



City Research Online

City St George's, University of London

Citation: Luffarelli, J., Gonçalves, D. & Stamatogiannakis, A. (2015). When Feedback Interventions Backfire: Why Higher Performance Feedback May Result in Lower Self-Perceived Competence and Satisfaction with Performance. *Human Resource Management*, 55(4), pp. 591-614. doi: 10.1002/hrm.21739

This is the submitted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/13943/>

Link to published version: <https://doi.org/10.1002/hrm.21739>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

**When Feedback Interventions Backfire:
Why Higher Performance Feedback May Result in Lower Self-Perceived Competence and
Satisfaction with Performance**

Jonathan Luffarelli, Dilney Gonçalves, and Antonios Stamatogiannakis

Abstract

In relative performance evaluation systems, appraisers may choose to adopt stricter or laxer evaluation criteria. When laxer (stricter) criteria are used, higher absolute performance evaluations become easier (harder) to achieve. Thus, each appraisee's absolute performance feedback and the mean of the distribution of absolute performance feedback are shifted upward (downward). Yet, relative performance remains constant. When evaluation outcomes depend solely on relative performance, can the adoption of laxer (stricter) criteria—leading to higher absolute performance feedback but no change in relative performance—influence appraisees' satisfaction with performance? Despite the ubiquity of such systems in organizations, research has not addressed this question. This article points to an important gap between practitioners' beliefs and research findings. We show that while most appraisers believe that higher absolute performance feedback will automatically result in more satisfaction with performance, the opposite may also happen. Specifically, we find that appraisees with a stronger (weaker) chronic or contextual need to engage in social comparison are more satisfied with lower (vs. higher) absolute performance feedback. Overall, we demonstrate why and how feedback interventions in relative performance evaluation systems may backfire, and suggest a set of practical guidelines for maximizing appraisees' satisfaction with performance in such systems.

Keywords: feedback, forced rankings, performance assessment, performance management, satisfaction, social comparison

Corresponding author: Jonathan Luffarelli, Department of Marketing at IE Business School – IE University, Maria de Molina 12, Planta Baja; Madrid, Spain 28006, Phone: +34 915689733, Fax: +34 915689710, jluffarelli.phd2015@student.ie.edu.

Introduction

A manager was reviewing the performance of her subordinates. In her company—as in many others—employees were *only* evaluated according to the ranking of their performance ratings compared to those of the other employees in the department (i.e., curved performance evaluation system). Since in such relative evaluation systems the absolute value of the ratings is inconsequential to employees, the manager decided to adopt less stringent evaluation criteria to shift the entire distribution of performance ratings upwards, without changing the relative ranking—and thus the relative performance evaluation—of each employee. Her intuition was that if anything, higher absolute performance feedback (keeping relative ranking constant) should increase employees' self-perception of their competence level, and so improve their satisfaction with performance. Little did she expect that soon after handing out the appraisals, employees at all performance-rating levels would start expressing their discontent. Puzzled by the unexpected outcome of her generous, but also objectively inconsequential evaluation scheme, the manager started wondering why her feedback strategy may have backfired.

In this paper, we provide an answer to her question and others like it. Specifically, under what circumstances can similar, seemingly generous and inconsequential feedback interventions backfire? When evaluation outcomes depend solely on relative feedback, could absolute feedback have a tangible effect on self-perceived competence and satisfaction with performance? And, despite the fact that, in this instance, relative performance did not change, what factors might explain why and how higher absolute feedback led to lower satisfaction in a strictly relative evaluation system? The answers to these questions could have numerous practical applications for HR managers and important implications for a wide array of research areas.

Relative appraisal systems are ubiquitous in organizations (Backes-Gellner & Pull, 2013) and companies such as Yahoo, AIG, Cisco, and Amazon, use curved rankings for employees'

performance assessments (Brustein, 2013; Ovide & Feintzeig, 2013). Yet research on both absolute and relative performance feedback has typically focused on non-relative evaluation settings (Harris & Smith, 2005; Klein, 1997; Moore & Klein, 2008). Therefore, our knowledge of relative performance appraisals is rather limited (Goffin, Jelley, Powell, & Johnston, 2009; Schleicher, Bull, & Green, 2009), and the effects of feedback interventions—similar to those in the anecdote above—are not well understood (Atwater, Brett, & Charles, 2007; Atwater, Waldman, & Brett, 2002; Kluger & DeNisi, 1996; 1998). It is still unknown whether individuals use absolute performance information in settings in which evaluation is strictly relative (e.g., curve-based evaluations), and whether self-perceived competence and satisfaction with performance are sensitive to absolute performance feedback in such settings.

By improving our understanding of the effect of absolute performance feedback in relative appraisal systems and by helping practitioners understand the circumstances under which feedback interventions may backfire, our paper makes important contributions to the literature on performance appraisals and feedback delivery (Biron, Farndale, & Paauwe, 2011; Bracken, Timmreck, Fleenor, & Summers, 2001; Farndale & Kelliher, 2013; Miller, 2001; Pichler, 2012). Precisely, we expose a new and important gap between practitioners' lay beliefs and research findings (Rynes, Colbert, & Brown, 2002). We demonstrate how and why, contrary to the commonly held belief of those making (i.e., appraisers) and those receiving (i.e., appraisees) evaluations, higher (lower) absolute performance feedback can make appraisees who are evaluated relative to each other feel less (more) satisfied with performance. Our results point toward the important moderating role of social comparison orientation and the mediating role of self-perceived competence in the relationship between absolute performance evaluation and satisfaction with performance. Building on these results, we provide a set of practical guidelines for maximizing appraisees' satisfaction with performance in relative appraisal systems.

The rest of this paper is structured as follows. First, we discuss the features of performance feedback in relative appraisal systems, the relationship between performance feedback and satisfaction with performance, and how this relationship may be moderated by social comparison orientation and mediated by self-perceived competence. In the process, we develop a set of three hypotheses. We then present four studies that establish appraisers' and appraisees' lay theories about performance feedback in relative evaluation settings, and that examine whether and why, contrary to those lay theories, higher (lower) feedback may result in employees' lower (higher) self-perceived competence and satisfaction with performance. We conclude by discussing the theoretical and managerial contributions, as well as some limitations and future research directions.

Performance Feedback in Relative Performance Appraisal Systems

CEOs and other executives (Gong, Li, & Shin, 2011), lower-level managers (Goffin et al., 2009), and even students (Goubeaud, 2010) are often evaluated in relative ways. When individuals are evaluated relatively, each appraisee typically has access to three pieces of information: *absolute performance feedback* (i.e., individual performance), the *distribution mean of everyone's absolute performance feedback* (i.e., average of everyone's performance), and *relative performance feedback* (i.e., standing compared to others). To illustrate using our introductory example, employees under the manager's supervision have access to their own performance rating (absolute performance feedback, e.g., 8 out of 10), the average performance rating of the employees working in the department (average of the absolute performance feedback of all the employees, e.g., 7 out of 10), and their quantile ranking (relative performance, e.g., 40th percentile). Importantly, the first two pieces of information are absolute in nature (i.e., the employee's score is 8/10; the department's average is 7/10); that is, they do not force a comparison with others. However, the third piece of information is relative (i.e., the employee's score is better than 60% of his/her peers) and does force a comparison with others.

Interestingly, in relative performance appraisal settings, appraisers may choose to adopt stricter or laxer evaluation criteria without affecting appraisees' relative performance. Stricter (laxer) evaluation criteria should make higher absolute performance evaluations harder (easier) to achieve for everyone. Thus, when stricter (laxer) evaluation criteria are used, each individual's *absolute performance feedback*, and therefore the *distribution mean of absolute performance feedback*, should be shifted downward (upward). Yet, *relative performance* (i.e., standing) would remain constant (Figure 1). When *only relative performance matters* for the evaluation outcomes (e.g., pay, career progression, etc.), would using stricter versus laxer evaluation criteria, and thus obtaining higher versus lower absolute performance evaluations alter (positively or negatively) appraisees' satisfaction with performance? We now review the relevant literature and develop hypotheses that answer this question.

Insert figure 1 here

Absolute and Relative Performance Feedback

In 1954, Festinger hypothesized that “when an objective, non-social basis for the evaluation of one's ability or opinion is readily available, persons will not evaluate their opinions or abilities by comparison with others.” Since then, research has examined the effects of absolute versus relative performance feedback. While some studies have provided support for this hypothesis (Harris & Smith, 2005), others have found that even when absolute information is available, people tend to compare themselves to others and to be influenced by relative information (Klein, 1997; Moore & Klein, 2008). For instance, in a between-participants experiment, Klein (1997; study 1) manipulated the absolute and relative probability that participants would cause a car accident in their lifetime. He told participants that they had either a 30% or a 60% chance of causing an

accident, and that this probability was either 20% lower or 20% higher than average. He found that participants' self-evaluation of driving ability was influenced by the relative, and not the absolute, probability of causing an accident. Similarly, Moore and Klein (2008; studies 1 and 2) manipulated individuals' relative performances at a "weight guessing game." In this game, participants were shown 10 pictures of individuals. They were asked to guess the weight of each of the 10 within a range of plus or minus five pounds. Participants were then presented with bogus feedback on their performance. They were told that they had guessed the weight of two or eight individuals correctly, and that this performance was either better than 23% or 77% of other participants. After receiving this feedback, participants were invited to play again with a different set of 10 pictures. This time, they were asked to bet an amount of money that they would achieve a given level of performance. In spite of the presence of absolute performance feedback (two vs. eight correct guesses), results showed that the amount participants were prepared to bet was sensitive to both absolute and relative performance feedback.

These important findings were typically made by examining simultaneously absolute and relative performance feedback and in settings for which the *a priori* importance of either type of performance feedback was unclear. For example, in Klein's (1997) study, it is *a priori* unclear if driving ability should be judged according to the absolute or the relative probability of causing an accident. We complement these findings by examining the effect of absolute feedback in settings where relative evaluation is *a priori* of clear importance. For example, monetary winnings at a NASCAR race (Bothner, Kang, & Stuart, 2007), medals at the Olympics (Medvec, Madey, & Gilovich, 1995), or employee compensation in many firms (Backes-Gellner & Pull, 2013) are clearly and strictly based on rankings (i.e., relative performance). For settings in which relative performance is clearly the only thing that matters for the evaluation outcomes, would appraisees' self-evaluation and satisfaction with performance still be influenced by absolute performance

feedback? And, if absolute performance feedback matters, when would lower or higher absolute performance feedback result in lower or higher self-evaluation and satisfaction with performance?

We propose that even when *only* relative performance matters, absolute performance feedback still influences appraisees' satisfaction with performance. In addition, we propose that whether lower or higher absolute performance feedback causes lower or higher satisfaction with performance depends on appraisee's social comparison orientation, which determines the extent to which our own performance versus the performance of others is used to judge self-competence. In the next sections, we formally develop the reasoning behind these propositions and subsequently test them empirically.

The Relationship Between Performance Feedback and Satisfaction

Is higher performance feedback associated with higher or lower satisfaction? First, research argues that higher (actual or perceived) performance feedback may result in intrinsic rewards from which individuals derive contentment (i.e., satisfaction with that level of performance) and thereby potentially cause higher job satisfaction (Judge, Thoreson, Bono, & Patton, 2001). Second, research about both the broader concept of job satisfaction and the narrower one of satisfaction with performance supports this argument. Regarding job satisfaction, conventional wisdom, as well as empirical evidence (Christen, Iyer, & Soberman, 2006; MacKenzie, Podsakoff, & Ahearne, 1998), suggests that higher (lower) performance or feedback might result in higher (lower) job satisfaction. Regarding the specific relationship between performance feedback and satisfaction with performance (which is the focus of this research), the limited existing evidence points in the same direction. For instance, in a sample of U. S. Air Force mechanics, supervisors' performance feedback had a significant and positive effect on subordinates' satisfaction with performance (Van Scotter, 2000). We hypothesize that the intuition of both appraisers and appraisees will align with

this evidence; that is, they will believe that higher (lower) absolute performance feedback will lead to superior (inferior) satisfaction with performance. More formally:

Hypothesis 1: Appraisers (H1a) and appraisees (H1b) believe that giving and receiving (respectively) higher (lower) absolute performance feedback will lead to superior (inferior) satisfaction with performance.

On the other hand, the literature also suggests that self-evaluation and satisfaction may be affected by the use of different reference points (Epstude & Roese, 2008; Jacowitz & Kahneman, 1995; Johnson, Ilies, & Boles, 2012; Mussweiler & Strack, 2000), such as expectations and goals (Larrick, Heath, & Wu, 2009) or counterfactuals (Markman, McMullen, & Elizaga, 2008). Interestingly, the evidence indicates that the use of certain reference points could lead individuals to be less (more) satisfied with higher (lower) actual or false performance feedback (Alicke, Zell, & Bloom, 2010). For example, Iyengar, Wells, and Schwartz (2006) found that although satisficers tended to do objectively worse than maximizers in a job search, they were more satisfied with their level of performance. They determined that this effect occurred because unlike satisficers, maximizers sought and used the best possible job offer as the reference point to evaluate their current job conditions. In the same way, Medvec et al. (1995) and Medvec & Savitsky (1997) found that bronze medalists were more satisfied with their performances than silver medalists in sports competitions as the salience of different reference points resulted in different counterfactual thoughts. In particular, they determined that bronze medalists used the prospect of “not winning a medal” as a reference while silver medalists used the prospect of “winning the gold medal”.

Building on social comparison theory, research suggests that the performance of others is an additional and important reference point that people may use in their judgments (Garcia & Tor, 2007; Heslin, 2005; Lane & Gibbons, 2007). In this paper, we propose that the extent to which the

performance of others (vs. own performance) is used to judge self-competence will determine whether lower or higher absolute performance feedback can cause higher satisfaction with performance. We also propose that the use of either one of these two available absolute pieces of information (i.e., others' vs. own performance) depends on appraisees' chronic or contextually activated social comparison orientation. We next describe how the extent to which appraisees engage in social comparison may affect the use of one's own versus others' average absolute feedback for self-competence evaluation, leading to distinct effects on self-perceived competence and satisfaction with performance.

When Superior Absolute Performance Feedback May Backfire

Social comparison refers to the tendency of individuals to define their self-concept in relation to their social environment, including how they stand in comparison to others (Buunk & Gibbons, 2007; Festinger, 1954; Suls, Martin, & Wheeler, 2002). Although everyone essentially engages in social comparison, individuals with a high social comparison orientation feel a much stronger drive to engage in relative self-appraisal than individuals with a low orientation (Gibbons & Buunk, 1999). This drive can be the result of chronic traits (Bogart, Benotsch, & Pavlovic, 2004; Locke & Nekich, 2000; Van der Zee, Oldersma, Buunk, & Bos, 1998), or influenced by contextual factors (Goodman & Haisley, 2007; Greenberg, Ashton-James, & Ashkanasy, 2007; Mussweiler & Epstude, 2009; Mussweiler, Rüter, & Epstude, 2004). According to social comparison theory, individuals' social comparison processes moderate the selection of information about others when making self-evaluative and satisfaction judgments (Gibbons & Buunk, 1999; Mussweiler 2003). Specifically, in the context of relative performance appraisals, the need to use information about others should be of fairly low importance to appraisees who have a low social comparison orientation. However this need should be of fairly high importance to those who have a high

orientation. Thus, when making judgments about themselves, for example, how competent they are in their jobs, appraisees low (high) in social comparison orientation should rely relatively more on non-comparative (comparative) information.

In particular, when it comes to evaluating their level of self-competence (i.e., how good they are at a given task), appraisees low in social comparison orientation should be relatively more likely to use their own performance ratings (vs. the average performance rating of their peers) as an input. In other words, because they have a fairly weak tendency to compare themselves to others, their own performance level can be a sufficient indicator of how good they are. Given that high (low) absolute performance feedback signals that one has performed well (poorly), appraisees with a lower inclination toward social comparison should feel more (less) competent when everyone's absolute performance feedback—including their own—is higher (lower). Put differently, in relative evaluation systems, appraisees who are more likely to use their own performance level to determine their level of self-competence (i.e., those low in social comparison orientation) should feel more competent when receiving an absolute performance rating of 8, rather than an absolute performance rating of 7 (out of 10).

Conversely, for appraisees who have a tendency to compare themselves with others, their own performance feedback may not be a sufficient indicator of self-competence. Instead, these appraisees should be relatively more likely to use the performance level of a relevant group (e.g., colleagues or class members) as a reference point to assess how good they are. In other words, because their self-competence is determined by how favorably their own performance compares with the performance of their peers, they will use cues about others' performance as comparison standards. The less (more) favorable the comparison, the less (more) competent they will feel. Less (more) favorable comparisons should occur when the average absolute performance feedback that serves as a reference point is higher (lower). In our case, however, higher (lower) average absolute

performance feedback also means higher (lower) individual performance feedback (Figure 1). As an example, adopting laxer (vs. stricter) evaluation criteria may create a situation in which a given appraisee receives an absolute performance rating of 9 (vs. 7) out of 10 and in which the average performance rating is 8 (vs. 6). In both cases, appraisees with a strong social comparison orientation will need to compare their own absolute performance with an external standard, thus focusing on the average. However, since the average is 8 with laxer evaluation criteria, but only 6 with stricter criteria, this comparison may be less favorable in the first case, as the comparison standard (i.e., the average) is higher. As a result, a surprising effect may arise: appraisees with a stronger inclination toward social comparison may feel more (less) competent when everyone's absolute feedback—including their own—is lower (higher). Formally, we propose the following hypothesis (Figure 2):

Hypothesis 2: When appraisees' tendency to engage in social comparison is high (low), they will feel more self-competent with lower (higher) absolute performance feedback.

Further, existing research in a wide array of domains has demonstrated that the extent to which appraisees feel competent and self-efficacious can be a powerful antecedent of job satisfaction (Judge & Bono, 2001). For example, research in entrepreneurship (Bradley & Roberts 2004), psychology (Judge, Bono, Erez, & Locke, 2005), and educational psychology (Caprara, Barbaranelli, Steca, & Malone, 2006) has respectively shown that entrepreneurs', employees' and teachers' self-perceived level of competence predicted satisfaction. Building on the aforementioned studies on job and performance satisfaction, we propose the following hypothesis (Figure 2):

Hypothesis 3: The interaction effect between absolute performance feedback and social comparison orientation on self-perceived competence will carry over to satisfaction with performance.

Insert figure 2 here

Overview of the Studies and Results

Our hypotheses are tested in four studies that combine a variety of methods. In all the studies, it was made explicit that only relative performance mattered. In the pilot study, we directly ask appraisers to indicate whether higher or lower absolute performance feedback would lead appraisees to be more satisfied with performance. We find that appraisers have the intuition that higher (lower) absolute performance feedback would translate into higher (lower) appraisee satisfaction with performance (H1a). Experiment 1 consists of two parts. First, complementing the pilot study, we directly ask appraisees (who are typically evaluated relatively) to indicate whether higher or lower absolute performance feedback would make them more satisfied with performance. Like appraisers, we find that they believe they would be more (less) satisfied with performance with higher (lower) absolute performance feedback (H1b). However, contrary to this shared belief, when the same appraisees are asked to estimate their satisfaction with performance in a domain with which they are very familiar, they indicate greater satisfaction with lower (vs. higher) absolute performance feedback (H3). Experiment 2 provides further evidence that lower (vs. higher) absolute performance feedback may result in greater satisfaction with performance. First, we directly measure social comparison orientation and find that appraisees with a high (low) chronic tendency to engage in social comparison are more satisfied with their level of performance when absolute performance feedback is low (high) (H3). Second, we demonstrate the generalizability and robustness of these findings by examining alternative explanations, and by using different levels of performance and evaluation settings (i.e., sports competition, loyalty program, and classroom). Experiment 3 replicates our findings and lifts the veil on the psychological mechanism behind our results (H2; H3). Here, we manipulate social comparison orientation. Results demonstrate that

social comparison orientation moderates the relationship between absolute performance feedback and self-perceived competence (H2), while self-perceived competence mediates the relationship between absolute performance feedback and satisfaction with performance (H3). The studies are described in detail below, and a concise overview is given in Table I.

Insert Table I here

Pilot Study: Appraisers' Lay Beliefs About Satisfaction with Performance in Relative Appraisal Systems

Method

Objective

Common wisdom holds that higher (lower) performance feedback will result in higher (lower) satisfaction with performance. However, as we discussed previously, the evidence suggests that there are contexts such as job hunting (Iyengar et al. 2006) or the Olympic Games (Medvec et al. 1995; Medvec & Savitsky 1997) in which satisfaction is higher (lower) with lower (higher) performance. Experienced appraisers may be aware of these findings, or may have personally experienced patterns that differ greatly from what common knowledge suggests. Therefore, the purpose of the pilot study was to examine whether experienced appraisers believed that appraisees would experience more satisfaction with performance with higher (vs. lower) absolute performance feedback, even when the feedback is objectively inconsequential for outcomes.

Participants

Due to their continuous involvement in performance evaluations, university professors were used as expert appraisers. This pilot study was administered to two different samples. The first sample consisted of 50 professors at a major Spanish business school (with a student and faculty

body with a very international profile) in which all performance feedback is relative (i.e., determined by percentile ranks). The second sample was recruited from an international academic society and consisted of 37 professors affiliated to diverse universities, schools, and departments.¹ Twenty-two of them indicated that they systematically use curved grading, and 15 that they never, rarely, or very rarely do so (“how frequently do you evaluate your students based on a curved grading system?” 1 = never, 7 = always). Because these questionnaires were administered to a very specific target population, we did not ask for information such as school affiliation, age, gender, or nationality in order to guarantee anonymity.¹ Due to their expertise in judgment and decision-making, participants in the second sample provide a very conservative sample for the purposes of this study. If anyone would make predictions that run counter to common wisdom, we would expect this sample to do so.

Procedure

Participants in both samples were asked to imagine that they just had administered an exam that would be scored using a described curved evaluation system. They were then asked to indicate whether they believed that a student would be more satisfied with performance when obtaining a score of 7 out of 10 when the average of the class is 8 or when obtaining a score of 4 out of 10 when the average of the class is 5, while ranking in the 40th percentile in both cases.

Results and Discussion

Because answers were not significantly different across the two samples or across appraisers who were familiar or unfamiliar with relative performance feedback systems, we report the pooled results. Results indicated that while only 15% ($n = 13$) of the appraisers in our sample believed that lower absolute performance feedback would result in higher satisfaction with performance, 85% ($n = 74$) believed that higher absolute performance feedback would result in higher satisfaction with performance ($X^2(1, N = 87) = 42.77, p \leq 0.001$). Therefore, in support of hypothesis 1a, regardless

of appraisers' familiarity with relative performance feedback systems, appraisers anticipate that appraisees will be more (less) satisfied with higher (lower) absolute performance feedback.

In what follows, we show why and when these beliefs may be inaccurate. We report three experiments that suggest that appraisers might unknowingly and systematically engage in behaviors that do not maximize, or even reduce, appraisees' satisfaction with performance.

Experiment 1: Higher (Lower) Absolute Performance Feedback Can Result in Lower (Higher) Satisfaction with Performance

Method

Objective

The aim of the first experiment was twofold. First, we wanted to examine whether appraisees who are used to being evaluated in a relative manner have accurate lay theories about the level of absolute feedback that would make them more satisfied with performance; or if, like experienced appraisers, they tend to believe that higher (lower) absolute feedback will result in higher (lower) satisfaction with performance. Second, we wanted to demonstrate that in a relative performance evaluation system, higher (lower) feedback can result in lower (higher) satisfaction with performance.

Participants

The participants were 54 postgraduate students at a major business school in which the outcomes of all evaluations are determined by rankings (the same business school as the 50 postgraduate professors in the first pilot study sample). They were offered class credit in exchange for their participation. All were in their final academic term, and thus very familiar with relative performance appraisal systems. By the time of the experiment, they had completed more than 20 courses in which they had been evaluated on a curve. Consequently, we expected all these

participants to have a strong tendency to engage in social comparison, given that social comparison orientation can be activated by context (Mussweiler & Epstude, 2009; Mussweiler et al., 2004), that postgraduate business students tend to be competitive (McCabe, Butterfield, & Treviño, 2006; McCabe & Treviño, 1995), and that closeness between organizational members can promote relative comparisons (Heslin, 2005). In addition, they were in a class with an average age of 23 and average work experience of one year. Finally, within about six months of the administration of this study, 93% of them had jobs in a variety of sectors, positions, and geographical regions, indicating that their profile was very similar to that of typical young professionals who end up working in organizations that might use relative performance evaluation systems.

Procedure and Measures

There were two parts to experiment 1. In the first, we measured participants' satisfaction with performance with high versus low absolute performance feedback in a relative evaluation setting. In the second, we asked participants to make a direct choice between obtaining either a high or low absolute performance feedback. Whereas rating satisfaction with a given performance outcome would reveal appraisees' reaction to a situation, the choice between two potential performance outcomes should reflect appraisees' lay theories about how each of the two outcomes would affect their satisfaction with performance. This two-part design allows us to probe whether appraisees have accurate lay theories about which outcomes are likely to maximize their satisfaction with performance, as well as what would actually maximize their satisfaction with performance.

In the first part of this experiment, participants were randomly assigned to one of the two experimental conditions of this single-factor experiment: high versus low absolute performance feedback (individual and average absolute performance feedback). Participants were asked to assess the satisfaction of a postgraduate student in their program with performance feedback using a

seven-point scale (1 = very dissatisfied, 7 = very satisfied with performance). Just like after a real exam, participants were given three pieces of information: the score (i.e., absolute performance feedback), their relative standing in the class (i.e., relative performance), and the average score of the class (i.e., average absolute performance feedback). In the low absolute performance feedback condition ($n = 31$), the score was 4 out of 10, and the average score of the class was 5 out of 10. In the high absolute performance feedback condition ($n = 23$), the score was 7 out of 10, and the average score of the class was 8. Relative ranking was kept constant across conditions (i.e., the student was ranked in the 40th percentile).

Based on our theorizing (H3), we expected participants to give higher performance satisfaction ratings for the low (vs. high) absolute performance feedback condition. Our pilot study had already shown that this finding would be at odds with appraisers' lay intuitions. In order to investigate whether it would also be at odds with appraisees' lay intuitions, we carried out a follow-up survey that was administered to 70 postgraduate students drawn from the same population (including the 54 who participated in the first part of experiment 1). Participants were asked to make a direct choice between two absolute performance feedback situations: either obtaining a score of 4 out of 10 when the average is 5, or of 7 out of 10 when the average is 8, while ranking in the 40th percentile in both cases. This direct choice is the same that the professors in the pilot study had to make.

Results and Discussion

Given our directional hypothesis (H3), we used one-tailed testing (Cho & Abe, 2013). Consistent with our predictions, an independent-samples t-test indicated that participants (who were familiar with relative performance appraisal systems, and presupposed to be high in social comparison orientation) provided higher performance satisfaction ratings in the low absolute performance feedback condition ($M = 2.48$, $SD = 0.85$) than in the high absolute performance

feedback condition ($M = 2.13$, $SD = 0.69$; $t(52) = 1.63$, $p = 0.054$). However, in the follow-up survey when participants made a direct choice, 77% ($n = 54$) preferred receiving a high absolute performance feedback ($\chi^2(1, N = 70) = 20.63$, $p \leq 0.001$). Thus, the results of the first part of this experiment contrast with those of the second part. Specifically, when asked to make a direct choice between a high versus a low absolute performance feedback score, participants seemed to focus only on the magnitude of the two scores, and chose the higher one. Effectively, as the first result shows, they are choosing the outcome that would make them (on average) less satisfied.

An informal debriefing with the participants of the follow-up survey showed that there were two reasons why some of them indicated that, for the same level of relative performance, they would prefer receiving lower (vs. higher) absolute performance feedback. First, they realized that the two options from which they had to choose were objectively equal, and, therefore, a lower absolute performance feedback was just as good as a higher absolute performance feedback. Second, based on their prior experience, they anticipated that comparing their performance with a very high average could make them unhappy, even if the former was high too.

Drawing conclusions from the pilot study and experiment 1, it appears that in relative performance appraisal systems, appraisers and appraisees share a common belief that satisfaction with performance is higher (lower) when receiving higher (lower) absolute performance feedback. This supports hypotheses 1a and 1b. Yet, contrary to this belief and consistent with hypothesis 3, we find that higher (lower) absolute performance feedback might actually result in lower (higher) satisfaction with performance.

Experiment 2: The Moderating Role of Social Comparison Orientation

Method

Objectives

Experiment 2 generalizes our results in four ways. First, as experiment 1 showed the effect of absolute performance feedback on performance satisfaction in a context where social comparison orientation was naturally high, we examine the moderating role of social comparison orientation by measuring it directly. Second, we focus on different contexts where individuals may possibly be evaluated relatively. Third, we seek to replicate our effect using a set of participants that is relatively representative of the U.S. population. Fourth, in order to control for potential alternative explanations, we include cases where absolute performance feedback was above and below the average absolute performance feedback score.

Participants and Procedure

One hundred and twenty-six participants ($M_{age} = 34$; 57% female) were recruited for monetary compensation from Amazon Mechanical Turk (Hereafter MTurk), a crowdsourcing website (the benefits and limitations associated with the use of MTurk are discussed in the limitation section of this article). Participants were randomly assigned to a $2 \times 2 \times 3$ between-participants factorial design experiment. Fifteen were excluded for not following the instructions (Oppenheimer, Meyvis, & Davidenko, 2009). We manipulated absolute performance feedback (high vs. low individual and average absolute performance feedback), relative performance (above vs. below average absolute performance feedback), and context replicates² (classroom vs. sports competition vs. loyalty program). A full presentation of the manipulations is provided in the appendix. Similar to experiment 1, the high versus low absolute performance feedback manipulation consisted of having one condition with low and another with high absolute performance feedback (e.g., average absolute performance feedback of 16,300 [6,600] points in the

high [low] absolute performance feedback condition of the loyalty program context replicate). The above versus below average relative performance conditions were created by adding or removing a constant percentage from the average absolute performance feedback (e.g., plus/minus 25% of points from the average of 16,300 [6,600] points in the high [low] absolute performance feedback condition of the loyalty program context replicate). In every condition, participants were explicitly told that only relative performance mattered in determining outcomes. For generalizability, we used different ranking percentiles for each context replicate (respectively 55th and 45th, 60th and 40th, and 70th and 30th for the classroom, sports competition, and loyalty program replicates).

Measures

Satisfaction with performance was measured with two items, “How satisfied is [character’s name] with his/her performance?” and “How happy is [character’s name] with his/her performance?” (1 = not satisfied at all/very unhappy, 7 = very satisfied/very happy with performance; $\alpha = 0.99$) presented in counterbalanced order. We also measured social comparison orientation using a validated 11-item scale (cf., Gibbons & Buunk, 1999). Each item was measured on a five-point scale (1 = I disagree strongly, 5 = I agree strongly).

Analysis

The dependent variable was standardized. There was no significant effect of question order. The items of the comparison orientation scale were also averaged ($\alpha = 0.88$) and also standardized. Given the hypothesis (H3) of an interaction between a continuous variable (i.e., social comparison orientation) and an experimental factor (i.e., absolute performance feedback), we performed a linear regression (Aiken & West, 1991; Fitzsimons, 2008; Irwin & McClelland, 2001). Precisely, we regressed satisfaction with performance on absolute feedback (low vs. high), social comparison orientation, absolute feedback $\times$ social comparison orientation interaction, relative performance (below vs. above the mean), and two context replicates dummies.

Results and Discussion

We find that participants' relative performance (above vs. below the mean) had a very strong main effect on participants' satisfaction with performance ($B = 2.14$; $t(104) = 8.63$; $p \leq 0.001$). This result is to be interpreted as a manipulation check, indicating that participants understood that only relative performance mattered. We also find that the main effect of social comparison orientation on satisfaction with performance was marginally significant ($B = 0.33$; $t(104) = 1.73$; $p = 0.086$) and that the main effect of absolute performance feedback was not statistically significant ($B = 0.02$; $t(104) = 0.08$; $p = 0.937$). Finally, the context replicates dummies were statistically significant both for the sports scenario ($B = -0.97$; $t(104) = -3.24$; $p \leq 0.001$) and for the classroom scenario ($B = -1.60$; $t(104) = -5.30$; $p = 0.002$). These effects simply indicate that different scenarios produced different satisfaction with performance ratings.

More interestingly, our analysis revealed a significant absolute performance feedback $\times$ social comparison orientation interaction ($B = -0.60$; $t(104) = -2.37$; $p = 0.020$). We tested the effect of low versus high absolute performance feedback at low and high levels of social comparison orientation by a spotlight analysis (Aiken & West, 1991; Fitzsimons, 2008; Irwin & McClelland, 2001) at two standard deviations above (high social comparison orientation) and below (low social comparison orientation) the mean (e.g., Tepper, Duffy, Henle, & Lambert, 2006; Vess, Arndt, Cox, Routledge, & Goldenberg, 2009). Supporting hypothesis 3, we found that appraisees high in social comparison orientation were less satisfied with their performance with high (vs. low) absolute performance feedback ($B = -1.19$; $t(104) = -2.11$; $p = 0.037$). Conversely, appraisees low in social comparison orientation were more satisfied with their performance with high (vs. low) feedback ($B = 1.23$; $t(104) = 2.15$; $p = 0.034$). The results of this analysis are reported in Figure 3.

Insert figure 3 here

Given our full-factorial experimental design, we tested an additional model to examine if relative performance (above vs. below the average) interacted with either absolute feedback (high vs. low) or social comparison orientation. Specifically, we estimated an alternative model adding the following interaction terms to the model presented above: relative performance $\times$ absolute performance feedback, relative performance $\times$ social comparison orientation, and relative performance $\times$ absolute performance feedback $\times$ social comparison orientation. The results showed no statistically significant relative performance $\times$ absolute performance interaction ($B = -0.19$; $t(101) = -0.05$; $p = 0.702$) or relative performance $\times$ social comparison orientation interaction ($B = -0.18$; $t(101) = -0.08$; $p = 0.649$) effects. The three-way interaction was not significant either ($B = 0.16$; $t(101) = 0.01$; $p = 0.977$). Finally, adding the relative performance interactions did not improve the model fit ($R^2\text{-change} = 0.003$; $F(3, 101) = 0.20$; $p = 0.895$).

An alternative hypothesis could be that even when only relative performance matters for the evaluation outcomes, appraisees with a high social comparison orientation would consider only their relative standing compared to others, while appraisees with a low social comparison orientation would consider only their absolute performance feedback. If this were the case, low social comparison orientation appraisees should remain unaffected by their relative positions around the average of everyone's performance. Providing support for our hypotheses, the results show that this is not the case. The fact that the significant absolute performance feedback $\times$ social comparison orientation interaction is not affected by participants' position (above or below) relative to the average performance indicates that in relative evaluation systems, both high and low social comparison orientation appraisees attend in a similar way to relative information (information based on which they are ultimately assessed). However, supporting our predictions, this significant

absolute performance feedback $\times$ social comparison orientation interaction also indicates that appraisees do not attend to absolute performance information in a similar way. Specifically, while appraisees with a low social comparison orientation seem to attend more to absolute cues relating to their own individual performances, those with a high orientation seem to attend more to absolute cues relating to others' performances. Overall, these results provide additional support for our predictions, and leave little room for alternative explanations. More precisely, the fact that relative performance (below vs. above the mean) did not interact with any other variables makes our results hard to reconcile with any alternative explanation based on an individual's position relative to the average. One such explanation could be that when participants find themselves below the average, they are more satisfied with their performance when the average and their own absolute performance feedback are both low, because this gives a convenient excuse for the bad absolute performance feedback (e.g., the task was hard, so I did poorly).

Finally, a comparison of relative effect sizes shows that the effect size of being above versus below the average is much bigger than the effect size of the absolute performance feedback $\times$ social comparison orientation interaction. In other words, our data suggest that much of the observed variance in participants' satisfaction with performance can be attributed to their relative position around the average absolute performance feedback. This is not surprising since our research takes place in relative settings and we made it clear to participants that the final outcomes depended only on relative performance. Yet, even after capturing the variance that should "normatively" affect satisfaction with performance, the social comparison orientation $\times$ absolute performance feedback interaction still remained significant. This indicates that for the same position around the average, absolute performance feedback can significantly impact satisfaction with performance.

Experiment 3: The Mediating Role of Self-Perceived Competence

Method

Objectives

Results from the experiments 1 and 2 demonstrate that for the same level of relative performance, appraisees with a greater tendency to engage in social comparison are more (less) satisfied with performance with lower (higher) absolute performance feedback. In experiment 3, we lift the veil on the psychological mechanism that would explain why inferior (superior) absolute performance feedback can lead to higher (lower) satisfaction with performance. In addition, we complement previous findings by experimentally manipulating (rather than measuring) social comparison orientation.

Participants and Procedure

Two hundred and fifty-six participants ($M_{\text{age}} = 33$; 54% female) were recruited for monetary compensation from MTurk. They were randomly assigned to a 2×2 between-participants factorial design with absolute performance feedback (high vs. low) and social comparison orientation (primed vs. control) as between-participants factors. Since in experiment 2 there was only a main effect of relative performance (i.e., above vs. below average absolute performance feedback) but no interactions, and since the participants in experiment 1 were all positioned below the average absolute performance feedback, all participants in experiment 3 were positioned above the average absolute performance feedback.

Manipulations

To manipulate social comparison orientation, we followed a procedural priming manipulation adapted from Mussweiler and Epstude (2009). Participants were presented with a picture and a verbal profile of an individual of the same gender. The verbal profile contained geographic, demographic, psychographic, and behavioral information. Participants were asked either to compare themselves with (social comparison priming condition) or simply describe (control condition) the person pictured.

Subsequent to the priming manipulation, participants were asked to assess the satisfaction with performance of a student with the grade obtained in a course. They were randomly assigned to one of two absolute performance feedback conditions (high vs. low). In the high absolute performance feedback condition ($n = 131$), the score of a target student was 7.5 out of 10, and the average score of the class was 6.5. In the low absolute performance feedback condition ($n = 125$), the score of a target student was 5, and the average score of the class was 4. Importantly, relative ranking was kept constant across conditions (i.e., the student ranked in the 75th percentile).

Measures

Similar to experiment 2, satisfaction with performance was measured with two items (1 = not satisfied at all/very unhappy, 7 = very satisfied/very happy with performance; $\alpha = 0.89$) presented in counterbalanced order. We also measured perceived competence (1 = very bad, 7 = very good), perceived task difficulty (1 = very difficult, 7 = very easy), and performance feedback expectations (percentage of appraisees who should have been outperformed based on absolute performance level).

Results and Discussion

The dependent variable was standardized. First, we subjected our measure of satisfaction with performance to a simple moderation analysis with absolute performance feedback (low vs. high), social comparison orientation (control vs. primed), and their interaction as independent variables. This regression showed significant main effects of absolute performance feedback ($B = 0.50$; $t(252) = 2.89$; $p = 0.004$) and social comparison orientation ($B = 0.45$; $t(252) = 2.88$; $p = 0.005$) on satisfaction with performance. These main effects were also qualified by a significant and negative absolute performance feedback $\times$ social comparison orientation interaction ($B = -0.51$; $t(252) = -2.92$; $p = 0.004$). This interaction replicates the results of experiment 2 and suggests that participants primed to be more socially comparative reported significantly lower satisfaction with

performance when they were presented with higher (vs. lower) absolute performance feedback. The results of this moderation analysis are reported graphically in Figure 4.

Insert figure 4 here

To test whether the observed performance feedback $\times$ social comparison orientation interaction on satisfaction with performance was mediated by self-perceived competence, we subjected our measure of satisfaction to a moderated mediation analysis (Preacher, Rucker, & Hayes, 2007; Zhao, Lynch, & Chen, 2010) with absolute performance feedback (low vs. high), social comparison orientation (control vs. primed), and their interaction as independent variables, and standardized self-perceived competence as a mediator. Consistent with a moderated mediation, our social comparison orientation manipulation significantly moderated the relationship between absolute performance feedback and self-perceived competence. In turn, self-perceived competence significantly mediated the relationship between absolute performance feedback and satisfaction with performance. The results of this moderated mediation are reported in Table II and Figure 5.

Insert Table II and figure 5 here

Results show a significant main effect of absolute performance feedback ($B = 0.44$; $t(252) = 2.51$; $p = 0.013$; Table II.A), and a significant main effect of social comparison orientation ($B = 0.45$; $t(252) = 2.94$; $p = 0.004$; Table II.A). Both these main effects are qualified by a significant absolute performance feedback $\times$ social comparison orientation interaction ($B = -0.45$; $t(252) = -0.57$; $p = 0.011$; Table II.A) on our mediator, self-perceived competence. This interaction indicates that compared to participants in the control condition, participants who were primed to be

more socially comparative reported significantly lower self-perceived competence with higher absolute performance feedback. In turn, we found that higher self-perceived competence led to higher satisfaction with performance ($B = 0.41$; $t(251) = 7.00$; $p \leq 0.001$; Table II.B). The strength and significance of the indirect effects of absolute performance feedback were further assessed using 5000 bootstrap samples (Preacher et al., 2007; Zhao et al., 2010). The conditional indirect effect of absolute performance feedback on satisfaction with performance was significant and negative for participants who were primed to engage in social comparison (95% CI of $[-0.007, -0.002]$; Table II.C) but was significant and positive (95% CI of $[0.058, 0.330]$; Table II.C) for participants in the control group. These conditional indirect effects indicate that participants who were primed to be more socially comparative felt more dissatisfied with high (vs. low) absolute performance feedback, while participants in the control condition felt more satisfied with high (vs. low) absolute performance feedback. Finally, consistent with a partial moderated mediation, the absolute performance feedback $\times$ social comparison orientation interaction ($B = -0.33$; $t(251) = -2.03$; $p = 0.043$; Table II.B) was also directly and significantly correlated with satisfaction with performance.

With a view toward comprehensiveness, and since expectation (Ilgen, 1971) and perceived task difficulty (Ewen, 1973) are important determinants of satisfaction, we tested alternative moderated mediation models with these two variables as mediating variables. Neither performance expectation nor perceived task difficulty was found to be a significant mediator, allowing little space for these potential alternative mediators.

General Discussion

Overview of the Findings

Comparative evaluation systems are common in organizations (Backes-Gellner & Pull, 2013; Goffin et al., 2009), universities and colleges (Goubeaud, 2010), sports competitions (Medvec & Savitsky, 1997), and other occasions. Although in these settings it is often the appraiser's intention to base evaluations on relative standings only, we demonstrate that absolute performance feedback also impacts appraisees' satisfaction with performance. Specifically, we find that contrary to common wisdom and practice, higher (lower) feedback may result in lower (higher) self-perceived competence and satisfaction with performance in relative evaluation systems. More precisely, we propose and demonstrate that chronic or contextual social comparison orientation moderates the relationship between absolute performance feedback and self-perceived competence, while self-perceived competence mediates the relationship between absolute performance feedback and satisfaction with performance. We find that appraisees who have a low (high) social comparison orientation feel more self-competent and satisfied with performance with higher (lower) absolute performance feedback.

Theoretical Contributions

Understanding the gaps between practitioners' beliefs and research findings is an important inquiry in human resource management (Rynes et al., 2002). Our findings indicate that in relative evaluation systems, both appraisers and appraisees believe that higher (lower) absolute performance feedback will translate into superior (inferior) satisfaction with performance. Yet, our findings also indicate that in such system, higher (lower) absolute performance feedback can often lead to inferior (superior) satisfaction with performance. This unsubstantiated belief may have adverse effects for organizations in industries where appraisees' satisfaction is essential. For instance, student satisfaction is a critical determinant of the quality of alumni networks, reputation,

enrolment, and financial performance of higher education institutions (Alves & Raposo, 2007; 2009; Athiyaman, 1997; Carter, 2009; Henning-Thurau, Langer, & Hansen, 2001). More generally, this unsubstantiated belief may have undesirable effects in organizations that evaluate employees relatively, since satisfaction is linked to employees' physical and mental health (Faragher, Cass, & Cooper, 2005), as well as to individual (Riketta, 2008; Rodwell, Kienzle, & Shadur, 1998; Saari & Judge, 2004), group (Whitman, Van Rooy, & Viswesvaran, 2010), and organizational performance (Edmans, 2012; Zhou, Li, Zhou, & Su, 2008).

To date, only a handful of studies have shown that higher (lower) performance can result in lower (higher) satisfaction. Although these studies have taken place in different contexts, they share some theoretical underpinnings with our research. Specifically, in prior research higher (lower) performance resulted in lower (higher) satisfaction with performance when certain reference points such as counterfactuals (Medvec & Savitsky, 1997; Medvec et al., 1995) or ideals (Iyengar et al., 2006) were used for the purpose of self-evaluation. The counterintuitive phenomenon we demonstrate in our research stems from judging self-competence based on an individual's own performance versus that of others. We establish that appraisees who have a high social comparison orientation (i.e., are more likely to use others' performance as a reference) tend to feel more (less) self-competent and satisfied with their level of performance when the mean of everyone's absolute performance feedback (and, thus, their own absolute performance—Figure 1) is low (high). We also establish that this effect is reversed for appraisees who have a low social comparison orientation (i.e., are more likely to judge self-competence based on their own performance), who tend to feel more (less) self-competent and satisfied with their level of performance when the mean of everyone's absolute performance feedback (and, thus, their own absolute performance) is high (low).

The crucial role that social comparison plays in this process places our study in a research stream that has demonstrated the importance of social comparisons in organizational settings (Brown, Ferris, Heller, & Keeping, 2007; Conner, 2003; Greenberg et al., 2007; Harris, Anseel, & Lievens, 2008; Hu & Liden 2013). Precisely, we demonstrate that in contexts in which individuals are evaluated relatively, social comparison orientation can be a dispositional and situational moderator of predictors of satisfaction with performance. Yet, the performance-satisfaction literature (Iaffaldano & Muchinsky, 1985; Judge et al., 2001; Petty, McGee & Cavender, 1984; Saari & Judge, 2004; Schleicher, Watt, & Greguras, 2004) has ignored the potential role of social comparison orientation. In light of our results, future performance-satisfaction studies might want to measure or control for participants' social comparison orientation.

Another important theoretical contribution of our paper lies in the examination of the effects of absolute performance feedback on self-evaluation and satisfaction with performance. We complement prior research by manipulating the effects of absolute performance feedback in contexts for which only relative performance (which was kept constant in all of our studies but experiment 2 in which it was manipulated but did not interact with other treatments) mattered for the ultimate outcome. To the best of our knowledge, this is the first research to do this. Indeed, past studies have usually manipulated both types of feedback concurrently (Harris & Smith, 2005; Klein, 1997; Moore & Klein, 2008). Past studies have also not examined contexts in which the *a priori* importance of relative over absolute feedback is clearly greater to participants. As such, we provide some support for Festinger's (1954) claim that individuals will tend to take into consideration absolute information even when relative feedback is available. Indeed, the results of our research show that in the presence of relative information, individuals still use absolute performance information to make inferences about their competence. This is extremely surprising, given that we examine contexts in which only relative performance matters. Therefore, managers

would be ill advised to ignore the potential role of absolute performance evaluation in relative evaluation settings.

Finally, this paper also contributes to the literature on relative evaluation systems (Blume, Rubin, & Baldwin, 2013; Garcia & Tor, 2007; Goffin et al., 2009; Schleicher et al., 2009; Scullen, Bergey, & Aiman-Smith, 2005) and performance appraisals and feedback delivery (Biron, Farndale, & Paauwe, 2011; Bracken et al., 2001; Farndale & Kelliher, 2013; Miller, 2001). Our results indicate that in addition to being potentially subjected to contextual influences (Ferris, Munyon, Basik, & Buckley, 2008; Pichler, 2012) and to the features and characteristics of performance evaluations systems (Brown, Hyatt, & Benson, 2010), appraisees' reactions to relative evaluation systems are subject to certain personality traits, such as social comparison orientation. Importantly, our results also demonstrate the potential detrimental impact of relative performance evaluation methods on satisfaction with performance if appraisees' social comparison orientation profiles are not taken into account. As such, they shed some needed light on the conditions under which feedback interventions (Atwater et al., 2002; 2007; Kluger & DeNisi, 1996; 1998) in curved grading settings produce positive and negative effects on appraisees' self-perceived competence and satisfaction with performance.

Managerial Implications

From a practical perspective, we have shown that when appraisees are evaluated relatively, a fine balance—one that would make everyone happy—is hard to achieve. Indeed, appraisees with a high social comparison orientation tend to be more satisfied with higher absolute performance feedback, while those with a low orientation tend to be more satisfied with lower absolute performance. As a strategic remedy, we propose that appraisers should consider the average social comparison orientation profiles of appraisees being evaluated, as well as the contextual factors that make social comparison processes more or less likely. Subsequently, appraisers can decide whether

they should adopt more stringent or lenient evaluation criteria to maximize appraisees' satisfaction with performance. For instance, imagine a sales manager evaluating salespeople on a 100-point scale, and then ranking these scores in order to determine bonus allocation. If those salespeople tend to compare themselves with others a lot (little), obtaining one point should be made harder (easier). For example, one point could correspond to \$1,000 (\$1,500) in sales. Without affecting the actual rank-based outcome (i.e., bonus allocation) of the evaluation (Figure 1), the adoption of more stringent (lenient) evaluation criteria should result in lower (higher) absolute performance feedback, and thus higher satisfaction with performance for salespeople with high (low) social comparison orientation. Obviously, and very importantly, the strategic adoption of stricter versus laxer evaluation criteria must be both ethical and consistent with the current practices of the organization. As such, it should probably be subject to the approval of senior managers and the ethics committee.

If the adoption of stricter versus laxer evaluation criteria is not feasible (for ethical or organizational reasons), our work still offers recommendations for increasing satisfaction with performance. First, as the satisfaction with performance of appraisees with a high social comparison orientation critically depends on the average absolute performance feedback, organizations could choose strategically whether or not to disclose this information. If the average is high, it would reduce these appraisees' satisfaction with performance, so the information should not be revealed. If the average is low, it would increase their satisfaction with performance, so the information should be revealed.³ Alternatively, appraisers can also use procedural priming techniques similar to ours in order to induce appraisees to behave more comparatively when absolute performance feedback is low and less comparatively when absolute performance feedback is high. The results of our pilot study demonstrate that this set of guidelines for maximizing appraisees' satisfaction with performance in relative evaluation settings has yet to be communicated to most appraisers. Finally, based on our results, some appraisers could be tempted to maximize appraisees' satisfaction with

performance by shifting absolute performance feedback downward when appraising high social comparison orientation individuals and upward when appraising low social comparison orientation individuals (e.g., by removing or adding a constant from everyone's absolute performance feedback). Such a practice might be questionable and probably unethical. Thus engaging in it would require consultation with higher management and/or the organization's ethics board.⁴

Going back to our opening example, what could the manager have done differently to increase her employees' satisfaction? According to our data, she should have realized that due to extensive exposure to relative performance appraisal systems, the employees she supervised tended to engage in social comparison to a significant extent. Consequently, in order to increase satisfaction with performance, she should have implemented one of the strategic remedies discussed above. For instance, she could have adopted more stringent evaluation criteria in order to lower appraisees' individual and average absolute performance feedback. However, interesting considerations arise about such feedback intervention. We found that an attempt to increase satisfaction with performance with higher absolute performance feedback might result in a decrease in satisfaction in appraisees with a high social comparison orientation. However, recent research suggests that appraisees who receive inflated absolute performance feedback may be viewed more positively in subsequent selection processes (e.g., job applications; Swift, Moore, Sharek, & Gino, 2013). So, in such cases, should a manager evaluate employees generously, leading to better career prospects but lower satisfaction, or strictly, leading to worse career prospects but greater satisfaction? Similar considerations arise for professors who evaluate students relatively.

Limitations and Generalizations to Other Organizational Settings

The consistency of our results with our hypotheses and across studies indicates a high degree of internal validity. Nevertheless, the fact that some of our findings come from an educational context is a limitation and a potential threat to the external validity of our findings.

However, there are several reasons to be confident that the theoretical contributions and practical implications discussed in the previous sections may extend to many organizations that employ relative evaluation systems. First, although we use (among others) the performance evaluations of students in some (but not all) of our experiments, the core phenomenon we demonstrate and the theory behind it do not depend on some unique characteristics of classroom evaluations. Instead, they are specific to curve-based appraisal systems, which are commonly found in organizations (Backes-Gellner & Pull, 2013; Brustein, 2013; Ovide & Feintzeig, 2013). Second, the student participants we used in experiment 1 joined the workforce of several organizations a few months after the study was conducted. Thus, it is unlikely that immediately after graduation, they would start processing performance feedback very differently from what our results show. Third, the phenomenon we demonstrate is contingent on social comparison—a psychological process that is common (to varying extents) to every individual (Gibbons & Buunk 1999) and in organizational settings (Goodman & Haisley, 2007; Greenberg et al., 2007). Fourth, the effect we show in this paper was replicated across many research conditions, including variations in samples (i.e., student vs. not student) evaluation settings (i.e., classroom, sports competition, loyalty programs), positions relative to the average, and so on. Unless specific factors in particular organizational settings are demonstrated to interfere with the basic process on which we base our model, it seems natural to assume that our findings would hold in the case of organizational evaluations, too.

Another potential threat to the external validity of our findings lies in the use of respondents recruited from MTurk. Specifically, to what extent do participants recruited on MTurk differ systematically from individuals working in organizations? Would treatment effects and parameter estimates obtained from our sample of MTurk respondents differ from those that would be obtained from different samples (e.g., a sample of front-line employees)? Research has indeed found that females, democrats, Caucasians, and non-religious participants tend to be over-represented

(Behrend, Sharek, Meade, & Wiebe, 2011; Berinsky, Huber, & Lenz, 2012; Paolacci, Chandler, and Ipeirotis, 2010). Yet, the same studies also indicate that respondents recruited on MTurk are relatively representative of the U.S. workforce. Almost all of them are aged 18—55, a large majority is employed and has at least a two-year degree, they work in a variety of industries and geographical regions, and have a median income of \$45,000. Therefore, being relatively representative of employees working in various organizations, MTurk respondents should process the performance feedback they received at work in a similar manner to the feedback we provide in our experiments. There are also other biases and limitations associated with the use of MTurk. For example, there are legitimate worries about the attention level of respondents and the quality of the answers collected. Following the procedure suggested by Behrend et al. (2011), we assessed the quality of our data based on excessive response consistency and random responding. Since this procedure requires a pair of Likert scale items that should have similar responses and another pair that should have opposite responses, we did so only for experiment 2, the sole experiment where we used such a scale (i.e., social comparison scale; Gibbons & Buunk, 1999). Excessive consistency was assessed by flagging respondents who gave the exact same answer to two items that should have opposite responses.⁵ Since the two items used for this test are not perfect opposites, this is a very strict test of data quality. Ten cases (7.9% of our sample) were flagged by this test. Random responders were flagged when two items that should have similar responses had responses that were more than two points apart.⁶ Only two cases (1.6% of our sample) were flagged by this test. The proportion of flagged cases in our sample is much lower than the 24% found in Behrend et al. (2011), which suggests that our data are of good quality.⁷ The difference could be explained by the fact that we required our participants to have at least 98% positive feedback on MTurk, which removed a large portion of less attentive respondents. Attesting to the validity of the quality criteria proposed by Behrend et al. (2011), removing the flagged participants from the analyses of

experiment 2 only made the results stronger.⁸ Since we used the same selection criteria to recruit participants of experiment 3, we believe that the quality of the data should be similar to that of experiment 2.

A final limitation of our experiments is the sole use of hypothetical scenarios to test our hypotheses, since we cannot assess the magnitude of our effects in a naturalistic, more complex, and more realistic setting. At the same time, however, hypothetical scenarios allow us to test unambiguously the causal relationship between absolute performance feedback and satisfaction with performance. Besides, given the basic nature of the cognitive processes pertaining to appraisees' reactions to feedback, previous research conducted with hypothetical scenarios that are not specific to organizational settings (e.g., weight-guessing game) (e.g., Garcia & Tor 2007; Klein 1997) has been able to productively inform managers, employees, and policy makers about the effects of performance feedback an evaluation.

Directions for Future Research

Our work offers several avenues for future research. As stated above, future research might want to replicate our findings in different samples to establish their validity. Future research could also examine the extent to which our theory affects behavior in organizations when individuals receive real monetary incentives for their relative performance but none for the absolute performance. In addition, for the purpose of generalizability, this paper has demonstrated the role of social comparison in relatively straightforward and common evaluation settings (e.g., classroom, sports competition, loyalty programs). Future research could attempt to extend our work into more complex settings, for example, situations in which appraisees have more or different evaluation information, in which appraisees engage in complex social interactions, or in which they have robust evaluation expectations may constitute boundary conditions for our research. Another avenue for future research lies in the fact that we focused on satisfaction with performance. Future

studies could attempt to replicate and extend our findings using a broader conceptualization of satisfaction (e.g., job satisfaction and its facets). Finally, our research indicates that appraisers can influence appraisees' satisfaction with performance by using stricter (laxer) evaluation criteria to inflate (deflate) each appraisee's absolute performance feedback by the same amount. However, we do not explore changes in appraisees' satisfaction with appraisers, perceived fairness, or changes in future performance and motivation. Future research should explore potential results for other dependent variables. It is possible that a strict appraiser might increase the satisfaction with performance and future motivation of appraisees with a high social comparison orientation, but nevertheless be very disliked by appraisees, thus giving rise to an interesting agency problem.

Conclusion

Relative evaluation and reward systems are common in organizations (e.g., Backes-Gellner & Pull, 2013). To date, the particularities of such systems, even for traditionally important variables such as satisfaction with performance, are not well understood (e.g., Goffin et al., 2009). As a result, some widely held assumptions about the effects of performance feedback in these systems may be unjustified, potentially leading to the adoption of sub-optimal evaluation practices and thus to undesirable effects on employees and organizational performance. We demonstrate that one such assumption (held by both appraisers and appraisees) is that appraisees' satisfaction with performance will be higher (lower) with higher (lower) absolute performance feedback in relative evaluation systems. We show that this common assumption holds for appraisees who have a low social comparison orientation but not for those who have a high social comparison orientation. Indeed, while keeping relative performance constant, appraisees reported lower (higher) performance satisfaction ratings with higher (lower) absolute performance feedback. Importantly, we also demonstrate why and how this intriguing effect occurs, regardless of whether social

comparison orientation is chronically or contextually activated. This effect is also expected to hold for organizations that typically hire employees with a high social comparison orientation (e.g., MBA's) and organizations that foster internal relative comparisons (e.g., use tournament compensation systems).

Moreover, our paper proposes concrete solutions for increasing satisfaction with performance by taking into account the interactive effect between social comparison orientation and absolute performance feedback level. First, as social comparison orientation can be a chronic individual tendency, an organization could include questionnaires that assess it in its hiring and monitoring procedures (cf., Gibbons & Buunk, 1999). If social comparison orientation is found to be relatively high (low), the organization should adopt stricter (laxer) evaluation criteria. This would result in lower (higher) absolute performance feedback scores for everyone and, according to our findings, higher satisfaction with performance. Moreover, if these arrangements are not feasible, the organization could consider dropping the use of relative evaluation systems (e.g., as Microsoft did in 2013; Brustein, 2013; Ovide & Feintzeig, 2013), or simply hire employees who would be a good match with the evaluation criteria used for the organization: high (low) social comparison orientation individuals should be hired by organizations with strict (relaxed) evaluation criteria. Second, as social comparison can also be activated when individuals repeatedly compare themselves with each other, organizations that make heavy use of relative performance evaluation systems should adopt strict evaluation criteria for absolute performance feedback. Alternatively, they could try to adopt other policies that reduce social comparison within the organization (e.g., team-based bonuses), or simply refrain from disclosing absolute performance feedback.

More generally, our paper shows that relative evaluation systems can have consequences that are not immediately obvious. Further investigations, based on both organizational and psychological principles, are required in order to guarantee the optimal use of such system

¹ Participants were recruited via the Society for Judgment and Decision Making mailing list. Department affiliation of participants: Decision Science 10%; Economics 8%; Marketing 18%; OB/HR 10%; Psychology 36%; Others 18%

² The results for the three context replicates were virtually identical (all interaction terms including context replicate were not significant). Since our context replicates served only generalizability but not theoretical purposes, we pooled results across the three replicates for our analyses, leaving us with a 2 (absolute performance feedback: high vs. low) $\times$ 2 (relative performance: above vs. below average). We still tracked the effects of the different context replicates by running a fixed-effects model with dummy variables representing different replicates.

³ We thank an anonymous reviewer for pointing out this implication to us.

⁴ We thank an anonymous reviewer for the suggestion to emphasize and discourage potentially unethical applications of our work.

⁵ The two items of the social comparison scale used for the excessive consistency test were: “I am not the type of person who compares often with others” and “I often compare myself with others with respect to what I have accomplished in life.”

⁶ The two items of the social comparison scale used for the random responder test were: “I always pay a lot of attention to how I do things compared to how others do things” and “If I want to find out how well I have done something, I compare what I have done to how others have done.”

⁷ We could not check the Long String Index (the longest sequence of answers in the same scale point—e.g., strongly agree—provided by each respondent) because our social comparison scale was presented to each participant in a random order that was not recorded at the time by our survey software.

⁸ Analyses with and without flagged participants are available on request from the authors.

Acknowledgments: The authors gratefully acknowledge the helpful comments of Xianchi Dai, Matthias Seifert, Tara Wernsing, and Haiyang Yang, on earlier versions of this article. They also gratefully acknowledge the helpful comments of the Associate Editor and three anonymous reviewers.

All three authors have contributed equally to this work and should be considered as co-first authors. The order of authorship was determined randomly

This research received partial funding from IE Business School's Seed Funds Programme, from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement No. 298420, and from the Spanish Ministry of Economy and Competitiveness ECO2011-26384.

Jonathan Luffarelli is a PhD Candidate in Marketing and an adjunct professor of Marketing at IE Business School – IE University. He holds a Master in Research Methodology in Management Science from IE University, a Master in International Management from IE Business School, and a B.A. in Marketing from the Dublin Business School. Most of his research focuses on the effects of corporate social responsibility and brand aesthetics on brand equity. His work has been published in numerous international conference proceedings.

Dilney Gonçalves has been an assistant professor of Marketing at IE Business School - IE University since 2010. He holds a PhD in Marketing and a MSc. in Management from INSEAD. He has more than 9 years of professional experience in the IT industry. His research focuses on judgment and decision making, with particular interest in consumer choices, social comparison processes, and their effects on well-being. His work has been presented in international conferences like the Society for Personality and Social Psychology, the Society for Consumer Psychology, and the Association for Consumer Research.

Antonios Stamatogiannakis has been an assistant professor of Marketing at IE Business School - IE University since 2011. He holds a PhD from INSEAD. His current research interests lie in the area of motivation, and specifically conscious goal-setting, as well as on subtle effects of visual perception on human behavior. He has received competitive research funding from several international organizations, as well as from industry sources. His work has been published in the *Journal of Consumer Research*, as well as in numerous international conference proceedings.

References

- Aiken, L.S., & West, S.G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA:Sage.
- Alicke, M.D., Zell, E., & Bloom, D.L. (2010). Mere categorization and the frog-pond effect. *Psychological Science*, 21:2, 174–177.
- Alves, H., & Raposo, M. (2007). Conceptual model of student satisfaction in higher education. *Total Quality Management*, 18:5, 571–588.
- Alves, H., & Raposo, M. (2009). The measurement of the construct satisfaction in higher education'. *The Service Industries Journal*, 29:2, 203–218.
- Athiyaman, A. (1997). Linking student satisfaction and service quality perceptions: The case of university education. *European Journal of Marketing*, 31:7, 528–540.

- Atwater, L.E., Brett, J.F., & Charles, A.C. (2007). Multisource feedback: Lessons learned and implications for practice. *Human Resource Management*, 46:2, 285–307.
- Atwater, L.E., Waldman, D.A., & Brett, J.F. (2002). Understanding and optimizing multisource feedback. *Human Resource Management*, 41:2, 193–208.
- Backes- Gellner, U., & Pull, K. (2013). Tournament compensation systems, employee heterogeneity, and firm performance. *Human Resource Management*, 52:3, 375–398.
- Behrend, T.S., Sharek, D.J., Meade, A.W., & Wiebe, E.N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43:3, 800–813.
- Berinsky, A.J., Huber, G.A., & Lenz, G.S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20:3, 351–368.
- Biron, M., Farndale, E., & Paauwe, J. (2011). Performance management effectiveness: Lessons from world-leading firms. *International Journal of Human Resource Management*, 22:6, 1294–1311.
- Blume, B.D., Rubin, R.S., & Baldwin, T.T. (2013). Who is attracted to an organisation using a forced distribution performance management system? *Human Resource Management Journal*, 23:4, 360–378.
- Bogart, L.M., Benotsch, E.G., & Pavlovic, J.D. (2004). Feeling superior but threatened: The relation of narcissism to social comparison. *Basic and Applied Social Psychology*, 26:1, 35–44.
- Bothner, M.S., Kang, J-h, & Stuart, T.E. (2007). Competitive crowding and risk taking in a tournament: Evidence from NASCAR racing. *Administrative Science Quarterly*, 52:2, 208–247.
- Bracken, D.W., Timmreck, C.W., Fleenor, J.W., & Summers, L. (2001). 360 feedback from another angle. *Human Resource Management*, 40:1, 3–20.
- Bradley, D.E., & Roberts, J.A. (2004). Self- employment and job satisfaction: Investigating the role of self-efficacy, depression, and seniority. *Journal of Small Business Management*, 42:1, 37–58.
- Brown, D.J., Ferris, L.D., Heller, D., & Keeping, L.M. (2007). Antecedents and consequences of the frequency of upward and downward social comparisons at work. *Organizational Behavior and Human Decision Processes*, 102:1, 59–75.
- Brown, M., Hyatt, D., & Benson, J. (2010). Consequences of the performance appraisal experience. *Personnel Review*, 39:3, 375–396.
- Brustein, J. (2013). Microsoft kills its hated stack rankings. Does anyone do employee reviews right? in, *BusinessWeek*, <http://www.businessweek.com/articles/2013-11-13/microsoft-kills-its-hated-stack-rankings-dot-does-anyone-do-employee-reviews-right>.
- Buunk, A.P., & Gibbons, F.X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, 102:1, 3–21.
- Caprara, G.V., Barbaranelli, C., Steca, P., & Malone, P.S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44:6, 473–490.

- Carter, R.E. (2009). The impact of perceived service quality on MBA student satisfaction and recommendations: Do expectations matter? *Services Marketing Quarterly*, 30:3, 234–248.
- Cho, H-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research*, 66:9, 1261–1266.
- Christen, M., Iyer, G., & Soberman, D. (2006). Job satisfaction, job performance, and effort: A reexamination using agency theory. *Journal of Marketing*, 70:1, 137–150.
- Conner, D.S. (2003). Social comparison in virtual work environments: an examination of contemporary referent selection. *Journal of Occupational and Organizational Psychology*, 76:1, 133–147.
- Edmans, A. (2012). The link between employee satisfaction and firm value, with implications for corporate social responsibility. *The Academy of Management Perspectives*, 26:4, 1–19.
- Epstude, K., & Roese, N.J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, 12:2, 168–192.
- Ewen, R.B. (1973). Pressure for production, task difficulty, and the correlation between job satisfaction and job performance. *Journal of Applied Psychology*, 58:3, 378–380.
- Faragher, E.B., Cass, M., & Cooper, C.L. (2005). The relationship between job satisfaction and health: A meta-analysis. *Occupational and Environmental Medicine*, 62:2, 105–112.
- Farndale, E., & Kelliher, C. (2013). Implementing performance appraisal: Exploring the employee experience. *Human Resource Management*, 52:6, 879–897.
- Ferris, G.R., Munyon, T.P., Basik, K., & Buckley, M.R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review*, 18:3, 146–163.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7:2, 117–140.
- Fitzsimons, G.J. (2008). Death to dichotomizing. *Journal of Consumer Research*, 35:1, 5–8.
- Garcia, S.M., & Tor, A. (2007). Rankings, standards, and competition: Task vs. scale comparisons. *Organizational Behavior and Human Decision Processes*, 102:1, 95–108.
- Gibbons, F.X., & Buunk, A.P. (1999). Individual differences in social comparison: The development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, 76:1, 129–142.
- Goffin, R.D., Jelley, B.R., Powell, D.M., & Johnston, N.G. (2009). Taking advantage of social comparisons in performance appraisal: The relative percentile method. *Human Resource Management*, 48:2, 251–268.
- Gong, G., Li, L.Y., & Shin, J.Y. (2011). Relative performance evaluation and related peer groups in executive compensation contracts. *The Accounting Review*, 86:3, 1007–1043.
- Goodman, P.S., & Haisley, E. (2007). Social comparison processes in an organizational context: New directions. *Organizational Behavior and Human Decision Processes*, 102:1, 109–125.

- Goubeaud, K. (2010). How is science learning assessed at the postsecondary level? Assessment and grading practices in college biology, chemistry and physics. *Journal of Science Education and Technology*, 19:3, 237–245.
- Greenberg, J., Ashton-James, C.E., & Ashkanasy, N.M. (2007). Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes*, 102:1, 22–41.
- Harris, M.M., Anseel, F., & Lievens, F. (2008). Keeping up with the Joneses: A field study of the relationships among upward, lateral, and downward comparisons and pay level satisfaction. *Journal of Applied Psychology*, 93:3, 665–673.
- Harris, P.R., & Smith, V. (2005). When the risks are low: The impact of absolute and comparative information on disturbance and understanding in US and UK samples. *Psychology & Health*, 20:3, 319–330.
- Henning-Thurau, T., Langer, M., & Hansen, U. (2001). Modeling and managing student loyalty: An approach based on the concept of relationship quality. *Journal of Service Research*, 3:4, 331–344.
- Heslin, P.A. (2005). Conceptualizing and evaluating career success. *Journal of Organizational Behavior*, 26:2, 113–136.
- Hu, J., & Liden, R.C. (2013). Relative leader–member exchange within team contexts: How and when social comparison impacts individual effectiveness. *Personnel Psychology*, 66:1, 127–172.
- Iaffaldano, M.T., & Muchinsky, P.M. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin*, 97:2, 251–273.
- Ilgen, D.R. (1971). Satisfaction with performance as a function of the initial level of expected performance and the deviation from expectations. *Organizational Behavior and Human Performance*, 6:3, 345–361.
- Irwin, J.R., & McClelland, G.H. (2001). Misleading heuristics for moderated multiple regression models. *Journal of Marketing Research*, 38:1, 100–109.
- Iyengar, S.S., Wells, R.E., & Schwartz, B. (2006). Doing better but feeling worse. *Psychological Science*, 17:2, 143–150.
- Jacowitz, K.E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21:11, 1161–1166.
- Johnson, M.D., Ilies, R., & Boles, T.L. (2012). Alternative reference points and outcome evaluation: The influence of affect. *Journal of Applied Psychology*, 97:1, 33–45.
- Judge, T.A., & Bono, J.E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86:1, 80–92.
- Judge, T.A., Bono, J.E., Erez, A., & Locke, E.A. (2005). Core self-evaluations and job and life satisfaction: The role of self-concordance and goal attainment. *Journal of Applied Psychology*, 90:2, 257–268.
- Judge, T.A., Thoresen, C.J., Bono, J.E., & Patton, G.K. (2001). The job satisfaction-job performance

relationship: A qualitative and quantitative review. *Psychological Bulletin*, 127:3, 376–407.

- Klein, W.M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, 72:4, 763–774.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119:2, 254–284.
- Kluger, A.N., & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current Directions in Psychological Science*, 7:3, 67–72.
- Lane, D.J., & Gibbons, F.X. (2007). Social comparison and satisfaction: Students' reactions after exam feedback predict future academic performance. *Journal of Applied Social Psychology*, 37:6, 1363–1384.
- Larrick, R.P., Heath, C., & Wu, G. (2009). Goal-induced risk taking in negotiation and decision making. *Social Cognition*, 27:3, 342–364.
- Locke, K.D., & Nekich, J.C. (2000). Agency and communion in naturalistic social comparison. *Personality and Social Psychology Bulletin*, 26:7, 864–874.
- MacKenzie, S.B., Podsakoff, P.M., & Ahearne, M. (1998). Some possible antecedents and consequences of in-role and extra-role salesperson performance. *Journal of Marketing*, 62:3, 87–98.
- Markman, K.D., McMullen, M.N., & Elizaga, R.A. (2008). Counterfactual thinking, persistence, and performance: A test of the reflection and evaluation model. *Journal of Experimental Social Psychology*, 44:2, 421–428.
- McCabe, D.L., Butterfield, K.D., & Treviño, L.K. (2006). Academic dishonesty in graduate business programs: Prevalence, causes, and proposed action. *Academy of Management Learning & Education*, 5:3, 294–305.
- McCabe, D.L., & Treviño, L.K. (1995). Cheating among business students: A challenge for business leaders and educators. *Journal of Management Education*, 19:2, 205–218.
- Medvec, V.H., Madey, S.F., & Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology*, 69:4, 603–610.
- Medvec, V.H., & Savitsky, K. (1997). When doing better means feeling worse: The effects of categorical cutoff points on counterfactual thinking and satisfaction. *Journal of Personality and Social Psychology*, 72:6, 1284–1296.
- Miller, J.S. (2001). Self-monitoring and performance appraisal satisfaction: An exploratory field study. *Human Resource Management*, 40:3, 321–332.
- Moore, D.A., & Klein, W.M. (2008). Use of absolute and comparative performance feedback in absolute and comparative judgments and decisions. *Organizational Behavior and Human Decision Processes*, 107:1, 60–74.
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences.

Psychological Review, 110:3, 472–489.

- Mussweiler, T., & Epstude, K. (2009). Relatively fast! Efficiency advantages of comparative thinking. *Journal of Experimental Psychology*, 138:1, 1–21.
- Mussweiler, T., Rüter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology*, 40:5, 689–696.
- Mussweiler, T., & Strack, F. (2000). The “relative self”: Informational and judgmental consequences of comparative self-evaluation. *Journal of Personality and Social Psychology*, 79:1, 23–38.
- Oppenheimer, D.M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45:4, 867–872.
- Ovide, S., & Feintzeig, R. (2013). Microsoft abandons “stack ranking” of employees, in, *The Wall Street Journal*, <http://online.wsj.com/news/articles/SB10001424052702303460004579193951987616572>.
- Paolacci, G., Chandler, J., & Ipeirotis, P.G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5:5, 411–419.
- Petty, M.M., McGee, G.W., & Cavender, J.W. (1984). A meta-analysis of the relationships between individual job satisfaction and individual performance. *Academy of Management Review*, 9:4, 712–721.
- Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta- analysis. *Human Resource Management*, 51:5, 709–732.
- Preacher, K.J., Rucker, D.D., & Hayes, A.F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42:1, 185–227.
- Riketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology*, 93:2, 472–481.
- Rodwell, J.J., Kienzle, R., & Shadur, M.A. (1998). ‘The relationship among work- related perceptions, employee attitudes, and employee performance: The integral role of communications’, *Human Resource Management*, 37:3–4, 277–293.
- Rynes, S.L., Colbert, A.E., & Brown, K.G. (2002). HR professionals’ beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, 41:2, 149–174.
- Saari, L.M., & Judge, T.A. (2004). Employee attitudes and job satisfaction. *Human Resource Management*, 43:4, 395–407.
- Schleicher, D.J., Bull, R.A., & Green, S.G. (2009). Rater reactions to forced distribution rating systems. *Journal of Management*, 35:4, 899–927.
- Schleicher, D.J., Watt, J.D., & Greguras, G.J. (2004). Reexamining the job satisfaction–performance relationship: The complexity of attitudes. *Journal of Applied Psychology*, 89:1, 165–177.
- Scullen, S.E., Bergey, P.K., & Aiman- Smith, L. (2005). Forced distribution rating systems and the

improvement of workforce potential: A baseline simulation. *Personnel Psychology*, 58:1, 1–32.

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

Swift, S.A., Moore, D.A., Sharek, Z.S., & Gino, F. (2013). Inflated applicants: Attribution errors in performance evaluation by professionals. *Public Library of Science ONE*, 8:July, 10.1371/journal.pone.0069258.

Tepper, B.J., Duffy, M.K., Henle, C.A., & Lambert, L.S. (2006). ‘Procedural injustice, victim precipitation, and abusive supervision. *Personnel Psychology*, 59:1, 101–123.

Van der Zee, K., Oldersma, F., Buunk, B.P., & Bos, D. (1998). Social comparison preferences among cancer patients as related to neuroticism and social comparison orientation. *Journal of Personality and Social Psychology*, 75:3, 801–810.

Van Scotter, J.R. (2000). Relationships of task performance and contextual performance with turnover, job satisfaction, and affective commitment. *Human Resource Management Review*, 10:1, 79–95.

Vess, M., Arndt, J., Cox, C.R., Routledge, C., & Goldenberg, J.L. (2009). Exploring the existential function of religion: The effect of religious fundamentalism and mortality salience on faith-based medical refusals. *Journal of Personality and Social Psychology*, 97:2, 334–350.

Whitman, D.S., Van Rooy, D.L., & Viswesvaran, C. (2010). Satisfaction, citizenship behaviors, and performance in work units: A meta- analysis of collective construct relations. *Personnel Psychology*, 63:1, 41–81.

Zhao, X., Lynch Jr., J.G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37:2, 197–206.

Zhou, K.Z., Li, J.J., Zhou, N., & Su, C. (2008). Market orientation, job satisfaction, product quality, and firm performance: Evidence from China. *Strategic Management Journal*, 29:9, 985–1000.

Appendix

Context replicates used in experiment 2

For presentation purposes, the high absolute performance feedback, above average absolute performance feedback condition is presented as the default. Variations of the low absolute performance feedback, above average absolute performance feedback condition appear in (parentheses). Variations of the high absolute performance feedback, below average absolute performance feedback condition appear in [brackets]. Variations on the low absolute performance feedback, below average absolute performance feedback condition appear in {curled brackets}.

Classroom

“Emma is a Master’s student in a business school. She knows that the letter grade she will receive for the course depends only on her relative position compared to others. Today she received her exam and found that she got 8.5 (6.5) [7.5] {5.5} out of 10 and that the average grade of her class in the exam was 8 (6) [8] {6}. She also found that her performance was better (better) [worse] {worse} than that of 60% of the students; worse (worse) [better] {better} than 40%.”

Sport competition

“Danny, the striker of a soccer team from a student league, knows that the more goals he scores compared to other strikers, the more likely he is to be drafted into professional teams. He realizes that strikers from other teams scored 9 (5) [9] {5} goals in average and that he scored 12 (8) [6] {2}. Finally, he finds that his performance was better (better) [worse] {worse} than that of 70% of the strikers engaged in the competition; worse (worse) [better] {better} than that of 30%.”

Loyalty program

“Maria is a member of the loyalty program of a supermarket chain, and she scores points on the basis of the total amount of Euros she spends. At the end of the year, each member gets a gift whose value only depends on the number of points accumulated compared to other participants. Today, Maria receives a mail indicating that, over the year, she has received more (more) [fewer] {fewer} points than 55% of the program’s members; fewer (fewer) [more] {more} points than 45%. Other participants have received an average of 16 300 (6 600) [16 300] {6 600} points. She has received 20 375 (12 225) [8 250] {4950} points.

TABLE 1. Overview of the studies and results

Pilot study	
Objective:	To examine appraisers' lay beliefs about relative performance appraisals.
Method:	Survey: Appraisers directly choose whether they think that high versus low absolute performance feedback would make appraisees more satisfied with performance.
Results:	The large majority of appraisers believe that appraisees' satisfaction with performance will be higher with higher (lower) absolute performance feedback. This result is true regardless of whether appraisers have experience with relative appraisals or not.
Experiment 1	
Objectives:	<i>Part A:</i> To demonstrate that higher (lower) feedback can result in lower (higher) satisfaction with performance, in a relative context where appraisees are used to compare themselves with others. <i>Part B:</i> To examine appraisees' lay beliefs about relative performance appraisals.
Method:	<i>Part A:</i> Two between-participants conditions: Appraisees provide ratings of satisfaction with a given (high vs. low) level of absolute performance feedback. Based on theoretical reasons, social comparison orientation is assumed to be naturally high among all participants. Relative performance is kept constant and absolute performance is below the mean. <i>Part B:</i> Survey: The same appraisees directly choose between receiving high versus low absolute performance feedback.
Results:	<i>Part A:</i> Appraisees provide higher (lower) satisfaction with performance ratings with lower (higher) absolute performance feedback. <i>Part B:</i> In contrast to the results of part A, most appraisees believe that satisfaction with performance will be higher with higher (vs. lower) absolute performance feedback.
Experiment 2	
Objectives:	To examine the moderating role of chronic social comparison orientation in the performance-satisfaction relationship. To replicate the finding that lower (higher) absolute performance feedback can cause higher (lower) satisfaction with performance across performance situations and regardless of whether absolute performance is above or below than the mean.
Method:	Four between-participants conditions: Appraisees provide ratings of satisfaction with a given (high vs. low performance $\times$ above vs. below the mean) level of absolute performance feedback. Social comparison orientation is measured on a validated multi-item scale. Three different evaluation situations are used.
Results:	Appraisees with a high (low) social comparison orientation provide higher satisfaction ratings with lower (higher) absolute performance feedback. These results hold across performance situations and when appraisees are above or below the mean.
Experiment 3	
Objectives:	To replicate the finding that lower (higher) absolute performance feedback can cause higher (lower) satisfaction with performance by priming (vs. measuring) social comparison orientation. To examine the mediating role of self-perceived competence in the performance-satisfaction relationship.
Method:	Four between-participants conditions: Social comparison orientation is manipulated by priming (primed vs. control). Then, appraisees provide ratings of satisfaction with a given (high vs. low) level of absolute performance feedback. Self-perceived competence is measured. Relative performance is kept constant and absolute performance is above the mean.
Results:	The absolute performance feedback $\times$ social comparison orientation interaction effect on satisfaction with performance observed in Experiment 2 is replicated and mediated by self-perceived competence.

TABLE 2. Experiment 3—moderated mediation analysis results of the effects of absolute performance feedback, social comparison orientation, and self-perceived competence on satisfaction with performance

A. Self-perceived competence				
<i>Model summary</i>	<i>R</i>	<i>R</i> ²	<i>F</i>	<i>p</i>
	0.28	0.08	7.36	≤ 0.001
<i>Predictor</i>	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Constant	−0.27	0.12	−2.17	0.031
Absolute performance feedback (APF)	0.44	0.17	2.51	0.013
Social comparison orientation (SCO)	0.45	0.15	2.94	0.004
APF × SCO	−0.45	0.17	−2.57	0.011
B. Satisfaction with performance				
<i>Model summary</i>	<i>R</i>	<i>R</i> ²	<i>F</i>	<i>p</i>
	−0.46	−0.21	16.85	≤ 0.001
<i>Predictor</i>	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>
Constant	−0.20	0.12	−1.72	.087
Self-perceived competence	0.41	0.06	7.00	≤ 0.001
Absolute performance feedback (APF)	0.33	0.16	2.02	0.045
Social comparison orientation (SCO)	0.26	0.15	1.82	0.070
APF × SCO	−0.33	0.16	−2.03	0.043
C. Conditional indirect effect of APF at values of SCO				
<i>Social comparison orientation</i>	<i>Effect</i>	<i>SE</i>	<i>Lower 95% CI</i>	<i>Upper 95% CI</i>
0 (i.e., control condition)	0.18	0.070	0.058	0.330
1 (i.e. primed condition)	−0.004	0.001	−0.007	−0.001

FIGURE 1. Distribution of absolute performance feedback following stricter and laxer evaluation criteria

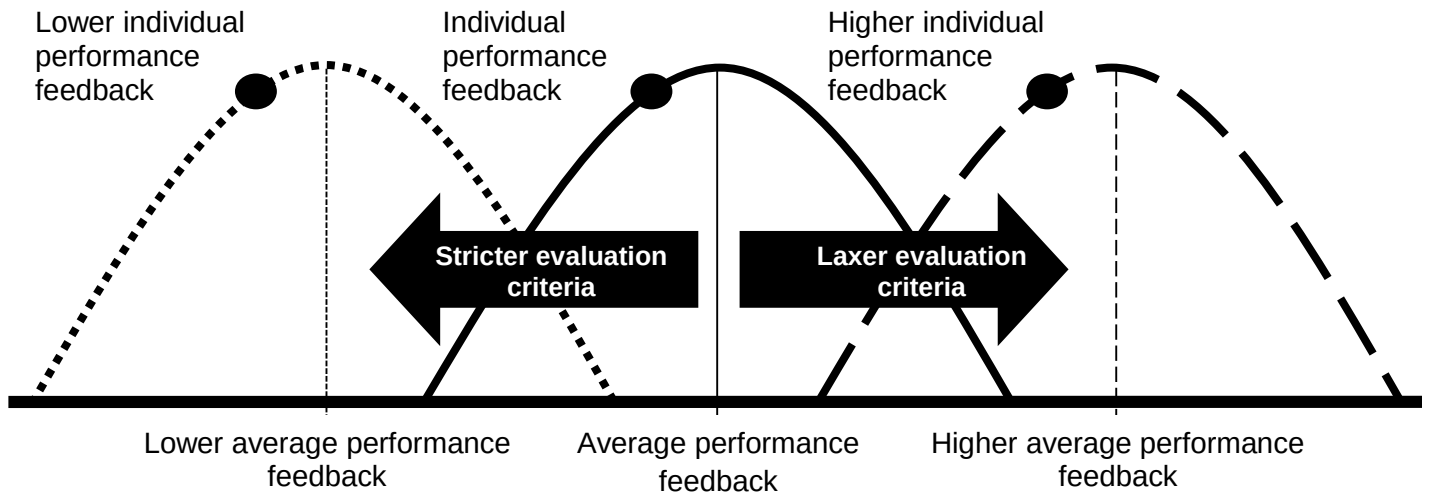


FIGURE 2. Hypothesized moderated mediation of the effects of absolute performance feedback, social comparison orientation, and self-perceived competence on satisfaction

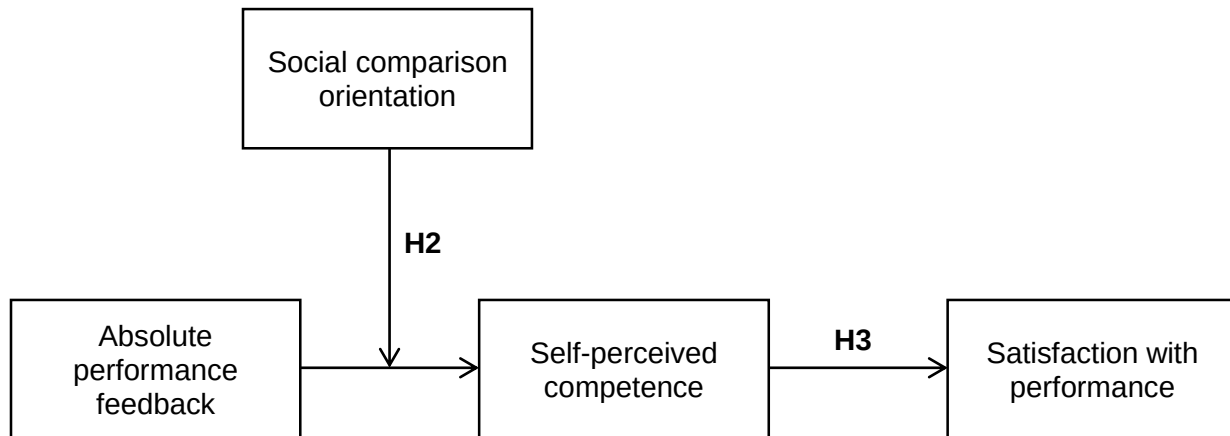


FIGURE 3. Experiment 2—results of the spotlight analysis for low versus high absolute performance feedback at low and high levels of social comparison orientation

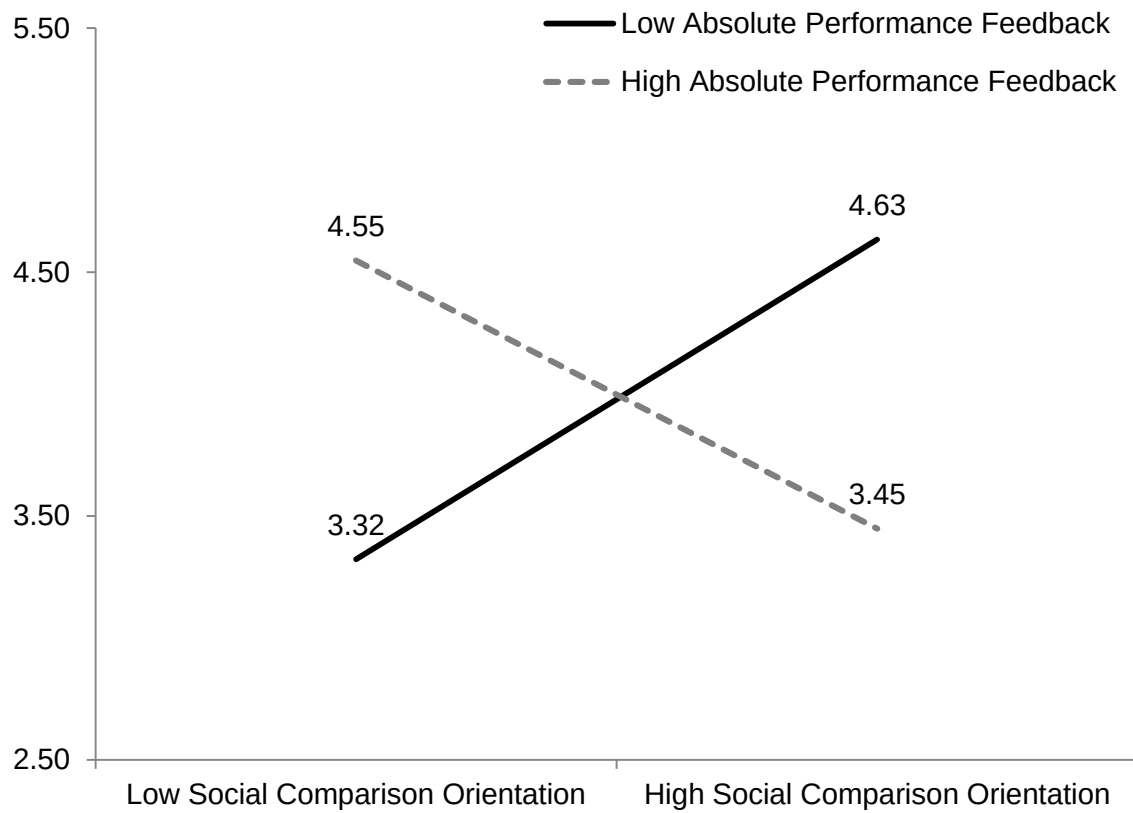


FIGURE 4. Experiment 3—moderation diagram of the effects of absolute performance feedback and social comparison orientation on satisfaction with performance

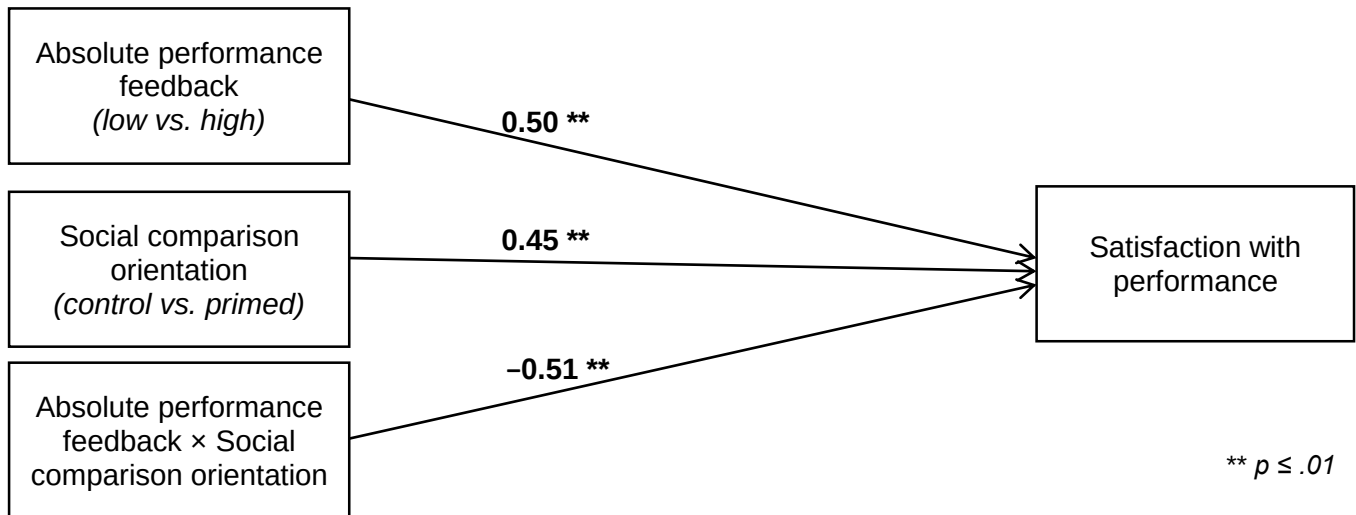


FIGURE 5. Experiment 3—moderated mediation diagram of the effects of absolute performance feedback, social comparison orientation, and self-perceived competence on satisfaction

