportunity to make comparisons with the very person who had succeeded against them, which offered the potential of reversing the earlier failure or of besting the person on other dimensions. Supporting our suspicion that high-self-esteem subjects were trying to compensate for the failure is the finding that they expected their comparisons to reveal that they were superior.

Low-self-esteem success
Low-self-esteem subjects who succeeded also expected that their comparisons would be favorable—or at least, they expected more favorable comparisons than low-self-esteem subjects typically did. Why were their expectations not absolutely favorable, as they were in Experiment 1? One possibility is that the subjects in Experiment 2 were lower in self-esteem. Recall that even the low-self-esteem subjects who failed in Experiment 1 rated themselves (nonsignificantly) higher than the other. Perhaps more important, in Experiment 1, subjects believed that their responses on the rating measure, as well as their scores on the tests they selected for the comparison measure, would be known only by themselves, whereas in Experiment 2, subjects had no such assurance of privacy. Past research has indicated that people are more modest when they believe that their expectations will be put to the test than when they do not (for references, see Schlenker et al., 1990).

In any case, these low-self-esteem success subjects’ expectations were unusually favorable for people with low self-esteem in this sample. We suspect that they chose to compare further with someone who had just failed against them in hopes of boosting their self-esteem. Why would low-self-esteem people seek to boost their self-esteem when they had just succeeded? Perhaps because, as Wills (1981) proposed, their self-esteem is in a sense chronically threatened. Our results suggest that people with low self-esteem may try to self-enhance if given a safe chance to do so—“safe” in the sense that it promises success and is unlikely to expose their deficiencies.

Exploration of motivations
The correlations that emerged in Experiment 2 offer preliminary support for our view that the motivations underlying the selection of matches differ for low- and high-self-esteem people. The positive correlation between expectations and matches indicated that for people with low self-esteem, the more they expected to perform better than the other, the more they sought comparisons. This correlation suggests that they were seeking self-enhancement. For people with high self-esteem, the less favorable their expectation, the more they sought
comparisons. However, this negative correlation does not indicate that they were seeking the opposite of self-enhancement. Recall that the high-self-esteem subjects were expecting their comparisons to yield favorable outcomes, which suggests they were expecting self-enhancing comparisons. It is also important to note that the negative correlation was accounted for largely in the high-self-esteem success group \( r = -0.63 \); the other high-self-esteem subjects had correlations close to zero \( r_s \) for no-feedback and failure groups were \(-0.02\) and \(-0.05\), respectively. We interpret the negative correlation to mean that, given the expectation of a favorable comparison, high-self-esteem people are less likely to seek comparisons, the more favorable their expectation. This is true, we believe, because especially favorable comparisons are not unusual or interesting for people with high self-esteem. It seems that expectations alone do not determine comparison-seeking, at least for high-self-esteem people. Otherwise, the high-self-esteem success subjects, who expected the most favorable comparisons, would have been most likely to seek comparisons. We believe that, in addition to a favorable expectation for comparison, high-self-esteem people must have the need or desire for favorable comparisons. They should have this desire when they have just failed relative to the other, and when they believe that they can reverse that failure by outperforming the other. Our high-self-esteem failure subjects expected to perform better than the other, which suggests that they were hoping to turn an upward comparison into a downward one.

Like the correlations involving expectations, the correlations between mood and matches suggest that the motivations underlying comparisons differed between low- and high-self-esteem people. For low-self-esteem subjects, the better they felt, the more they sought comparisons. This result is consistent with the idea that low-self-esteem subjects were seeking to revel in their success. In contrast, positive mood was not associated with comparison-seeking among high-self-esteem subjects. This result suggests that they were not motivated to maximize their pleasure, which is consistent with our notion that the expectation of a favorable comparison is not sufficient to induce high-self-esteem people to seek comparisons. High-self-esteem subjects may have been more motivated by negative moods; the worse they felt, the more they sought comparisons. This correlation, coupled with the finding that high-self-esteem failure subjects expected favorable comparisons, suggests that these subjects were trying to overcome their negative feelings by reversing the failure. Whereas low-self-esteem success subjects seemed to be trying to maximize their pleasure, high-self-esteem failure subjects seemed to be trying to right a wrong.

We admit that our interpretation of these correlational results remains speculative at this point and that we have not ruled out all possible alternative interpretations. An alternative explanation suggested by a reviewer of this article, for example, was that information-seeking motives prompted subjects’ comparison-seeking. The unusual experiences of success for low-self-esteem subjects and of failure for high-self-esteem subjects may have led them to want to learn more about themselves and the other. It is possible that this explanation is not ruled out by the items used to address the puzzled–skeptical hypothesis, because one could believe the test results, yet still want to understand those results better. However, we believe
that this information-seeking idea does not encompass all of our findings. First, responses in
the interview study suggested that many people would experience self-esteem-related
motives in this situation. Second, high-self-esteem failure subjects expected that they would
perform better than the other in their future comparisons. If they sought comparisons with a
superior other solely because such comparisons would be informative, that implies that they
should believe that the superiority would be stable. Third, low-self-esteem success subjects
were more likely to compare, the more pleased they were with their results. Fourth, the
correlations involving expectations, moods, and comparison-seeking differed between low-
and high-self-esteem people; if they were guided by the same motive, namely, information
seeking, one would not expect differences in the correlates of comparison-seeking. We are
not claiming that each of these findings necessarily contradicts the information-seeking
explanation; rather, we are saying that the information-seeking explanation does not provide
an account for these findings, whereas our explanation does.

Of course, these arguments do not ensure that the explanation that we favor is correct.
Further studies providing more direct evidence for these motivations are needed. Experiment
3 illustrates how such studies may be conducted.

Self-Protection Among Low-Self-Esteem Subjects?
Although low-self-esteem subjects were not especially interested in comparisons after failure,
we did not find that they truly avoided comparisons; they did not compare less than the no-
feedback group. These results are inconsistent with recent emphases on low-self-esteem
people’s self-protectiveness (Baumeister et al., 1989). It is possible that avoidance of
comparisons would be found in a sample of people who are lower in self-esteem than our
subjects; researchers have argued that university students tend to be high in self-esteem (Swann,
1987). It is also possible that our low-self-esteem subjects in the no-feedback
condition themselves actually avoided comparisons, which would have made it difficult to
demonstrate a lower degree of comparison in the failure subjects. A no-feedback condition
may not be a neutral experience for people with low self-esteem; these subjects apparently
assumed that they would perform worse than the other person. In Experiment 3, we
examined the hypothesis that low-self-esteem people are reluctant to make comparisons
even after they succeed, unless they are confident of their relative superiority to the other.

Experiment 3
To test the hypothesis that low-self-esteem subjects will reach beyond their usual self-
protectiveness and actually seek self-enhancement if the opportunity is safe, we presented to
low-self-esteem subjects who succeeded the opportunity for safe versus relatively risky
comparisons. In Experiments 1 and 2, low-self-esteem subjects who succeeded had
available safe comparisons, because they already knew that they were superior to the
comparison other. In Experiment 3, we manipulated how safe the social comparisons would
be by informing only half of the subjects of the other person’s failure. For these subjects, our
predictions were the same as in Experiments 1 and 2: Armed with the knowledge that the
other was inferior to them, these low-self-esteem subjects should seek further comparisons to take advantage of an almost certain opportunity to self-enhance. In contrast, for subjects not informed of the other’s failure, comparisons would be relatively risky, and hence they should not seek them. This is a strong test of the hypothesis, because subjects could be so emboldened by their success that they seek comparisons. However, without knowledge of the other’s social accuracy ability, these subjects could not be guaranteed a favorable comparison outcome; the other may well have succeeded too and may even be superior. To avoid possible harm to their fragile self-esteem, then, they should not seek comparisons as much as subjects who knew of the other’s inferiority.

Method

Pretest Phase
Several weeks before the experiment, 197 introductory psychology students completed various measures in large group-testing sessions, including the Rosenberg Self-Esteem Scale.

Subjects and Design
Subjects whose self-esteem score fell below the median were recruited over the telephone. Twenty-four subjects (13 women and 11 men; $M_{\text{self-esteem score}} = 27.13$; $M_{\text{age}} = 22.5$ years) were randomly assigned to the “other-failure” or the “other-no-feedback” condition with the restriction that an approximately equal number of men and women were assigned to each condition. For their participation, subjects received an experimental credit.

Procedure
The procedure was the same as in Experiment 2, except where we note here. Each subject in the present study completed the social accuracy test in the presence of a high-self-esteem subject who also completed the same test, but who then participated in a separate experiment. Subjects scored their own tests using the bogus answer keys. (Subjects did not score each other’s tests because they were not always paired with a person who received failure feedback.) These keys provided success feedback to all subjects in the present study (i.e., the low-self-esteem subjects). Once the tests were “scored,” the high-self-esteem subject went to a different room and completed different measures for the other study, which is not reported here.

The manipulation of knowledge of the other’s score occurred once the subjects were separated. In the “other-failure” condition, the experimenter showed the subjects the “other subject’s” below average test score, ostensibly because they would need to know the other subject’s test score to do the impression-formation task. In the “other-no-score” condition, there was no mention of the other subject’s test score. To make the procedure as similar to Experiment 2’s as possible, subjects were told that their own test score had been shown inadvertently to the other subject. Subjects then completed the attribution questionnaire that
was designed to enhance the success feedback as well as the same manipulation check questions about performance used in Experiment 2.

Comparison-Selection Measure
The comparison-selection measure was then presented in the same way as in Experiment 2, except that subjects were informed that each of the tests was related to the social accuracy test. This modification was made to ensure that subjects would have a reasonable expectation of the outcome of any comparisons that they made (and hence would know whether they were “safe”). The list of tests included 14 of those used in Experiment 2 that sounded most related to social accuracy, plus: ability to form relationships, coping skills inventory, need for affiliation, friendliness index, social support index, and conflict resolution skills. Subjects also were told that in addition to receiving their own and the other’s scores, they would receive normative information for any tests that they selected. [12]

Results and Discussion

Manipulation Checks
Because there were no failure subjects in the present study, subjects’ responses to the questions concerning performance and satisfaction with performance were compared with those of the high-self-esteem failure subjects who participated in the separate study. The averages of these items ( = .92) were entered into an ANOVA with success–failure and sex as between-subjects factors. Although the success–failure factor was confounded with self-esteem (Experiment 3’s low-self-esteem subjects succeeded and the high-self-esteem subjects failed), this confound militated against finding that Experiment 3’s success manipulation was effective, because low-self-esteem subjects typically have more negative views of their performance than do high-self-esteem subjects (as was found in Experiment 2). A main effect for success–failure, $F (1, 44) = 420.07, p < .001$, revealed that success (low self-esteem) subjects thought that they had done better and were more satisfied ($M = 6.60$) than failure (high self-esteem) subjects ($M = 1.77$). Success subjects in the other-failure and other-no-score conditions did not differ ($M_s = 6.81$ and 6.36, respectively), $t (55) = -1.32, p = .195$. Thus, we conclude that the success feedback was effective.

Avoiding Versus Selecting Comparisons
We predicted that subjects who were shown the other subject’s below average test result would seek more comparisons than subjects who were not informed of the other’s failure. An ANOVA with the factors of condition and sex revealed that subjects in the other-failure condition made more matches ($M = 2.77$) than did subjects in the other no-score condition ($M = 1.73$), $F (1, 20) = 3.64, p = .035$, one-tailed. The other-failure subjects matched at rates above chance, $t (20) = 4.10, p < .001$, but the other-no-score subjects did not, $t (20) = 1.18$. These results support our hypothesis that low-self-esteem people seek comparisons after success only if they promise a favorable outcome.
General Discussion
In this research, we investigated ways in which people may seek social comparisons to benefit their self-esteem. We attempted to extend the literature on this topic in three ways. First, we examined two types of self-esteem bolstering, namely, self-protection and self-enhancement, and we predicted that low-self-esteem subjects would engage in both types, but under different circumstances. Second, we examined the strategy of avoiding comparisons, which has received relatively little empirical attention. Third, we introduced a new measure of social comparison. We discuss this last topic first.

A New Comparison-Selection Measure
The results of our new measure of comparison selection are very promising. The interview study suggests that our interpretation of the selection of matches as comparison seeking is reasonable, in that naive respondents easily recognized that subjects could select matches for social comparison purposes. Experiments 1 and 2 demonstrate that the measure yields consistent results concerning comparison selections and avoidance, and all three experiments indicate that it can provide clues about the motivations that underlie these strategies. To explore motivations, we examined subjects’ expectations for comparison in Experiments 1 and 2 and we varied the available comparison targets in Experiment 3. We believe that our measure more closely parallels naturalistic circumstances of comparison than do most previous measures of comparison selection (Wood, 1994). Most measures require subjects to choose a comparison target; because this one does not, it allowed subjects to seek or avoid comparisons. Most measures confine subjects’ comparisons to a single dimension; because our measure does not, it allowed subjects to choose their comparison dimensions. This measure may be used, then, to investigate the particular dimensions that people select for comparison. People may select their comparison dimensions strategically (Wood, 1989). For example, in a study of cancer patients, respondents often drew comparisons on dimensions on which they were advantaged rather than disadvantaged (Wood, Taylor, & Lichtman, 1985). By not constraining subjects’ comparisons, then, this new measure enables researchers to examine strategies of comparison that have received relatively little attention in the past, namely, avoiding comparisons and selecting comparison dimensions. Most studies of self-enhancing social comparisons have been focused solely on target selection (e.g., downward comparisons). These other strategies allow the individual more flexibility than does target selection, because the individual is not limited by whatever comparison targets happen to be available (Wood & Taylor, 1991; Wood et al., 1985). We believe that continued theoretical growth in social comparison requires the development of measures such as the present one.

Across the three experiments, we varied some aspects of the procedure involved in the test-selection measure. The results of these variations offer advice to researchers who may use our measure. First, one may obtain a wider range of comparison-seeking by having subjects select 5 out of 20 tests rather than 3 out of 13. Second, if one’s overall context is private,
subjects may make more comparisons. However, subjects also may be less affected by the manipulation, and their degree of comparison-seeking may not vary as widely.

Avoiding Versus Selecting Comparisons
Both Experiments 1 and 2 yielded an interaction that indicated that low- and high-self-esteem people have different patterns of comparison-seeking after success and failure. Experiment 1 indicated that low-self-esteem subjects who succeeded were more likely to seek comparisons than were low-self-esteem subjects who failed and high-self-esteem subjects who succeeded. Experiment 2 replicated this pattern and also indicated that high-self-esteem subjects who failed were, like the low-self-esteem success subjects, especially eager to seek comparisons. Thus, both studies obtained a very similar pattern of findings despite different manipulations of success–failure, different measures of self-esteem, and other differences in procedure. It is clear, then, that we have identified a phenomenon.

What is the explanation for this phenomenon? We obtained evidence against one account, namely that low-self-esteem success and high-self-esteem failure subjects were puzzled by or disbelieved the results and sought comparisons to resolve their confusion. However, further research is required to identify the correct explanation for the phenomenon. We suspect that the explanation lies in motivations that differ between people with low and high self-esteem after success and failure, and we have gathered some preliminary evidence for this explanation in the present research.

Comparisons After Failure: Self-Protection and Compensation
Specifically, our tentative explanation is based on the view that low- and high-self-esteem people use different strategies of self-esteem bolstering (Baumeister et al., 1989; Brown et al., 1988). Low-self-esteem people are oriented mainly toward self-protection, that is, they wish to avoid exposing their deficiencies. We did not find that low-self-esteem subjects avoided comparisons in that they chose fewer than the number of chance-expected matches, or fewer matches than in a no-feedback condition. However, as noted earlier, our chance baseline may have been underestimated, and the no-feedback subjects may themselves have avoided comparison. Consistent with the view that people with low self-esteem are self-protective, our low-self-esteem subjects were less inclined to make further comparisons after failing than after succeeding.

In contrast to the self-protective orientation of low-self-esteem people, high-self-esteem people typically seek self-enhancement (Baumeister et al., 1989). Not content merely to avoid their failures, they attempt to compensate for them (Baumeister, 1982; Baumeister & Tice, 1985). We found that people with high self-esteem sought comparisons after they failed with someone who appeared to be superior to them. We suspect that they were trying to compensate for the failure, because they had predicted that the comparisons would reveal them to be superior.
Comparisons After Success: Lack of Interest and Reveling
When they had succeeded, however, subjects with high self-esteem were less interested in comparing further. This result is consistent with past research, which has indicated that, except under conditions of threat, people do not like to make comparisons with clearly inferior opponents (e.g., Gastorf et al., 1978). The present studies indicate that this is true only for high-self-esteem people; low-self-esteem people do like to make further comparisons with people who failed against them. Past research has indicated that low-self-esteem people take pleasure in their successes (e.g., McFarlin & Blascovich, 1981; Shrauger, 1975; Swann et al., 1987), and that they often take pleasure in downward comparisons specifically (e.g., Aspinwall & Taylor, 1993; Gibbons & Gerrard, 1989). This research indicates that low-self-esteem people may go beyond merely enjoying their success; by comparing themselves with someone who has already been shown to be inferior to them, they can take advantage of a safe opportunity to revel in their success.

Experiment 3 supported this interpretation of low-self-esteem people’s comparison-seeking after success. They were more likely to seek comparisons when they knew that the other had failed, and hence was inferior to themselves, than when they did not know of the other’s score. Only the first condition provided an opportunity for self-enhancement that was “safe,” that is, that guaranteed a favorable comparison outcome.

Our discussion of safe comparisons implies that the safety of a comparison targets determined externally. However, some research has suggested that people may actively “create” comparison target is that suit their needs (Wood, 1989). People may distance themselves from comparison others cognitively (Gibbons, Gerrard, Lando, & McGovern, 1991) or physically (Pleban & Tesser, 1981). They may derogate others (Salovey & Rodin, 1984) or even imagine comparison others who are inferior (Goethals, 1986; Wood & Taylor, 1991). Is it possible that our low-self-esteem subjects used such strategies to “create” safe comparison targets? Some of our data suggest this possibility. After receiving favorable feedback in one domain, low-self-esteem subjects had favorable performance expectancies in additional domains—even domains they were told were unrelated to their success. On the other hand, some evidence supports the view that safety must be externally provided: Experiment 3 indicates that it is not enough for low-self-esteem people to believe that they have succeeded; they also need to know that the other person has failed before they will venture to compare. Perhaps, then, low-self-esteem people attempt some creation of targets, but within some constraints. Further research is needed to explore this intriguing issue.

Beyond Self-Protection Among People With Low Self-Esteem
Increasingly, a picture is emerging of low-self-esteem people as cautious people, oriented toward avoiding exposure of their weaknesses, rather than toward self-enhancement (Baumeister et al., 1989). Our evidence suggests, however, that low-self-esteem people will venture beyond their self-protective orientation and seek self-enhancement when it is safe to do so. It is possible, then, that if researchers provide similarly safe opportunities to use other self-esteem boosting strategies, low-self-esteem people will use them. Such evidence would
argue against the view (e.g., Taylor & Brown, 1988) that self-enhancement strategies are the exclusive province of people with high self-esteem.

References

Social Psychology.
Footnotes

1 ^ Three out of four of the suspicious subjects were in the failure condition. Subjects who claimed that they did not notice the feedback also tended to be in the failure condition (5 of 7).

2 ^ The experimenter showed the subject two slips of paper on which were written “rater” and “person completing questionnaires,” respectively. The slips of paper were then inserted into the top of a box that had a secret compartment inside the lid, into which the slips of paper fell. Two other “rater” slips had been placed earlier in the main compartment of the box, so that when the experimenter removed the lid, the subject chose one of the two “rater” slips.

3 ^ These interactions qualified lower-order main effects of success–failure, $F (1, 85) = 7.72, p = .007$, such that failure subjects made higher average ratings than did success subjects; and self-esteem, $F (1, 85) = 10.54, p = .002$, such that high-self-esteem subjects made higher ratings than did low-self-esteem subjects; and self–other, $F (1, 85) = 7.14, p = .009$, such that most subjects rated the self higher than the other; as well as a Self-Esteem ×Self–Other interaction, $F (1, 85) = 5.42, p = .022$, such that high-self-esteem subjects tended to have bigger self–other differences than did low-self-esteem subjects.

4 ^ Subjects also had been told that test norms were available for each test and that they should select three of these. Their choices could involve tests that were the same or different from those they chose for the self or other. Results revealed that to a high degree, subjects chose to see norms for the same tests that they had selected for the self. Forty-five percent (42) chose to see norms for all three tests that they planned to take, and only 2.2% of the subjects (2) did not choose to see norms for any of the tests that they planned to take. By comparing these results with those in Table 2, one can see that subjects' selections were greatly higher than chance expectations and that they even exceeded matches with the other subject. There were no differences between conditions. Subjects’ strong interest in how they compared with norms suggests that although their ostensible task was to form an impression of the other subject, they used the opportunity to learn more about the self.

5 ^ All the contrasts reported in this article used the mean square error term from the corresponding complete ANOVA design with self-esteem and condition as between-subjects variables. Contrasts involving the self- and other-ratings were repeated-measures contrasts.
in which the error term was the within-subjects mean square error term from ANOVAs with the same between-subjects variables, and with average self- and other-ratings as repeated measures (Rosenthal & Rosnow, 1985).

6 ^ We thank Rick Gibbons for pointing this out.

7 ^ The first 10 tests were listed under the heading, “Tests related to social accuracy,” and the second 10 were listed under “Tests unrelated to social accuracy.” We designated “related” and “unrelated” dimensions so that we could determine whether subjects would favor one or the other. However, most subjects chose related and unrelated matches to an equal degree, possibly because they perceived a demand to distribute their selections between these labeled categories.

8 ^ This interaction qualified main effects for condition, \( F (2, 60) = 7.67, p = .001 \), and self-esteem, \( F (1, 60) = 8.83, p = .004 \), such that failure subjects sought the most matches (\( M = 2.67 \), \( M_s \) for no-feedback and success = 1.58 and 2.38, respectively), and high-self-esteem subjects sought more matches (\( M = 2.56 \)) than did low-self-esteem subjects (\( M = 1.86 \)). The presence of these effects after their absence in Experiment 1 is attributable to the addition of the no-feedback group, in which a lower level of matches was observed (enhancing the condition main effect), along with higher numbers of matches made by high-self-esteem subjects in that condition (enhancing the self-esteem main effect). Indeed, these main effects are absent when the no-feedback condition is removed from the analysis.

9 ^ This study was conceived to answer a question by Russell Geen. Ten of the 20 subjects were provided with information about the relatedness of the dimensions to social accuracy. They were told that “Included with each description of an ability is information about whether the ability is related or unrelated to a particular ability, namely social accuracy. Social accuracy involves the ability to make accurate judgments about other people. You are not required to use this information about relatedness to social accuracy in any particular way, but you may take it into account if it seems relevant.” Separate “best five” and “all 25” comparison scores were constructed for the subjects in Experiment 2 from the overlap ratings made by these 10 subjects, from the 10 subjects who were told nothing about relatedness, and from the 20 subjects combined. The result for the 20 subjects are reported in the text, but results of all six ANOVAs yielded the same Condition × Self-Esteem interaction (all \( p < .001 \)), regardless of whether subjects received the information about relatedness.

10 ^ When the critical conditions were examined separately, the puzzled–skeptical composite and the perceived validity composite also were not correlated significantly with the number of matches. In the low-self-esteem success condition, these correlations were .03 and .45, respectively; in the high-self-esteem failure condition, the correlations were -.37 and .06, respectively. These values, although not significant, were in the direction opposite to the puzzled–skeptical hypothesis; the .45 suggests that the more confident subjects were in the validity of the test, the more they chose matching tests; and the -.37 suggests that the less puzzled and skeptical they were, the more they sought matching tests.

11 ^ Several additional effects were not pertinent to the puzzled–skeptical hypothesis: a main effect for condition, \( F (2, 60) = 265.57, p < .001 \), showing that success subjects were most likely to believe the results, failure subjects were least likely, and no-feedback subjects were in between (\( M_s \) = 6.05, 2.10, and 4.30, respectively); a main effect of self-esteem, \( F (1, 60) = 6.21, p = .016 \), with high-self-esteem subjects more likely to believe the results (\( M = 4.33 \)) than low-self-esteem subjects (\( M = 3.98 \)); a Condition × Sex interaction, \( F (2, 60) = 5.30, p = .008 \), indicating that men who failed were least likely to believe the results; and a
Self-esteem × Sex interaction, $F (1, 60) = 4.81, p = .032$, indicating that low-self-esteem men were least likely to believe the results ($M = 3.87$), high-self-esteem men were most likely to believe the results ($M = 4.53$), and low- and high-self-esteem women did not differ ($M_s = 4.10$ and $4.13$, respectively). Is it worrisome that subjects who failed questioned the validity of the test; that is, did our manipulation fail? This result is very common (e.g., Schlenker et al., 1990; see Taylor & Brown, 1988, for a review). We interpret our results to mean that when provided with a possible face-saving attribution for their failure, namely, that the test may have been in valid, failure subjects seized on it. These same subjects, however, had reported earlier that they were displeased and humiliated by their failure, and their expectations for comparison were less favorable than those of success subjects.

12 ^ We told subjects that normative data would be available because of the possibility that subjects may choose tests for the self that match those for the other subject because they could not interpret the other’s scores without some basis of comparison. We had thought that subjects would assume that the other subject’s scores would be presented in a way that would allow them to interpret those scores (e.g., “below average”). However, it is probably best to tell subjects explicitly that they will receive normative information, and we thank Ladd Wheeler for making this suggestion. It is important to note that this possibility, that is, that subjects may choose matches to help them to interpret the others’ scores, does not jeopardize the interpretation of our findings, because it cannot account for the interaction that we obtained between self-esteem and condition. Indeed, it would work against finding group differences. (It is also interesting that although subjects in Experiment 1 had normative information available for each test, they sought such information primarily for the tests they picked for themselves, not for the tests they chose for the other.)

13 ^ We thank a reviewer for raising this point.

Address for Correspondence:
Joanne V. Wood, Department of Psychology, University of Waterloo, Waterloo, Ontario Canada N2L 3G1

© 1994 American Psychological Association

Indexing (details)

Subjects Feedback (major), Impression Management (major), Self Esteem (major), Social Comparison (major)

Classification 3040: Social Perception&Cognition

Age Adulthood (18 yrs&older)

Population Human
Keywords (identifiers) failure vs success feedback, self enhancement vs self protection in social comparison, college students with high vs low self esteem

Methodology Empirical Study

Title Strategies of social comparison among people with low self-esteem: Self-protection and self-enhancement.

Authors Wood, Joanne V.; Giordano-Beech, Maria; Taylor, Kathryn L.; Michela, John L.; Gaus, Valerie

Publication title Journal of Personality and Social Psychology

Volume 67

Issue 4

Pages 713-731

Publication year 1994

Format Availability Electronic, Print

Format Covered Print

Publisher American Psychological Association

Country of publication United States

ISSN 0022-3514

eISSN 1939-1315

Peer reviewed Yes

Document type Journal, Journal Article, Peer Reviewed Journal

Publication history

Accepted date 03 May 1994

Revised date 26 Apr 1994, 08 Mar 2010

First submitted date 17 Oct 1992

DOI 10.1037/0022-3514.67.4.713

Release 01 Feb 1995 (PsycINFO)

10 Jul 2006 (PsycARTICLES)

Correction date 26 Apr 1994

08 Mar 2010 (PsycINFO)

08 Mar 2010 (PsycARTICLES)

Accession number 1995-05356-001