CURRENT RESEARCH IN SOCIAL PSYCHOLOGY
http://www.uiowa.edu/~grpproc/crisp/crisp.html

Submitted: July 27, 2010
First Revision: February 1, 2011
Accepted: February 13, 2011
Published: April 5, 2011

THE PERFORMANCE VERSUS ABILITY DISTINCTION FOLLOWING SOCIAL COMPARISON FEEDBACK

Justin T. Buckingham
Towson University

Lavonia Smith LeBeau
Human Services Research Institute

William M. P. Klein
University of Pittsburgh

ABSTRACT

Previous research suggests that self-evaluations can be influenced by social comparison feedback. The present study tested whether social comparison feedback has stronger effects on self-evaluations of performance than ability. Participants received social comparison feedback indicating that they had performed above or below average. In addition to rating their performance and ability, participants listed the criteria they used to make these ratings. The results show that social comparison had a stronger effect on performance judgments than ability judgments and that most participants reported using test feedback to evaluate performance and existing self-knowledge to evaluate ability.

INTRODUCTION

It has now been over 50 years since Festinger (1954) introduced his influential social comparison theory. In the past 20 years, research on social comparison has shifted to some degree from examining its antecedents to investigating its consequences. Many published studies show that the typical consequence of explicit social comparison (i.e., using others as a standard for evaluating oneself) is a contrast effect (Broemer and Diehl 2004; Stapel and Koomen 2000; 2001; Stapel and Blanton 2004). That is, comparing to an upward target lowers self-evaluations and comparing to a downward target enhances self-evaluations. Although contrast effects are typical, a number of factors have been shown to moderate the effects of social comparison on self-evaluation. Among these factors are psychological closeness to the target (Tesser, 1988), distinctiveness of the target (Stapel and Koomen 2000), mutability of self (Lockwood and Kunda 1997; Stapel and Koomen 2000), self-construal (Stapel and Koomen 2001), self-esteem (Jones

and Buckingham 2005), and self-certainty (Pelham and Wachsmuth 1995). As exemplified by this list, research on moderators of social comparison effects has largely focused on characteristics of the target and the individual making the comparison.

Until recently, researchers have neglected to investigate the dimension on which individuals evaluate themselves as a possible moderator of social comparison effects; the present study is among the first to do so. There is reason to believe that the specificity of the evaluation dimension moderates the extent to which people apply relevant information. Logically, feedback about a specific instance should be more readily applied to self-evaluations of performance than to self-evaluations of the more general ability tested. For self-evaluations of performance, the feedback is the only information available. For self-evaluations of ability, the feedback may be weighed in the context of existing self-knowledge and therefore lose much of its impact.

Recent studies (Buckingham and Alicke 2002; Chambers and Windschitl 2009) have supported this logic by finding that effects of comparative feedback sometimes depend on the specificity of the evaluation dimension. That is, social comparison has been shown to have greater effects on more specific self-evaluations than on more global self-evaluations. In Buckingham and Alicke's studies, participants took a test of lie detection ability and received feedback indicating that they had scored above or below average and better or worse than a co-participant. Both comparison standards had significant orthogonal effects on participants' ratings of how well they had performed, but only comparison to the co-participant significantly affected their ratings of general lie detection ability. It appears that whereas participants were bound to use social comparison information to evaluate their *test performance*, they were impervious to at least some of the feedback when rating their *lie detection ability*. In Buckingham and Alicke's studies, it may have been easier for participants to withstand comparisons to the statistical average than to a salient individual.

The aim of the present research was to provide a further test of the hypothesis that social comparison feedback has stronger effects on self-evaluations of performance than ability. Furthermore, the present research investigates a mechanism that may explain the performance versus ability distinction. We describe this mechanism in the next section.

Ambiguity of trait ratings

We suspect that one mechanism that helps people maintain desirable self-evaluations of ability following feedback is the use of a broader range of self-evaluation criteria. We assume that whenever people evaluate themselves they must balance their preference for self-enhancement with a need for accuracy (Klein and Buckingham 2002; Kunda 1990). When people receive test feedback, their performance is verifiable. Therefore, when rating performance, it is difficult to use a biased self-evaluation process while also maintaining accuracy. Thus, ratings of performance are likely to be influenced by recent feedback. However, when rating ability, people are more likely to use additional criteria (such as their existing self-views) to minimize the effects of feedback (particularly unfavorable feedback) while still maintaining an "illusion of objectivity" (Armor 1999). This is evidenced by Buckingham and Alicke's (2002) finding (in Study 5) that feedback had a significant effect on ability ratings when the ability rating made reference to the specific performance (i.e., "As measured by this test, how would you rate your

lie detection ability?"). In addition, Dunning, Meyerowitz, and Holzberg (1989) have demonstrated that participants who are allowed to use a broader range of criteria have more positive self-evaluations than those who are forced to use very specific criteria. The general rule seems to be that ambiguity facilitates self-enhancement whereas specificity impairs it (Klein & Buckingham, 2002). This is also true when rating oneself relative to a target – self-enhancement is greater when comparing with less specific targets (Alicke, Klotz, Breitenbecher, Yurak, and Vredenburg 1995).

The present study

In the present study, we expanded on Buckingham and Alicke's (2002) previous work by examining the criteria people use to evaluate performance and ability; this has not been done in previous studies and is an important step for understanding why social comparison has stronger effects on self-evaluations of performance than ability.

Participants were led to believe they scored above or below average on a social sensitivity test and then evaluated their performance and social sensitivity. Participants also listed the information they used to rate performance and social sensitivity. Consistent with Buckingham and Alicke (2002), we hypothesized that comparative feedback would have a stronger effect on self-evaluations of performance than ability. We also hypothesized that the distinction between performance and ability would be greatest in the condition in which participants made upward comparisons—which would suggest that participants are motivated to avoid using unfavorable feedback. On the criteria lists, we hypothesized that participants would be more likely to report using test feedback as a criterion for evaluating performance than social sensitivity; in contrast, we hypothesized that participants would be more likely to list chronic self-views and previous behavior as criteria for evaluating their social sensitivity than performance.

METHOD

Participants and design

136 Towson University psychology students volunteered to participate in the experiment. They were randomly assigned to the upward or downward social comparison conditions. We also varied the objective scores used for feedback. An initial set of 83 participants were assigned scores of 11 out of 20 on the test of social sensitivity. These participants were led to believe the average score was either 16.02 (upward comparison) or 6.02 (downward comparison). During the same academic term, we collected data from a second set of 53 participants from the same participant pool. These participants were assigned scores of 14 out of 20 and averages of either 8.02 or 18.02. Other than the numerical change in the feedback, all procedures remained the same. Because the change in scores resulted only in a main effect [1] and this factor did not qualify any of our findings, we will report results from the combined data set and not discuss objective scores further. The dependent variables were ratings of performance and social sensitivity as well as open-ended lists of the criteria participants used to evaluate performance and ability.

Procedure

Participants arrived in groups of 2 or 3 and were seated in individual laboratory cubicles. Following from the procedure used in Pyszczynski, Greenberg, and LaPrelle (1985), participants read an instruction sheet that defined social sensitivity as "the ability to perceive other peoples' reactions to events; to know what they're thinking and feeling" (p. 199) and explained that they would be taking the Illinois Social Sensitivity Test (ISST). In addition, participants were told that people who do well on this test tend to have happier, more successful friendships, marriages, and other interpersonal relationships. The fictitious ISST we used required that participants select the word most commonly (based on a large survey) associated with a target word.

Participants were allowed 7 minutes to complete the 20-item test. After collecting the tests, the experimenter went to another laboratory room ostensibly to score the tests. After 2 minutes, the experimenter returned with feedback indicating the participant's score and the average score, which was said to be based on 250 Towson students who had already taken the test.

Participants then completed a questionnaire in which they were asked to recall their test score and the average score. Participants also rated their performance, "How well do you think you performed on the social sensitivity test?" and ability, "How would you rate your social sensitivity?" on scales from 0 (very poor) to 10 (very good). In two separate items on the same page, participants were asked to list all the information they used to rate their test performance and social sensitivity, respectively.

RESULTS

Excluded data

Six participants failed to accurately recall the average score and/or their own score and their data were therefore excluded from the analyses.

Self-Evaluations

The primary analysis was a 2 (Social Comparison: upward vs. downward) X 2 (Evaluation Dimension: performance vs. ability) ANOVA with repeated measures on the evaluation dimension factor because each participant self-evaluated on both performance and ability. This analysis revealed a significant main effect of evaluation dimension such that participants evaluated their social sensitivity ability ($M = 7.23$, $SD = 1.53$) more favorably than their test performance ($M = 6.22$, $SD = 1.76$), $F(1, 128) = 60.32$, $p < .001$, $r = .57$ [2]. Furthermore, there was a significant main effect of social comparison direction such that participants in the downward comparison condition ($M = 7.24$, $SD = 1.14$) had higher self-evaluations (collapsed across the performance vs. ability dimension) than participants in the upward comparison condition ($M = 6.11$, $SD = 1.50$), $F(1, 128) = 24.05$, $p < .001$, $r = .40$, suggesting that the manipulation of feedback valence was successful.

Of most importance, these main effects were qualified by the hypothesized Social Comparison X Evaluation Dimension interaction, $F(1, 128) = 19.93$, $p < .001$. Simple effects tests showed that social comparison had a strong and significant effect on self-evaluations of performance, $F(1,$

128) = 41.67, $p < .001$, $r = .50$. Participants in the downward comparison condition ($M = 7.01$, $SD = 1.22$) had more favorable evaluations of their performance than participants in the upward comparison condition ($M = 5.27$, $SD = 1.84$). By comparison, the effect of social comparison on self-evaluations of social sensitivity was weaker and non-significant, $F(1, 128) = 3.86$, $p > .05$, $r = .17$. Participants in the downward comparison condition ($M = 7.46$, $SD = 1.39$) had slightly more favorable self-evaluations of their ability than participants in the upward comparison condition ($M = 6.94$, SD = 1.65), but the difference was much smaller than on the performance dimension. As predicted, the effect size for performance was significantly larger than for ability, $z = 3.02$, $p < .05$.

To specifically examine the hypothesis that performance versus ability differences would be found in the upward comparison condition more so than in the downward comparison condition, we also tested the simple effects of evaluation dimension within each feedback condition. There was a moderately strong and significant effect of evaluation dimension in the downward comparison condition, $F(1, 70) = 8.44$, $p < .01$, $r = .33$, with participants rating their social sensitivity ($M = 7.46$, $SD = 1.39$) more favorably than their performance ($M = 7.01$, $SD = 1.22$). As expected, a stronger effect was found in the upward comparison condition, $F(1, 58) = 50.79$, $p < .001$, $r = .68$ with participants again rating their social sensitivity ($M = 6.94$, $SD = 1.65$) significantly higher than their performance ($M = 5.27$, $SD = 1.84$). Importantly, the effect size in the upward comparison condition was significantly larger than in the downward comparison condition, $z = 2.70$, $p < .05$.

Criteria lists

A student who was unfamiliar with our predictions and blind to condition coded participants' lists of the criteria they reported using to evaluate their performance and ability. The coder judged whether participants had mentioned each of the following three types of information in their criteria lists: (a) test score, (b) the average score or performance of others, and (c) behaviors or self-views unrelated to performance on this specific test (e.g., "I'm very sensitive with my friends' feelings"). We checked the reliability of the coder's ratings by having another student independently code responses from 25% of the sample. Inter-rater reliability for each of the judgments was acceptable (α's ranged from .78 to .94).

We analyzed the coded responses for each type of information with a 2 (Social Comparison: upward vs. downward) X 2 (Mention Information: yes vs. no) chi-square analysis. On the lists of information used to evaluate their test performance, the majority of participants mentioned social comparison information (55%) and their score (64%), but very few participants (15%) mentioned behaviors or self-views unrelated to this specific test. Social comparison feedback had no significant effect on whether participants mentioned their score, social comparison, or self-views, $p > .05$.

We observed a very different pattern of results on the social sensitivity criteria lists. Here, most participants mentioned behaviors or self-views unrelated to the specific test (76%) and relatively few mentioned their scores (23%) or the comparison feedback (13%). In addition, there was a significant effect of social comparison on whether participants mentioned social comparison

information as a criterion. This effect is consistent with a self-serving bias: Participants who received downward comparison feedback mentioned the social comparison information more frequently than participants who received upward comparison feedback (19% vs. 5%) Chi-square (1) = 5.35, $p < .05$. Social comparison did not significantly affect whether participants mentioned their scores or self-views, *p's* > .05.

DISCUSSION

The current study replicates and extends previous research by Buckingham and Alicke (2002) showing that aggregate social comparison feedback has stronger effects on ratings of performance than corresponding ability. The main purpose of the present study was to investigate the mechanism(s) underlying the performance versus ability distinction and to explore the role of self-enhancement. We hypothesized and found that participants reported using different criteria to evaluate performance vis-à-vis ability. In the present study, most participants reported using test feedback to evaluate performance and information about their previous behavior or existing self-views to evaluate social sensitivity.

The present data also show that the performance versus ability distinction is larger when people receive unfavorable feedback than when they receive favorable feedback. We expected participants to show a similar preference for downward comparison information on their lists of criteria used to evaluate ability. In support of this hypothesis, social comparison direction had a significant effect on whether participants listed the feedback as a criterion for evaluating their abilities. Participants who scored above average were almost four times as likely to mention using the average as a criterion for evaluating their social sensitivity than those who scored below average.

Are differences in criteria responsible for differences in self-evaluations of performance and ability? Correlations indicated that listing one's test score as a criterion for evaluating ability was negatively related to ratings of social sensitivity ($r = -.25$, $p < .01$) whereas listing self-views as a criterion was positively related to such ratings ($r = .25$, $p < .01$). However, these data do not rule out possible third variables or the possibility that differences in ratings led to differences in criteria lists. For instance, it is possible that participants used the criteria lists as a way to justify their ratings. In addition, the fact that criteria lists and self-evaluations were only modestly correlated suggests that the criteria lists only captured a small portion of the criteria participants actually used. This result is probably due the fact that people do not always have accurate insight into the criteria they use to self-evaluate (cf. Nisbett and Wilson 1977). Nevertheless, the overall pattern of results for criteria lists matches the pattern for quantitative self-evaluations quite well. On ratings of performance, most participants said they used social comparison as a criterion and social comparison had a significant effect. On ratings of ability, most participants said they did not use social comparison as a criterion and social comparison had no significant effect.

The present findings suggest an important limitation to social comparison effects. Previous research has shown that social comparison can lead people to see themselves as more or less intelligent, skilled, or attractive (e.g., Stapel and Koomen 2001). Such effects have most reliably been demonstrated when the relevance of the social comparison for self-evaluation is carefully hidden (e.g., by exposing participants to a social comparison target as part of a bogus social

perception study). When people receive explicit social comparison feedback (e.g., comparative performance feedback), they may be better prepared to resist being influenced by such feedback – particularly if they have the cognitive resources to do so (Gilbert, Giesler, and Morris 1995). Therefore, despite the affective consequences of upward and downward comparison, people may sometimes be able to avoid fluctuations in their self-views at the trait level.

The present research suggests that making a distinction between specific and global evaluations may serve to maintain favorable overall perceptions of the self. In the achievement domain, the distinction may have a self-evaluation maintenance function. Distinguishing between performance and ability may allow people to overcome minor obstacles such as poor performance and maintain favorable self-evaluations so that they continue pursuing important goals.

ENDNOTES

1. The effect of objective information was significant, $F(1, 126) = 11.64$, $p < .01$, $r = .29$. Participants who received higher objective scores ($M = 7.19$, $SD = 1.18$) had more favorable self-evaluations than participants who received lower objective scores ($M = 6.41$, $SD = 1.50$).

2. Effect sizes in this paper are all reported in terms of the Pearson correlation (r), as suggested by Rosenthal and Rosnow (1991). According to Cohen (1992), r values of .10, .30, and .50 are considered small, medium, and large, respectively.

REFERENCES

Alicke, Mark D., M. L. Klotz, David L. Breitenbecher, Tricia J. Yurak, and Debbie S. Vredenburg. 1995. "Personal Contact, Individuation, and the Better-Than-Average Effect." *Journal of Personality and Social Psychology* 68:804-25.

Armor, David A. 1999. "The Illusion of Objectivity: A Bias in the Perception of Freedom from Bias." *Dissertation Abstracts International: Section B* 59:5163.

Broemer, Philip and Michael Diehl. 2004. "Evaluative Contrast in Social Comparison: The Role of Distinct and Shared Features of the Self and Comparison Others." *European Journal of Social Psychology* 34:25-38.

Buckingham, Justin T. and Mark D. Alicke. 2002. "The Influence of Individual versus Aggregate Social Comparison and the Presence of Others on Self-Evaluations." *Journal of Personality and Social Psychology* 83:1117-30.

Chambers, John D. and Paul D. Windschitl. 2009. "Evaluating One Performance Among Others: The Influence of Rank and Degree of Exposure to Comparison Referents." *Personality and Social Psychology Bulletin* 35:776-92.

Cohen, Jacob. 1992. "A Power Primer." *Psychological Bulletin* 14:155-9.

Dunning, David A, Judith A. Meyerowitz, and Amy D. Holzberg. 1989. "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability." *Journal of Personality and Social Psychology* 57:1082-90.

Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations* 7:117-40.

Gilbert, Daniel. T., R. Brian Giesler, and Kathryn A. Morris. 1995. "When Comparisons Arise." *Journal of Personality and Social Psychology* 69:227-36.

Jones, April M. and Justin T. Buckingham. 2005. "Self-Esteem as a Moderator of the Effect of Social Comparison on Women's Body Image." *Journal of Social and Clinical Psychology* 24:1086-1109.

Klein, William M. P. and Justin T. Buckingham. 2002. "Self-Other Biases in Judgments of Ambiguous Performance and Corresponding Ability." *Psychologica Belgica* 42:43-64.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108:480-98.

Lockwood, Penelope and Ziva Kunda. 1997. "Superstars and Me: Predicting the Impact of Role Models on the Self." *Journal of Personality and Social Psychology* 73:91-103.

Nisbett, Richard E. and Timothy D. Wilson. 1977. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84:231-59.

Pelham, Brett W. and Jeff Orson Wachsmuth. (1995). "The Waxing and Waning of the Social Self: Assimilation and Contrast in Social Comparison." *Journal of Personality and Social Psychology* 69:825-38.

Pyszczynski, Tom, Jeff Greenberg, and John LaPrelle. 1985. "Social Comparison after Success and Failure: Biased Search for Information Consistent with a Self-Serving Conclusion." *Journal of Experimental Social Psychology* 21:195-211.

Rosenthal, Robert and Ralph L. Rosnow. 1991. *Essentials of Behavioral Research: Methods and Data Analysis* (2nd edition). New York: McGraw-Hill.

Stapel, Diederik A. and Hart Blanton. 2004. "From Seeing to Being: Subliminal Social Comparisons Affect Implicit and Explicit Self-Evaluations." *Journal of Personality and Social Psychology* 87:468-81.

Stapel, Diederik A. and Willem Koomen. 2000. "Distinctness of Others, Mutability of Selves: Their Impact on Self-Evaluations." *Journal of Personality and Social Psychology* 79: 1068-87.

Stapel, Diederik A. and Willem Koomen. 2001. "I, We, and the Effects of Others on Me: How Self-Construal Level Moderates Social Comparison Effects." *Journal of Personality and Social Psychology* 80:766-81.

Tesser, Abraham. 1988. “Toward a Self-Evaluation Maintenance Model of Social Behavior.” Pp. 181-227 (vol. 21) In *Advances in experimental social psychology: Social Psychological Studies of the Self: Perspectives and Programs,* edited by Leonard Berkowitz. San Diego: Academic Press.

APPENDIX: CORRELATION MATRIX FOR ANOVA

	Comparison Feedback	Performance Rating	Social Sensitivity Rating
Comparison Feedback	1.00	-.50*	-.17
Performance Rating		1.00	.50*
Social Sensitivity Rating			1.00
Mean		6.22	7.23
Standard Deviation		1.76	1.53

Note. N = 130. For comparison feedback, 1= Downward 2= Upward.
*$p < .01$.

AUTHORS’ NOTE

This research was presented at the 2004 meetings of the Midwestern Psychological Association and the Southeastern Society of Social Psychologists.
We are grateful to Ryan Schurtz, Aaron Wheeler, Kristina Vitucci, and Rich Feinberg, who helped collect data for the present study. We also thank April Jones and Erin Greenbaum for their assistance in data coding.

AUTHOR BIOGRAPHIES

Justin Buckingham is an associate professor of psychology and director of the experimental psychology master’s program at Towson University. His research interests include social comparison and self-evaluation. E-mail: jbucking@towson.edu

Lavonia Smith LeBeau is a senior research associate at Human Services Research Institute conducting research to evaluate and monitor federally funded substance abuse prevention programs. E-mail: llebeau@hsri.org

William Klein is adjunct associate professor of psychology at the University of Pittsburgh and Associate Director of the Behavioral Research Program at the National Cancer Institute/NIH. His research interests are in the areas of optimism, risk perception, social comparison, and self-judgment. E-mail: kleinwm@mail.nih.gov