Social Comparison in Performance Appraisal

Jinseok S. Chun

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY
2018
ABSTRACT

Social Comparison in Performance Appraisal

Jinseok S. Chun

This dissertation examines to what extent social comparison is emphasized in performance evaluations of work organizations, how employees react to it, and whether there is an alternative to it. Operationalizing social comparison as an evaluation process that compares an employee’s performance to their coworkers’ performance, Studies 1 and 2 demonstrate that social comparison is emphasized to a stronger extent in collectivistic cultures than in individualistic cultures. Studies 3 and 4 find that employees in collectivistic cultures perceive higher procedural fairness when they receive social comparison evaluations as compared to employees in individualistic cultures. The mediation analyses from Studies 2 and 4 indicate that these findings are explained by the perceived descriptive and injunctive norms of social comparisons within collectivistic versus individualistic cultures, which shape people’s general attitudes toward using social comparison in evaluation settings. In collectivistic cultures that put strong emphasis on people’s social context, social comparison is considered to be a necessary component of performance evaluations. In contrast, in individualistic cultures where people focus on the specific characteristics of each person, social comparison is believed to be more or less irrelevant.

Given the aversive effect of social comparison in individualistic cultures, the second chapter of this dissertation investigates whether there is a proper alternative to social comparison in the context of performance evaluations. It finds that temporal comparison—which compares
an employee’s performance to his or her own past performance—can be such an alternative. Temporal comparison secures employees’ perceptions of fairness by providing the beliefs that their evaluators are focusing on them and their specific characteristics. These findings imply that employees in individualistic cultures want their independent identities to be acknowledged at work, and providing temporal comparison evaluations is one way to fulfill such needs.
TABLE OF CONTENTS

List of Tables ........................................................................................................................................... iv
List of Figures............................................................................................................................................... v
Dedication ................................................................................................................................................... vi

CHAPTER 1

Introduction .................................................................................................................................................. 2
Social Comparisons and Competition ................................................................................................. 6
Strengths of Social Comparisons in Collectivistic versus Individualistic Cultures ...................... 7
Perceived Fairness of Social Comparisons in Collectivistic Versus Individualistic Cultures........ 9
  Overview of Studies ............................................................................................................................. 10
Study 1 .................................................................................................................................................... 11
  Method .................................................................................................................................................. 12
  Results ................................................................................................................................................ 13
  Discussion .......................................................................................................................................... 14
Study 2 .................................................................................................................................................... 15
  Method ................................................................................................................................................ 16
  Results ............................................................................................................................................... 18
  Discussion .......................................................................................................................................... 21
CHAPTER 2

Introduction ........................................................................................................69

Temporal Comparisons, Social Comparisons, and Fairness .................................72

The Effects of Temporal and Social Comparisons on Fairness Perceptions ........73
LIST OF TABLES

Table 1. Descriptive Statistics and Intercorrelations from Study 1 ........................................50
Table 2. Different Types of Comparisons from Study 1 .............................................................51
Table 3. Descriptive Statistics and Intercorrelations from Study 2 ........................................52
Table 4. Social Comparisons and Absolute Levels from Study 2 ..............................................53
Table 5. The Hypothesized Mediators from Study 2 ...............................................................54
Table 6. Descriptive Statistics and Intercorrelations from Study 3 ...........................................55
Table 7. Procedural Fairness from Study 3 ..................................................................................56
Table 8. Descriptive Statistics and Intercorrelations from Study 4 ............................................57
Table 9. Procedural Fairness from Study 4 ..................................................................................58
Table 10. The Hypothesized Mediators from Study 4 .................................................................59
Table 11. Descriptive Statistics and Intercorrelations from Study 5 .........................................109
Table 12. Descriptive Statistics and Intercorrelations from Study 6 ..........................................110
Table 13. Perceptions of Fairness from Study 6 .........................................................................111
Table 14. Descriptive Statistics and Intercorrelations from Study 7 .........................................112
Table 15. Perceptions of Fairness and Individualized Treatment from Study 7 ......................113
Table 16. Mediation Analyses from Study 7 ..............................................................................114
LIST OF FIGURES

Figure 1. Different types of comparisons from Study 1 ..........................................................60
Figure 2. Performance reports used in Study 2 (English version) .........................................61
Figure 3. Social comparisons and absolute levels from Study 2 .............................................62
Figure 4. Procedural fairness from Study 3 .............................................................................63
Figure 5. Procedural fairness from Study 4 .............................................................................64
Figure 6. Perceptions of fairness from Study 6 .......................................................................115
Figure 7. Perceptions of fairness and individualized treatment from Study 7 .........................116
Dedicated to my family, my friends, and my mentors,

with love and respect.
CHAPTER 1: SOCIAL COMPARISONS IN DIFFERENT CULTURES
Introduction

Aiming to harness the greatest potential of their human resources, companies often compare their employees’ performance against one another (Bandiera, Barankay, & Rasul, 2005; Besley & Ghatak, 2008; Card, Mas, Moretti, & Saez, 2012; Frey, 2007; Garcia & Tor, 2007; Oliver, Eggers, & Nils, in press). Past studies in micro-economics examined how comparison among multiple employees’ performance—called social comparison or relative performance feedback—influences individuals’ performance (Azmat & Iriberri, 2010; Barankay, 2012; Blanes i Vidal & Nossol, 2011; Charness, Masclet, & Villeval, 2011; Hannan, Krishnan, & Newman, 2008). These studies attracted research attention to the effects of social comparison on employees’ performance. However, there has been a limited amount of work on the psychological experiences that employees go through when they are compared with others.

This lack of research has been driven by the past psychological literature that has understood social comparison as a process that individuals initiate themselves (Festinger, 1954; Garcia, Tor, & Schiff, 2013; Gibbons & Buunk, 1999; Goethals, 1986; Suls, Martin, & Wheeler, 2002; Wood, 1989). Because research in psychology has mainly conceptualized social comparison as a self-initiated process, the literature is scarce of evidence regarding the psychological processes that employees experience upon receiving performance evaluations that compare them with others. Although there are valuable findings from the previous work that involved individuals who were informed where they stood relative to others (e.g., Dunn, Ruedy, & Schweitzer, 2012; Garcia & Tor, 2009; Levine & Green, 1984; Novemsky & Schweitzer, 2004; Tesser, 1988; Wheeler, Koestner, & Driver, 1982; Zell & Alicke, 2009), none of them focused on employees’ reactions to social comparison evaluations in workplace settings.

Seeking to fill this gap, in the present research I examine (a) to what extent social
comparison is used to evaluate employees’ performance in workplaces and (b) how employees react to social comparison evaluations they receive. I take cultural perspectives to test these questions. Although social comparison can be used to inform employees of their performance relative to other employees, how much it is emphasized in performance evaluations and which outcomes it begets may vary across different countries. Organizations and managers are expected to engage with their employees within the boundary of cultural norms (Javidan, Dorfman, de Luque, & House, 2006), because management practices that clash with the norms of the societies are considered illegitimate and thus breed dysfunctional outcomes (e.g., Fehr & Fischbacher, 2004; Morse, Gergen, Peele, & van Ryneveld, 1977). If explicit comparisons among multiple people are seen to violate cultural norms, managers may not emphasize social comparison in providing performance evaluations, and employees may demonstrate negative reactions when they receive evaluations that involve social comparison.

It is important to note that social comparisons in performance evaluations can take multiple forms. In the present research, I operationalize social comparisons in two specific types: (a) direct comparisons among multiple employees’ performance (peer comparisons; Levine & Green, 1984) and (b) discussion of employees’ performance in terms of their positions in a performance ranking (ranking comparisons; Garcia & Tor, 2007). With respect to cultures, I focus on the distinction between collectivism versus individualism. As I explain below, collectivism and individualism effectively capture people’s attitudes and perceptions regarding social comparisons, and thus offers a valuable perspective to analyze to what extent social comparisons are emphasized in performance evaluations across different cultures and how employees react to such approaches of performance appraisals.

The specific dependent variable that I use to capture employees’ reactions to social
comparison evaluations is perceptions of procedural fairness (Colquitt, 2001; Colquitt, Greenberg, & Zapata-Phelan, 2005). Procedural fairness refers to the extent to which employees believe that the processes adhere to fair criteria, such as accuracy, unbiasedness, and ethicality (Leventhal, Karuza, & Fry, 1980). It is differentiated from the other dimensions of fairness such as distributive fairness, which focuses on the outcomes that employees receive. Procedural fairness has been demonstrated to have significant and unique effects on various organizational outcomes including employees’ job attitudes and performance (Colquitt, Conlon, Wesson, Porter, & Ng, 2001).

The present research contributes to the literature in the following manners. First, as noted, past research on social comparisons in workplace settings mainly addressed the question of how social comparisons affect employees’ performance, without much attention on the psychological experiences that employees go through when they are compared with other people. By showing that social comparisons significantly affect important workplace perceptions such as procedural fairness, my findings illustrate the relevance of psychological perspectives in analyzing the effects of social comparison evaluations at work.

Second, the current work represents an intersection of the research streams on social comparisons and cultures. I find that cultures can provide a useful lens to understand people’s experiences of social comparison evaluations. In collectivistic cultures, people are defined in relation to those around them (Markus & Kitayama, 1991; Varnum, Grossmann, Kitayama, & Nisbett, 2010). Therefore, people in collectivistic cultures may believe that employees’ performance should be appraised relative to one another so the employees can acquire information about how they have performed in their social context. In contrast, people in individualistic cultures may believe that what an individual employee has achieved can be
effectively appraised by examining the specific features of his or her achievement (Kühnen, Hannover, & Schubert, 2001; Nisbett, Peng, Choi, & Norenzayan, 2001). My findings demonstrate that these different beliefs determine not only to what extent managers emphasize social comparisons in evaluating their employees’ performance but also how employees react to the social comparison evaluations they receive.

The present research provides practical advice for managers who appraise their employees’ performance. Although performance evaluations are conducted to help employees accurately understand how they are performing their jobs and how they can improve their performance, these benefits can only be realized when employees accept the legitimacy and validity of the evaluations (Leung, Su, & Morris, 2001). If employees think that the evaluations are based on inaccurate information or shaped by biased or unethical processes, they may disregard the contents of the evaluations and not utilize the evaluations for developmental purposes (Mayer & Davis, 1999). By showing how the effects of social comparison evaluations on fairness perceptions change across different cultural contexts, I illuminate what managers in different cultures should take into account to adapt the ways in which they appraise their employees’ performance.

My findings demonstrate that social comparisons are emphasized to a stronger extent in performance evaluations of collectivistic cultures as compared to individualistic cultures. Moreover, I demonstrate that social comparisons are perceived to be procedurally fairer in collectivistic cultures than in individualistic cultures. To delve into the mechanisms of these differences, I concentrate on the normative aspects of social comparisons in different cultures. It is found that people in collectivistic cultures perceive cultural norms that accept using social comparisons to appraise achievements. They thus demonstrate favorable attitudes toward using
social comparisons in performance evaluations. In individualistic cultures, by contrast, social comparisons are perceived to clash with the cultural norms that consider each individual as an independent entity with unique characteristics. In such situations, comparisons across multiple people can be viewed to fail to appreciate their distinctiveness. I find that the different norms and attitudes across collectivistic and individualistic cultures explain to what extent managers emphasize social comparisons in evaluating performance and to what extent employees perceive procedural fairness from social comparison evaluations they receive.

**Social Comparisons and Competition**

By providing information on where individuals stand in terms of their performance relative to others, social comparisons can stimulate their status concern and competitive motivation (Anderson, Hildreth, & Howland, 2015; Charness et al., 2011). However, research has found that the effect of social comparisons on individuals’ subsequent performance significantly varies across different settings. Although a few studies from micro-economics suggest that social comparisons can increase people’s performance (Azmat & Iriberri, 2010; Blanes i Vidal & Nossol, 2011), there is also evidence indicating performance-reducing effects of social comparisons (Barankay, 2012). Moreover, it is possible that the effect of social comparisons on performance may vary depending on the specific characteristics of the evaluations (Hannan et al., 2008).

Given these findings, it appears to be important to investigate employees’ psychological reactions upon receiving social comparison evaluations. Such investigations should enable researchers to acquire fundamental knowledge regarding how employees think and feel when they are compared with other people. They will provide the building blocks to understand how social comparisons affect various outcomes in organizations, and how such effects may vary.
depending on the situations or other characteristics of the evaluations. In the following sections, I examine how social comparisons shape employees’ perceptions of procedural fairness. But before I delve into how fairness perceptions are affected by social comparisons, I discuss why managers may use social comparisons in appraising employees’ performance. In these analyses, I focus on the roles of cultures and norms.

**Strengths of Social Comparisons in Collectivistic versus Individualistic Cultures**

One of the most fundamental findings in cross-cultural psychology is that other people comprise more important information in collectivistic cultures than in individualistic cultures (Markus & Kitayama, 1991; Varnum et al., 2010). People in collectivistic cultures define themselves in terms of their relationships with others, whereas people in individualistic cultures focus on their specific characteristics independent of others (Brewer & Chen, 2007). This implies that people in collectivistic cultures attend to the information about others to a stronger extent, so they can incorporate such information in assessing their current states.

Past research supported this prediction by showing that people in collectivistic cultures compared themselves against others to a stronger extent than people in individualistic cultures: Chinese people were found to demonstrate stronger social comparisons in self-evaluations as compared to American people (Chung & Mallery, 1999). Similarly, studies conducted in Canada revealed that, after engaging in a task, individuals with Asian backgrounds sought information regarding their rankings to a stronger extent than those with European backgrounds (White & Lehman, 2005).

The stronger emphasis on social comparisons in collectivistic cultures than in individualistic cultures may extend to the context where employees are receiving performance
evaluations. Numerous studies have shown that people in collectivistic cultures subscribe to contextual theories in describing people (e.g., Cross, Bacon, & Morris, 2000; Heine & Lehman, 1997; Kühnen et al., 2001; Morris & Peng, 1994). People with collectivistic mindsets attend to the situations surrounding their targets of perceptions in addition to the specific characteristics of the targets. Therefore, people in collectivistic cultures are likely to experience and observe many situations in which individuals’ achievements are appraised in terms of their social context: other people’s achievements. People in collectivistic cultures may thus perceive social norms suggesting that people in their societies use social comparisons in appraising performance (i.e., descriptive norms) and they approve of such practices (i.e., injunctive norms; Leung & Morris, 2015). These perceptions of social norms may shape managers’ generalized attitudes toward using social comparisons in appraising employees’ performance (Bicchieri, 2006; Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007), ultimately inducing them to emphasize social comparisons in specific performance evaluations they conduct.

In contrast, managers in individualistic cultures are predicted to use social comparisons to a much weaker extent. Individualism focuses on the specific characteristic of the targets instead of their context (Morris & Peng, 1994). Therefore, people in individualistic cultures are likely to perceive descriptive and injunctive norms that exclude social comparisons in performance evaluations. These perceptions of descriptive and injunctive norms would lead people in individualistic cultures to develop general negative attitudes toward using social comparisons in any appraisal settings. Thus, managers in individualistic cultures may not emphasize social comparisons in specific evaluations they deliver to their employees (Liu, Morris, & Yang, 2018). In summary, I predict that social comparisons are emphasized to a stronger extent in performance evaluations of collectivistic cultures as compared to individualistic cultures. Moreover, I predict
that this difference is mediated by the managers’ domain-general attitudes toward using social comparisons, which are in turn explained by the perceived domain-general descriptive and injunctive norms of social comparisons.

_Hypothesis 1._ Social comparisons are emphasized to a stronger extent in performance evaluations of collectivistic cultures than in performance evaluations of individualistic cultures.

_Hypothesis 2._ The difference predicted in Hypothesis 1 is explained by managers’ domain-general attitudes toward using social comparisons in performance evaluations, which are in turn explained by managers’ perceptions of domain-general descriptive and injunctive norms of social comparisons.

**Perceived Fairness of Social Comparisons in Collectivistic Versus Individualistic Cultures**

I predict that the effect of social comparisons on employees’ perceptions of procedural fairness may vary across collectivistic and individualistic cultures (Brockner, Wiesenfeld, Siegel, Bobocel, & Liu, 2015). As noted, people in collectivistic cultures are likely to perceive stronger descriptive and injunctive norms of social comparisons and to demonstrate more positive attitudes toward using social comparisons, as compared to people in individualistic cultures. Perceived norms and attitudes regarding social comparisons may in turn determine the extent to which employees perceive procedural fairness when they receive social comparison evaluations.

Attitudes based on cultures norms determine people’s reactions to various events they encounter (Bicchieri, 2006), because they judge what is appropriate based on their perceived norms and attitudes (Schultz et al., 2007). When they encounter situations that clash with those norms and attitudes, they may perceive cultural misfit and develop negative reactions (Brockner,
Chen, Mannix, Kwok, & Skarlicki, 2000). Employees in individualistic cultures are thus predicted to reject the relevance of social comparisons, because comparing multiple people with one another conflicts with what they have observed and experienced. In contrast, employees in collectivistic cultures are more likely to accept the relevance of social comparisons in performance evaluations because the norms in their cultures dictate that social context should be incorporated in any appraisals, leading them to take more favorable attitudes toward social comparisons. Thus, I predict that social comparison evaluations may be perceived to adhere to fairer process criteria in collectivistic cultures than in individualistic cultures. Moreover, I predict that this difference may be explained by employees’ domain-general attitudes toward using social comparisons in performance evaluations, which are in turn explained by perceived domain-general descriptive and injunctive norms of social comparisons.

**Hypothesis 3.** Social comparisons are perceived to be procedurally fairer in collectivistic cultures than in individualistic cultures.

**Hypothesis 4.** The difference predicted in Hypothesis 3 is explained by employees’ domain-general attitudes toward using social comparisons in performance evaluations, which are in turn explained by employees’ perceptions of domain-general descriptive and injunctive norms of social comparisons.

**Overview of Studies**

The present research consists of two parts. In the first part, I examine the extent to which social comparisons are emphasized in performance evaluations across collectivistic versus individualistic cultures. In Study 1, I explore whether employees in collectivistic cultures perceive social comparisons in their performance evaluations to a stronger extent than employees in individualistic cultures. In Study 2, I test whether Chinese managers emphasize social
comparisons to a stronger extent than American managers in providing performance evaluations to their employees. I also test the hypothesized mechanism, examining whether the differences between Chinese versus American managers are explained by their perceived domain-general norms of social comparisons and domain-general attitudes toward using social comparisons in performance evaluations.

In the second part, I investigate employees’ reactions to social comparisons that are used to evaluate their performance. In Study 3, I experimentally manipulate social comparisons versus goal comparisons to test whether Chinese employees perceive higher procedural fairness from social comparison evaluations as compared to American employees. In doing so, I differentiate the valence of the evaluations to confirm that the hypothesized differences between Chinese versus American employees emerge in both positive and negative evaluations. In Study 4, I examine the mechanism of the difference observed in Study 3. I test whether Chinese employees perceive higher procedural fairness from social comparisons than American employees because of their perceived domain-general norms and domain-general attitudes toward using social comparisons in performance evaluations.

**Study 1**

In Study 1, I measured to what extent social comparisons were emphasized in collectivistic versus individualistic cultures by collecting data from business professionals who recently received performance evaluations from their workplaces. I investigated two different types of social comparisons. The first type was peer comparisons, which captured the extent to which employees’ performance was discuss in direct comparisons with the performance of their coworkers. The second was ranking comparisons, which captured the extent to which employees’ performance was discussed in terms of their positions in performance rankings.
To ensure that the differences between collectivistic versus individualistic cultures were observed specifically in social comparisons but not in other types of comparisons, I also measured the extent to which participants’ performance was compared with their goals (i.e., goal comparisons; Locke & Latham, 2002) and their own past performance (i.e., temporal comparisons; Albert, 1977).

Method

Sample. I collected data from 557 full-time MBA students at Columbia University. The mean age of the participants was 27.7 ($SD = 2.2$) and 43% were female. Participants identified their race with the following frequencies: 53% White, 32% Asian, 6% Latino, 5% Black, and 2% two or more races. Five participants (1%) did not report their race.

Measures. Before participants answered questions on social comparisons and procedural fairness, they were asked to think about the performance evaluation they had recently received. To make their memories of the performance evaluation more salient, they were also asked to provide the details of the evaluation in terms of when it took place, how long it was, who the evaluator was, and whether the evaluation was given in a live meeting.

After answering these questions, participants were asked to rate the extent to which peer comparisons and ranking comparisons were used in their performance evaluations. The items for peer and ranking comparisons were, “My performance was compared to that of a few particular coworkers” and “My performance was stacked up against comparable employees in a ranking,” respectively. The items for goal and temporal comparisons were, “My performance was compared to objectives that had been set for me” and “My performance was compared to my past performance,” respectively. I used a 5-point Likert-type scale ranging from 1 (“not at all”) to 5 (“very much”) for all items.
The information regarding participants’ work experiences were drawn from their profiles on LinkedIn.com. I coded the geographic locations of their past workplaces to find where participants received the evaluations that they referred to. Overall, they received the evaluations from 40 different countries. Among them, countries located in East Asia and South East Asia were categorized into collectivistic cultures: China \((n = 27)\), Singapore \((n = 11)\), Taiwan \((n = 6)\), South Korea \((n = 5)\), Japan \((n = 4)\), Mongolia \((n = 2)\), the Philippines \((n = 2)\), Thailand \((n = 3)\), and Vietnam \((n = 1)\). Countries located in Western Europe, Southern Europe, and Northern Europe were categorized into individualistic cultures: The United States \((n = 360)\), the United Kingdom \((n = 14)\), Spain \((n = 8)\), Canada \((n = 6)\), the Netherlands \((n = 2)\), Germany \((n = 1)\), Italy \((n = 1)\), and Norway \((n = 1)\). \(^1\)

**Results**

Among participants in the initial dataset, 41 of them were excluded for the following reasons: the coding of geographic locations was impossible because they did not provide sufficient information on their LinkedIn profiles \((n = 26)\); they reported that they had not received performance evaluations \((n = 6)\); they had no profiles on LinkedIn.com \((n = 6)\); they did not provide their names and thus could not be found on LinkedIn.com \((n = 3)\). After excluding these individuals, the final dataset contained the responses from 452 individuals.

**Social comparisons.** Table 1 reports descriptive statistics and correlations of the study.

---

\(^1\) Countries not included in the main analyses were from Africa (Angola, Congo, Kenya, Mauritius, Nigeria, South Africa), Central and South America (Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Panama, Peru), Eastern Europe (Georgia, Hungary, Russia), the Middle East (Israel, Jordan, Qatar, Turkey, the United Arab Emirates), and South Asia (India).
variables. Table 2 reports the means and standard deviations of different types of comparisons in collectivistic versus individualistic cultures (see Figure 1 for a graphical depiction). As predicted, participants from collectivistic cultures perceived higher levels of peer comparisons in their performance evaluations ($M = 2.80, SD = 1.34$) than those from individualistic cultures ($M = 2.21, SD = 1.34; d = .44, t = 3.16, p = .002, 95\% CI = [.221, .950])). Similarly, participants from collectivistic cultures perceived higher levels of ranking comparisons in their performance evaluations ($M = 3.52, SD = 1.30$) than those from individualistic cultures ($M = 3.08, SD = 1.59; d = .28, t = 2.04, p = .042, 95\% CI = [.016, .864])).

**Other types of comparisons.** To confirm that these results reflected differences specifically in perceptions of social comparisons but not in other types of comparisons, I explored whether there were differences between collectivistic and individualistic cultures in goal and temporal comparisons. Participants did not perceive significantly different levels of goal comparisons in collectivistic cultures ($M = 3.72, SD = 1.15$) versus individualistic cultures ($M = 3.91, SD = 1.21; d = -.16, t = -1.13, p = .257, 95\% CI = [-.516, .138])). Similarly, there was no significant difference of temporal comparisons between collectivistic cultures ($M = 3.20, SD = 1.30$) and individualistic cultures ($M = 3.37, SD = 1.36; d = -.13, t = -.91, p = .366, 95\% CI = [-.539, .199])).

**Discussion**

The results suggested that there were significant differences in perceptions of social comparisons in collectivistic and individualistic cultures. Both peer and ranking comparisons were perceived to a stronger extent in collectivistic cultures than in individualistic cultures. In contrast, there was no significant difference regarding how much employees’ performance was compared with their goals or their own past performance in collectivistic versus individualistic
cultures. Thus, the differences between collectivistic and individual cultures emerged specifically in the domains of social comparisons, not in any types of comparisons.

**Study 2**

A strength of Study 1 was that the data were drawn from performance evaluations that participants received in their workplaces. Different levels of social comparisons between collectivistic versus individualistic cultures were observed in performance evaluations that participants actually experienced from their workplaces. However, there were issues to be addressed to clearly demonstrate that social comparisons are emphasized to a stronger extent in performance evaluations of collectivistic cultures than in performance evaluations of individualistic cultures.

First, in Study 1, participants from collectivistic cultures might have spontaneously inferred social comparisons when their managers were not explicitly comparing their performance against their coworkers’ performance (D. R. Ames, 2004). Second, I did not measure the valence of the evaluations. Therefore, it was left unknown whether social comparisons were used to compliment employees for their achievement or to criticize them for their lack of success (Dunn et al., 2012). Third, I did not have the evidence regarding the psychological mechanism of the effects.

To account for these issues, in Study 2 I had participants in collectivistic and individualistic cultures *provide* performance evaluations and measured how much they would emphasize social comparisons in providing the evaluations. I also varied the valence of the evaluations to explore whether social comparisons were more emphasized in positive or negative evaluations. Finally, I explored the mediating mechanisms. I predicted that the extent to which
participants in collectivistic versus individualistic cultures emphasize social comparisons in specific performance evaluations is explained by their domain-general attitudes toward using social comparisons, which may be further explained by their perceptions domain-general norms of social comparisons. I ran the study in two countries, China and the United States, representing collectivistic and individualistic cultures, respectively.

Method

Sample. I originally recruited 141 Chinese participants from Zhubajie.com and 120 American participants from Amazon.com’s Mechanical Turk (Buhrmester, Kwang, & Gosling, 2011). To minimize the possibility that American participants may draw on their collectivistic mindsets (Hong, Morris, Chiu, & Benet-Martínez, 2000), I made it clear that only those without Asian racial background could participate in the study. All participants reported that they had experiences as full-time employees in organizations. The mean age of participants was 34.9 (SD = 9.8) and 59% were female. All Chinese participants identified themselves as Asian. There were two American participants who identified themselves as Asian at the end of the study. These participants were thus excluded from the data. The remaining American participants identified their race with the following frequencies: 78% White, 8% Black, 7% Latino, 1% Native American, and 6% two or more races.

Procedures and measures. Participants were asked to imagine that they were managers of sales employees in a mid-sized firm located in China or the United States. They thought that they had been directly supervising three employees (Luo Gang, Shi Qiang, Bai Feng in China; __________________________

2 The exclusion of these participants did not change the results reported in this study in terms of their statistical significance.
Michael D. Ross, James E. Smith, Richard M. Blackburn in the United States). Then, participants were asked to provide performance feedback to one of the three employees based on their performance reports. In the reports, employees’ performance was assessed in four domains: productivity, customer satisfaction, timeliness, and test scores. In all these domains, Luo Going and Michael D. Ross had the highest scores among the three employees. In contrast, Bai Feng and Richard M. Blackburn had the lowest scores across all four domains. The order in which participants viewed the three performance reports was randomly determined. See Figure 2 for the performance reports of the employees with the highest and lowest performance (English version).

Evaluation valence was manipulated by having participants provide performance feedback to the employee with the highest performance (Positive: Luo Gong or Michael D. Ross) versus the one with the lowest performance (Negative: Bai Feng or Richard M. Blackburn). Social comparisons were measured by asking participants how much they would emphasize different aspects of performance demonstrated by the employee (“X” hereafter). There were four items to measure social comparisons: “X’s productivity relative to the other two employees,” “X’s customer satisfaction relative to the other two employees,” “X’s timeliness relative to the other two employees,” and “X’s test score relative to the other two employees.”

Similar to Study 1, I sought to test whether the difference between collectivistic and individualistic cultures emerged specifically in the domain of social comparisons, but not in other domains of performance evaluations. Therefore, I also asked participants how much they would emphasize the absolute levels of the employee’s performance in the same four aspects of performance, using the following items: “The absolute level of X’s productivity,” “The absolute level of X’s customer satisfaction,” “The absolute level of X’s timeliness,” “The absolute level of
X’s test score.” The items on social comparisons and absolute levels were measured using a 7-point Likert scale ranging from 1 “not at all” to 7 “very much.”

I measured the hypothesized mediators for cultural differences in the extent to which participants in collectivistic versus individualistic cultures would emphasize social comparisons in specific evaluations they were asked to deliver. I tested whether participants’ domain-general attitudes toward using social comparisons and perceived domain-general descriptive and injunctive norms of social comparisons worked as the mediators. To measure these variables, I developed items that captured participants’ attitudes and perceived norms regarding social comparisons across six different domains: employees, students, children, young athletes, music lessons, and dance classes. I averaged participants’ responses across these six domains to generate the scores for participants’ domain-general attitudes toward using social comparisons and perceived domain-general norms of social comparisons. The full list of the items that I used can be found in the Appendix. These items were measured using a 5-point Likert scale ranging from 1 “not at all” to 5 “very much,” with “n/a” as an option. In addition to these items, I included one item for attention check.

Results

Social comparisons. There were 14 participants who failed to pass the attention check. These participants were excluded from the analysis. Table 3 reports descriptive statistics and intercorrelations of the study variables. Table 4 reports the means and standard deviations of social comparisons and absolute levels in each condition (see Figure 3 for a graphical depiction).

3 The exclusion of these participants did not change the results reported in this study in terms of the statistical significance.
The extent to which participants would emphasize social comparisons in their performance feedback (i.e., the average score of the four items on social comparisons) was submitted to a 2 (Country: China vs. the United States) X 2 (Evaluation valence: Positive vs. Negative) analysis of variance. The results demonstrated that the extent to which participants would emphasize social comparisons was significantly affected by the country ($F = 11.90, p < .001$). The evaluation valence did not have a significant effect ($F = 2.58, p = .110$). There was also a significant interaction between the country and evaluation valence ($F = 9.90, p = .002$).

Overall, Chinese participants reported that they would emphasize social comparisons ($M = 4.92, SD = 1.32$) to a stronger extent than American participants ($M = 4.27, SD = 1.69; d = .43$). The significant interaction suggested that this effect was stronger in the positive evaluation condition (simple effect $b = 1.25, t = 4.64, p < .001$, 95% CI $= [.718, 1.775]$) than in the negative evaluation condition (simple effect $b = .07, t = .25, p = .800$, 95% CI $= [-.449, .583]$).

**Absolute levels.** I also analyzed the extent to which participants would emphasize the absolute levels of employees’ performance (i.e., the average score of the four items on absolute levels) using a 2 (Country: China vs. the United States) X 2 (Evaluation valence: Positive vs. Negative) analysis of variance. The effect of the country was not significant ($F = .74, p = .390$) whereas the effect of evaluation valence was significant ($F = 15.57, p < .001$). The interaction was also significant ($F = 4.41, p = .037$).

These results indicated there was no significant overall difference between the extent to which Chinese participants would emphasize the absolute levels of employees’ performance ($M = 5.21, SD = .95$) versus the extent to which American participants would emphasize the absolute levels ($M = 5.31, SD = .97; d = -.10$). The significant interaction suggested that Chinese participants would emphasize absolute levels to a weaker extent than American participants in
the negative evaluation condition (simple effect $b = -.35, t = -2.13, p = .035, 95\% \text{ CI} = [-.682, -.026]$) whereas the opposite (but nonsignificant) pattern emerged in the positive evaluation condition (simple effect $b = .15, t = .86, p = .391, 95\% \text{ CI} = [-.189, .483]$).

**Mediation analyses.** Table 5 reports the values related to the hypothesized mediators. I use the average score of perceived domain-general descriptive norms and perceived domain-general injunctive norms, based on Schultz and colleagues’ suggestions (2007) that both of them capture essential components of norms. The hypothesized mediators demonstrated significant differences in the predicted direction. Chinese participants reported more favorable domain-general attitudes toward using social comparisons ($M = 3.51, SD = .78$) than American participants ($M = 2.63, SD = 1.06; d = .87, t = 7.52, p < .001, 95\% \text{ CI} = [.650, 1.111]$). Chinese participants also reported higher levels of perceived domain-general norms of social comparisons ($M = 3.99, SD = .58$) than American participants ($M = 3.55, SD = .80; d = .61, t = 5.01, p < .001, 95\% \text{ CI} = [.267, .614]$).

Finally, I conducted the mediation analysis, using domain-general attitudes as the mediator in the relationship between the country and social comparisons. A bootstrap analysis with 5,000 resampling demonstrated a significant indirect effect (bias-corrected 95\% CI: [.247, .675]). Then I examined whether perceived domain-general norms further explained the link between the country and domain-general attitudes (i.e., country—domain-general norms—domain-general attitudes—social comparisons). A significant indirect effect was found from a bootstrap analysis

---

4 The results did not change in terms of their statistical significance when I used a separate score for perceived domain-general descriptive and injunctive norms.
with 5,000 resampling (bias-corrected 95% CI: [.063, .256]).

**Discussion**

I found that social comparisons were emphasized to a stronger extent in performance evaluations of China as compared to the United States. This tendency was particularly strong when people could provide positive evaluations. There was no significant overall difference between Chinese versus American participants in the extent to which they would emphasize the absolute levels of employees’ performance. The mediation analyses indicated that the difference between the extent to which Chinese and American participants would emphasize social comparisons in performance evaluations was explained by their domain-general attitudes toward social comparisons, which were further predicted by perceived domain-general descriptive and injunctive norms of social comparisons.

**Study 3**

In the first two studies, I tested whether social comparisons in performance evaluations were emphasized to a stronger extent in collectivistic cultures as compared to individualistic cultures. Both when participants were on the receiving and giving ends of performance evaluations, social comparisons were emphasized to a stronger extent in collectivistic cultures than in individualistic cultures. In Studies 3 and 4, I tested a related but distinct question: how employees *react* when they receive social comparison evaluations from their managers. The

---

5 When the order of the domain-general attitudes and perceived domain-general norms were switched (i.e., country—domain-general attitudes—domain-general norms—social comparisons), the indirect effect was nonsignificant (bias-corrected 95% CI: [-.101, .097]).
specific dependent variable was perceptions of procedural fairness (Colquitt, 2001). I predicted that employees in collectivistic cultures may perceive higher procedural fairness from social comparison evaluations than employees in individualistic cultures. I examined how social comparisons (which compare employees’ performance to their coworkers’ performance), as compared to goal comparisons (which compare employees’ performance to their performance goal), influenced perceptions of procedural fairness. The specific type of social comparisons that I used in Study 3 was peer comparisons (i.e., direct comparisons between employees’ performance to their coworker’s performance).

**Method**

**Sample.** I collected data from 315 Chinese online participants via Sojump.com and 252 American online participants via Amazon.com’s Mechanical Turk platform. For American participants, for the same reason as in Study 2, I made it clear that only those without Asian racial background could participate in the study. The mean age of the participants was 33.9 (SD = 9.0) and 50% of them were female. All Chinese participants identified themselves as Asian. Five American participants reported that they were Asian at the end of the study and thus were excluded from the data. The remaining American participants identified their race with the following frequencies: 84% White, 6% Black, 3% Latino(a), 1% Native American, and 6% two or more races.

**Procedures.** Participants read vignettes that asked them to imagine a situation involving a performance evaluation. After reading the vignettes, participants were asked to share their

6 The exclusion of these participants did not change the results reported in this study in terms of the statistical significance.
reactions using the items that I report below. These vignettes were originally developed in English and then were translated to Chinese. To ensure that the translation process was successful, the original translation was back-translated to English by a bilingual research assistant. Based on this version, minor changes were made to confirm that the English and Chinese versions were parallel. I took similar procedures for the translation in Study 4.

In the vignette, participants were asked to imagine that they had been working in a mid-sized firm in China or the United States as sales employees for 2 years. It was further explained that they were members of a work team that consisted of their manager and employees including themselves. They were told that the performance evaluation in their company took place every 6 months, and they were receiving the evaluation from their manager based on how they had performed during the past 6 months. Subsequently, participants were randomly given one of the four versions of performance evaluations based on a 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) experimental design.

Participants in the positive evaluation conditions were first given an overall positive statement: “Generally, your performance is showing that you are doing a very good job. You’ve proven your strengths in many aspects. Of course, you might need some minor adjustments for the things you are struggling with. But on balance, your performance is strong.” After this statement, Chinese (American) participants in the positive social comparison condition read, “I’m saying this because, you have achieved better results than Zhao Di (Casey), the other member of your team who started at the same time as you. You have sold more units than Zhao Di (Casey). I believe the best way to evaluate performance is comparing it against the performance of comparable employees. So I would say you’ve done a good job.” In contrast, participants in the positive goal comparison condition read, “I’m saying this because you have
attained the performance goals set by the firm at the beginning of this period. You have sold more units than the goal that was set for you. I believe the best way to evaluate performance is comparing it against the employees’ goals. So I would say you’ve done a good job.”

In contrast, participants in the negative evaluation conditions were first given an overall negative statement: “Generally, your performance is showing that you are not doing well enough. There are several aspects in which you are falling short. Of course, there are certain areas where you are doing okay. But on balance, you need to improve.” After this statement, Chinese (American) participants in the negative social comparison condition further read, “I’m saying this because, you have not achieved better results than Zhao Di (Casey), the other member of your team who started at the same time as you. You have not sold as many units as Zhao Di (Casey). I believe the best way to evaluate performance is comparing it against the performance of comparable employees. So I would say you need to work harder.” In contrast, participants in the negative goal comparison condition read, “I’m saying this because you have not attained the performance goals set by the firm at the beginning of this period. You have not sold as many units as the goal that was set for you. I believe the best way to evaluate performance is comparing it against the employees’ goals. So I would say you need to work harder.”

After reading the scenarios, participants were asked to write 4 to 5 sentences to describe how they might have felt after they received the evaluations. Then participants completed items on their perceptions of procedural fairness, positive valence, and negative valence.

**Measures.** Perceptions of procedural fairness were measured by 3 items from Colquitt (2001), “the evaluation upheld ethical and moral standards,” “the evaluation was free of bias,” and “the evaluation was based on accurate information.” To check whether social versus goal comparisons affected participants’ perceptions of how positively or negatively their performance
was appraised, I measured their perceptions of positive and negative valence (“How much did the evaluation consider the positive side of your performance?” and “How much did the evaluation consider the negative side of your performance?”). All items used a 5-point Likert scale ranging from 1 (“not at all”) to 5 (“very much”).

Results

Perceptions of positive and negative valence. Table 6 reports descriptive statistics and intercorrelations of the study variables. First, I tested whether social or goal comparisons affected perceptions of positive and negative valence, and whether this effect differed between China and the United States. Thus, perceptions of positive and negative valence were submitted to 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) X 2 (Country: China vs. the United States) analyses of variance.

For perceptions of positive valence, there was the predicted main effect of evaluation valence ($F = 654.76, p < .001$) suggesting that participants perceived higher levels of positive valence in the positive evaluation condition ($M = 4.23, SD = .77$) than in the negative evaluation condition ($M = 3.49, SD = 1.02$). There was a significant main effect of the country ($F = 70.02, p < .001$), suggesting Chinese participants perceived higher levels of positive valence ($M = 3.61, SD = 1.07$) than American participants ($M = 3.04, SD = 1.40$). I found no significant main effect of the condition ($F = 2.42, p = .121$) and no significant interaction between the condition and country ($F = .18, p = .673$).

Similar results were found for perceptions of negative valence. There was the predicted main effect of evaluation valence ($F = 562.33, p < .001$) suggesting that participants perceived lower levels of negative valence in the positive evaluation condition ($M = 2.33, SD = .88$) than in the negative evaluation condition ($M = 4.06, SD = .97$). There was a significant main effect of the
country ($F = 36.46, p < .001$), suggesting Chinese participants perceived lower levels of negative valence ($M = 3.01, SD = 1.13$) than American participants ($M = 3.45, SD = 1.38$). The main effect of the condition ($F = .49, p = .483$) and the interaction between the condition and country were nonsignificant ($F = 2.13, p = .145$).

**Perceptions of procedural fairness.** Table 7 reports the means and standard deviations of procedural fairness in each condition (see Figure 4 for a graphical depiction). Perceptions of procedural fairness were submitted to a 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) X 2 (Country: China vs. the United States) analysis of variance. The results demonstrated significant main effects of all three variables (Condition $F = 56.93, p < .001$; Evaluation valence $F = 52.63, p < .001$; Country $F = 31.47, p < .001$).

The main effect of the condition suggested that participants perceived lower procedural fairness in the social comparison condition ($M = 3.36, SD = 1.02$) than in the goal comparison condition ($M = 3.91, SD = .83; d = -.57$). The main effect of evaluation valence suggested that participants perceived higher procedural fairness when they received positive evaluations ($M = 3.88, SD = .87$) than when they received negative evaluations ($M = 3.38, SD = .99; d = .51$). The main effect of the country suggested that Chinese participants perceived higher procedural fairness ($M = 3.81, SD = .82$) than American participants ($M = 3.40, SD = 1.09; d = .42$).

Importantly, the main effects of the condition and country were qualified by a significant interaction between them ($F = 22.30, p < .001$). In contrast, the interaction between the condition and evaluation valence ($F = 3.26, p = .072$), the interaction between the country and evaluation valence ($F = 2.18, p = .141$), and the three-way interaction ($F = .00, p = .967$) did not reach statistical significant at .05 confidence level.

The simple effect analysis of the interaction between the condition and country suggested
that in the social comparison condition, Chinese participants perceive significantly higher levels of procedural fairness ($M = 3.69, SD = .88$) than American participants ($M = 2.96, SD = 1.04$; simple effect $b = .73, t = 6.84, p < .001, 95\% CI = [.520, .939]$). In contrast, there was no significant difference in perceptions of fairness from the goal comparison condition across Chinese participants ($M = 3.94, SD = .74$) and American participants ($M = 3.87, SD = .94$; simple effect $b = .07, t = .64, p = .522, 95\% CI = [-.144, .283]$). The nonsignificance of the three-way interaction among the independent variables indicated that the interaction pattern between the condition and country emerged in both positive and negative evaluations. The results reported in this study did not change in terms of their statistical significance when perceptions of positive and negative valence were included in the model as control variables.

**Discussion**

Social comparison evaluations were perceived to be procedurally fairer by Chinese participants as compared to American participants. In contrast, Chinese and American participants did not perceive significantly different levels of procedural fairness from goal comparison evaluations. These patterns were found in both positive and negative evaluation conditions, and they were not explained by perceptions of how positive and negative the evaluations were.

**Study 4**

Study 3 demonstrated that Chinese participants, as compared to American participants, perceived higher levels of procedural fairness when social comparisons were used to evaluate their performance. The specific type of social comparisons used in Study 3 was peer comparisons (Levine & Green, 1984), which directly compared an employee’s performance against their
coworkers’ performance. In Study 4, I used the other type of social comparisons—ranking comparisons—and had participants receive information regarding their positions in performance rankings (Garcia & Tor, 2007).

I also examined whether the social comparison evaluations influenced another domain of fairness that captures employees’ experiences during the evaluation processes: interpersonal fairness (Bies, 1987; Colquitt, 2001). Interpersonal fairness is defined as the extent to which employees receive respectful, dignifying, and polite treatment. I tested whether social comparisons significantly affected participants’ perceptions of interpersonal fairness, and whether the effect of social comparisons on perceptions of interpersonal fairness varied depending on the country.

Most importantly, I tested the mediating mechanism that may explain why Chinese and American participants perceived different levels of fairness from social comparison evaluations. I predicted that the different levels of procedural fairness perceived by Chinese versus American participants in response to social comparison evaluations are explained by their domain-general attitudes toward using social comparisons, which are further explained by perceived domain-general norms of social comparisons.

**Method**

**Sample.** I collected data from 270 Chinese online participants via Sojump.com and 227 American online participants via Amazon.com’s Mechanical Turk platform. Again, for American participants, I emphasized that only those without Asian racial background could participate. The mean age of the participants was 34.7 (SD = 9.6) and 56% of them were female. All Chinese participants identified themselves as Asian. There were three American participants who reported
that they were Asian at the end of the study and I excluded them from the data. The remaining American participants identified their race with the following frequencies: 85% White, 5% Black, 4% Latino(a), .5% Native American, and 5% two or more races.

**Procedures.** Participants were given vignettes that described a context that was almost identical as the one described in Study 3; they imagined that they were sales employees receiving performance evaluations from their manager. The only difference in the description of the context was that in Study 4, participants were told that their evaluations took place every 3 months (not every 6 months as in Study 3). Subsequently, participants were randomly given one of the four performance evaluations based on a 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) design.

In the positive evaluation condition, participants were first given the manager’s compliment for their performance: “Generally, your performance is showing that you are doing a very good job. You have done well in many aspects of your performance. Of course, you might need some minor adjustments for the things you were struggling with. But on balance, this was a very good quarter for you.” After this statement, participants in the positive social comparison condition read, “I’m saying this because, you ranked in the upper half of your team's 5 employees. I believe the best way to evaluate an employee's performance is comparing it against the performance of other employees. It's a good sign that you are in the upper half of the team.” In contrast, participants in the positive goal comparison condition read, “I’m saying this because, you attained the performance goals set by the firm at the beginning of this period. I believe the

7 The exclusion of these participants did not change the results reported in this study in terms of the statistical significance.
best way to evaluate performance is comparing it against the employees’ performance goals. It's a good sign that you achieved your goal.”

In the negative evaluation condition, participants were given the manager’s criticism of their performance: “Generally, your performance is showing that you are not doing a very good job. You struggled in some aspects of your performance. Of course, there were certain areas in which you were doing okay. But on balance, this was not a very good quarter for you.” After this statement, participants in the negative social comparison condition further read, “I’m saying this because, you ranked in the lower half of your team's 5 employees. I believe the best way to evaluate an employee's performance is comparing it against the performance of other employees. It's not a good sign that you are in the lower half of the team.” In contrast, participants in the negative goal comparison condition read, “I’m saying this because, you did not attain the performance goals set by the firm at the beginning of this period. I believe the best way to evaluate performance is comparing it against the employees’ performance goals. It's not a good sign that you did not achieve your goal.”

As in Study 3, after they were given the evaluation statement, participants wrote 4 to 5 sentences to describe how they might have felt after the evaluations. Following this, participants completed items on perceptions of procedural fairness and interpersonal fairness, perceptions of positive and negative valence, domain-general attitudes toward using social comparisons, and perceived domain-general descriptive and injunctive norms of social comparisons.

**Measures**. The same items as in Study 3 were used to measure perceptions of procedural fairness, positive valence, and negative valence. Perceptions of interpersonal fairness were measured using 3 items from Colquitt (2001): “The manager treated you with respect,” “The manager treated you with dignity,” and “The manager treated you in a polite manner.” The same
items as in Study 2 were used to measure participants’ domain-general attitudes regarding social comparisons and perceived domain-general norms of social comparisons.

**Results**

**Perceptions of positive and negative valence.** Table 8 reports descriptive statistics and intercorrelations of the study variables. I first investigated whether perceptions of positive and negative valence were significantly affected by the type of comparisons or countries. Perceptions of positive and negative valence were thus analyzed by 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) X 2 (Country: China vs. the United States) analyses of variance.

For perceptions of positive valence, I found the predicted main effect of evaluation valence suggesting that participants perceived higher levels of positive valence in the positive evaluation condition \( (M = 4.24, SD = .68) \) than in the negative evaluation condition \( (M = 2.06, SD = .73; F = 1319.17, p < .001) \). The results suggested that the condition and country had also significant main effects on perceptions of positive valence (Condition \( F = 20.02, p < .001 \); Country \( F = 34.70, p < .001 \)). Participants in the social comparison condition perceived lower levels of positive valence \( (M = 3.01, SD = 1.29) \) than participants in the goal comparison condition \( (M = 3.28, SD = 1.30) \), and Chinese participants perceived higher levels of positive valence \( (M = 3.29, SD = 1.14) \) than American participants \( (M = 2.96, SD = 1.45) \). Importantly, there was no significant interaction between the condition and country \( (F = 2.09, p = .149) \).

For perceptions of negative valence, again I observed the predicted main effect of evaluation valence suggesting that participants in the positive evaluation condition perceived lower levels of negative valence \( (M = 2.28, SD = .73) \) than participants in the negative evaluation condition \( (M = 4.34, SD = .82, F = 1009.87, p < .001) \). In addition, there was a significant main
effect of the country ($F = 54.75$, $p < .001$), suggesting that Chinese participants perceived lower levels of negative valence ($M = 3.11$, $SD = 1.16$) than American participants ($M = 3.56$, $SD = 1.39$). The main effect of the condition on negative valence was not significant ($F = .44$, $p = .507$), and the interaction between the condition and country was not significant, either ($F = .65$, $p = .421$).

**Perceptions of procedural fairness.** Table 9 reports the means and standard deviations of procedural fairness in each condition (see Figure 5 for a graphical depiction). Perceptions of procedural fairness were submitted to a 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) X 2 (Country: China vs. the United States) analysis of variance. The results suggested that all three independent variables had significant main effects (Condition $F = 25.14$, $p < .001$; Evaluation valence $F = 124.69$, $p < .001$; Country $F = 5.83$, $p = .016$).

The main effect of the condition suggested that participants who received social comparison evaluations perceived lower procedural fairness ($M = 3.39$, $SD = 1.02$) than participants who received goal comparison evaluations ($M = 3.77$, $SD = .88$; $d = -.39$). The main effect of evaluation valence suggested that participants who received positive evaluations perceived higher procedural fairness ($M = 4.01$, $SD = .72$) than participants who received negative evaluations ($M = 3.14$, $SD = .99$; $d = .90$). The main effect of the country suggested that Chinese participants ($M = 3.66$, $SD = .89$) perceived higher procedural fairness than American participants ($M = 3.47$, $SD = 1.05$; $d = .20$).

More importantly, the main effects of the condition and country were qualified by a significant interaction between them ($F = 5.38$, $p = .021$). In contrast, the interaction between the condition and evaluation valence ($F = .23$, $p = .631$), the interaction between the country and
evaluation valence ($F = .90, p = .344$), and the three-way interaction were nonsignificant ($F = .29, p = .591$).

The simple effect analysis of the interaction between the condition and country suggested that, in the social comparison condition, Chinese participants perceived higher levels of procedural fairness ($M = 3.55, SD = .94$) than American participants ($M = 3.20, SD = 1.08$; simple effect $b = .35, t = 2.96, p = .003, 95\% CI = [.118, .586]$). In contrast, in the goal comparison condition, there was no significant difference between perceptions of procedural fairness from Chinese participants ($M = 3.76, SD = .84$) versus American participants ($M = 3.78, SD = .93$; simple effect $b = -.02, t = -.14, p = .889, 95\% CI = [-.259, .225]$). Similar to Study 3, the three-way interaction among the independent variables was nonsignificant, indicating that similar interactions between the condition and country emerged in positive and negative evaluation conditions. These results did not change in terms of their statistical significance when perceptions of positive and negative valence were included as control variables.

**Perceptions of interpersonal fairness.** Participants’ perceptions of interpersonal fairness were submitted to the 2 (Condition: Social comparison vs. Goal comparison) X 2 (Evaluation valence: Positive vs. Negative) X 2 (Country: China vs. the United States) analysis of variance. The results demonstrated different patterns from the results for procedural fairness.

There were three effects that were significant. First, there was a main effect of the condition ($F = 4.37, p = .027$), suggesting that participants perceived lower interpersonal fairness from social comparison evaluations ($M = 3.51, SD = 1.20$) than from goal comparison evaluations ($M = 3.70, SD = 1.06; d = -.17$). Second, there was a strong main effect of evaluation valence ($F = 213.14, p < .001$), suggesting that participants perceived higher interpersonal fairness from positive evaluations ($M = 4.23, SD = .71$) than negative evaluations ($M = 2.99, SD$}
Finally, there was a significant interaction between evaluation valence and country predicting interpersonal fairness ($F = 12.35, p < .001$). The simple effect analysis suggested that, in the positive evaluation condition, Chinese participants perceived significantly lower interpersonal fairness ($M = 4.09, SD = .71$) than American participants ($M = 4.40, SD = .68$; simple effect $b = -.31, t = -2.60, p = .010, 95\% CI = [-.550, -.076]$). In contrast, in the negative evaluation condition, Chinese participants perceived significantly higher interpersonal fairness ($M = 3.12, SD = 1.07$) than American participants ($M = 2.82, SD = 1.21$; simple effect $b = .29, t = 2.44, p = .015, 95\% CI = [.058, 530]$).

Importantly, there was no significant interaction between the condition and country ($F = .149, p = .699$). This effect indicated that participants who received social comparison evaluations perceived lower interpersonal fairness than those who received goal comparison evaluations, but the effect did not significantly differ across China and the United States.

**Mediation analyses.** Table 10 reports the values related to the hypothesized mediators. Replicating Study 2 results, the hypothesized mediators demonstrated significant differences in the predicted direction. Chinese participants reported more favorable domain-general attitudes toward using social comparisons ($M = 3.57, SD = .74$) than American participants ($M = 2.62, SD = .99; d = .96, t = 12.18, p < .001, 95\% CI = [.799, 1.106]$). Chinese participants also reported stronger perceived domain-general norms of social comparisons ($M = 4.20, SD = .49$) than American participants ($M = 3.48, SD = .84; d = .95, t = 11.86, p < .001, 95\% CI = [.602, .842]$).

Based on these patterns, I examined whether the effect of social comparisons on

8 As in Study 2, I used perceived domain-general norms by averaging the scores of perceived domain-general descriptive norms and perceived domain-general injunctive norms (Schultz et al., 2007).
perceptions of fairness for Chinese versus American participants were explained by their
domain-general attitudes and perceived domain-general norms. To do so, I tested whether the
interaction between the condition and country was mediated by the interaction between the
condition and domain-general attitudes. I tested the mediation effect using the path analysis
approach proposed by Edwards and Lambert (2007). Specifically, using 5,000 bootstrap
resampling, I estimated the indirect effect by computing the product of (a) the main effect of the
country on domain-general attitudes and (b) the interaction effect between the condition and
domain-general attitudes on procedural fairness, controlling for the interaction between the
condition and country (for an article with a similar analytic approach, see Grant & Berry, 2011).
From 5,000 bootstrap resampling, the indirect effect was significant (bias-corrected 95% CI:
[.015, .391]). This result indicated that the interaction effect between the condition and country
on procedural fairness was mediated by the interaction between the condition and domain-
general attitudes. Therefore, I expanded the analysis to test whether perceived domain-general
norms worked as a mediator in the relationship between the country and domain-general
attitudes. The results from 5,000 bootstrap resampling suggested a significant indirect effect
(bias-corrected 95% CI: [.030, .176]). Thus, the mediation effect of domain-general attitudes was
further mediated by perceived domain-general norms.

9 I computed the product of (a) the main effect of the country on perceived domain-general norms,
(b) the main effect of perceived domain-general norms on domain-general attitudes controlling for the
effect of the country, and (b) the interaction effect of the condition and domain-general attitudes on
procedural fairness controlling for the interaction between the condition and country and the interaction
between the condition and perceived domain-general norms.
Discussion

Using rankings to operationalize social comparisons, Study 4 replicated the Study 3 findings that indicated Chinese participants perceived higher levels of procedural fairness when they received social comparison evaluations as compared to American participants. Goal comparisons did not lead to significantly different perceptions of procedural fairness for Chinese versus American participants. Again, this pattern emerged in both positive and negative evaluation contexts.

Perceptions of interpersonal fairness demonstrated different results. Although there was a significant main effect of the condition, the non-significant interaction between the condition and country indicated that Chinese versus American participants demonstrated comparable responses to their evaluations.

Study 4 also provided evidence regarding why the effect of social comparisons on procedural fairness differed in China versus the United States. I found that the difference demonstrated by Chinese versus American participants who received social comparison evaluations was mediated by their general attitudes toward using social comparisons, which were further explained by their perceived norms of social comparisons. As compared to American participants, Chinese participants perceived stronger norms in their societies that accepted using social comparisons to appraise people’s achievements, and thus they demonstrated more favorable attitudes toward using social comparisons. Therefore, when social comparisons were used to appraise their performance, Chinese participants perceived higher levels of procedural fairness than American participants, both in positive and negative evaluations.
General Discussion

In the present research, I examined how people used and reacted to social comparisons in performance evaluations across different cultures. In Study 1, I found that employees perceived social comparisons—both peer comparisons and ranking comparisons—to a stronger extent in collectivistic than individualistic cultures. In Study 2, I focused on two specific countries representing collectivistic and individualistic cultures, China and the United States, to examine the mediating mechanism. As compared with American participants, Chinese participants perceived stronger descriptive and injunctive norms that accepted social comparisons in performance evaluations. Thus, Chinese participants had more favorable attitudes toward using social comparisons to appraise achievements in various domains. These differences in perceived norms and attitudes regarding social comparisons reported by participants in China versus the United States resulted in the differences in the extent to which they emphasized social comparisons in specific performance evaluations they conducted to appraise employees’ performance.

In Studies 3 and 4, I found cultural differences in employees’ reactions to social comparison evaluations they received: Chinese participants perceived higher procedural fairness from social comparison evaluations than American participants, both in positive and negative evaluations. Study 4 results revealed that these differential reactions to social comparison evaluations demonstrated by Chinese versus American participants were again driven by their perceived norms and attitudes toward using social comparisons.

These findings suggest that social comparisons play different roles in different cultures. In collectivistic cultures where people define themselves in terms of their social relationships (Brewer & Chen, 2007; Markus & Kitayama, 1991), assessing how they are performing relative
to others is a critical issue (Chung & Mallery, 1999; White & Lehman, 2005; Wilson & Ross, 2000). In contrast, people in individualistic cultures do not have strong needs to assess themselves relative to other people. Their primary motivation is in understanding the specific details of their individual characteristics and demonstrating these characteristics (Kim & Markus, 1999; Suls & Wan, 1987; Valins & Nisbett, 1972). Therefore, social comparisons may not be considered as a highly relevant component of performance appraisals (Teven & Hanson, 2004).

This does not mean that social comparisons will be emphasized to a stronger extent than any other elements of performance evaluations or lead to extremely high levels of procedural fairness in collectivistic cultures. In fact, in Studies 1 and 2, goal comparisons and absolute levels of employees’ performance were emphasized to a stronger extent than social comparisons in collectivistic cultures. Moreover, in Studies 3 and 4, goal comparisons always led to higher perceptions of fairness than social comparisons in China. However, my findings suggest that (a) social comparisons were emphasized to a stronger extent in collectivistic than individualistic cultures, and (b) the negative effect of social comparisons on procedural fairness was significantly weaker in collectivistic cultures than in individualistic cultures. Thus, relative to employees in individualistic cultures, employees in collectivistic cultures may react more proactively to social comparison evaluations they receive and demonstrate stronger willingness to leverage the evaluations for their future progress (Grant & Ashford, 2008; Leung et al., 2001; Mayer & Davis, 1999). In contrast, employees in individualistic cultures are unlikely to demonstrate such developmental reactions to social comparison evaluations. Based on the strong procedural unfairness they perceive, they may believe that the managers do not provide them with useful information and thus disregard the evaluations (Ilgen, Fisher, & Taylor, 1979).

My findings on interpersonal fairness from Study 4 provide important evidence regarding
the nature of social comparisons’ effects. Specifically, social comparisons had a negative main effect on perceptions of interpersonal fairness, but there was no significant interaction between social comparisons and country: social comparisons were considered to be less respectful than goal comparisons, but this tendency did not significantly differ across China and the United States. Thus, the reason why Chinese versus American participants perceived different levels of procedural fairness from social comparisons may not have been driven by their perceptions of how their managers delivered the evaluations. The actual reason may instead reside in what their managers delivered (i.e., the contents), and the extent to which participants consider the contents to be relevant.

**Theoretical Implications**

Previous research found that social comparisons were emphasized to a stronger extent in collectivistic cultures than in individualistic cultures for the purpose of self-evaluations (Chung & Mallery, 1999; White & Lehman, 2005). The present research not only extends these findings to the setting where people evaluate others, but also reveals how people react when they receive social comparison evaluations from others. My findings further suggest that the cross-cultural differences related to social comparisons in performance evaluations are explained by people’s perceived norms and attitudes regarding social comparisons.

My argument and findings contribute to the literature on contextual and relational views (Andersen & Chen, 2002; Cross et al., 2000). Research has found that people in collectivistic cultures care about their close others to a stronger extent than people in individualistic cultures (Brewer & Chen, 2007). The present findings suggest that people in collectivistic cultures also take relational views in evaluating others: they evaluated others relative to one another. Based on their perceptions of cultural norms and personal attitudes that accept using social comparisons to
appraise achievements and performance, people in collectivistic cultures emphasize social comparisons to a stronger extent in providing performance evaluations than people in individualistic cultures.

Given that social comparisons and competition mutually reinforce each other (Charness et al., 2011; Garcia et al., 2013; Kuhnen & Tymula, 2011), my findings may imply that competition is more prevalent in collectivistic cultures than in individualistic cultures. This interpretation seems inconsistent with the past findings demonstrating that people in collectivistic cultures pursue ingroup harmony (e.g., Triandis, Bontempo, Villareal, Asai, & Lucca, 1988). However, recent evidence from cross-cultural psychology suggested that competition and pursuit of ingroup harmony can coexist in collectivistic cultures. Specifically, Liu and colleagues (2018) demonstrated that people in collectivistic cultures find themselves in situations where they should balance two different and potentially conflicting goals: outperforming other members of their groups and maintaining interpersonal harmony within the groups. Therefore, people in collectivistic cultures compete with one another, but in a more covert manner so they can exert effort to outperform others and maintain ingroup harmony at the surface level. Combined with this evidence, my findings suggest that social comparisons may tap into competitive motivation, and such motivation can be stronger in collectivistic cultures than in individualistic cultures.

Practical Implications

Research has found that perceptions of procedural fairness have significant effects on various organizational outcomes such as job satisfaction, organizational commitment, and employee performance (Colquitt, 2001; Colquitt et al., 2001; Colquitt et al., 2005). Therefore, the present research suggests that social comparisons should be used carefully, especially in individualistic cultures. People in individualistic cultures may not believe that individuals should
be compared with one another, and thus they may not think that social comparison evaluations provide relevant information. Thus, managers in individualistic cultures may want to use different approaches of performance evaluations, such as goal comparisons (Locke & Latham, 2002) or temporal comparisons (Albert, 1977).

In collectivistic cultures, social comparisons may cause less damage on employee morale. Moreover, social comparison evaluations in collectivistic cultures may deliver information that employees consider to be more relevant. However, my results suggest that even in collectivistic cultures, the effects of social comparisons on perceptions of procedural and interpersonal fairness may not be positive. Therefore, when managers in collectivistic cultures use social comparisons to provide performance evaluations, they need to be careful in devising evaluation messages that minimize the potential negative reactions. For example, they may emphasize their willingness to provide support and guidance for the employees’ future development (Wang, Law, Hackett, Wang, & Chen, 2005).

Limitations and Future Research

In Study 1, I collected data from a wide range of countries representing collectivistic and individualistic cultures. However, because Study 1 was conducted in the United States, most participants reported their experiences from individualistic cultures, particularly from the United States. This resulted in dissimilar sample sizes for collectivistic and individualistic cultures. In Studies 2 to 4, I sought to address this issue by recruiting comparable numbers of participants from collectivistic and individualistic cultures. To achieve this goal, I focused on two countries, China and the United States, as the sources of my data. However, future research can collect data from other countries representing collectivistic and individualistic cultures (e.g., Japan for collectivistic cultures and Canada for individualistic cultures).
Relatedly, in Study 1, which was conducted in the United States, the study material was written in English and every participant received the same material. In contrast, Studies 2 to 4 were conducted in different countries (China and the United States) using different languages (Chinese and English, respectively). Therefore, for these studies I took very careful effort to secure the commensurability of the study materials, running multiple rounds of translation and back-translation. Reflecting this effort, participants in Studies 3 and 4 did not perceive different levels of procedural fairness in the goal comparison conditions. However, I cannot completely rule out the possibility that the findings from the present research might have been affected by the different languages that were used in different countries. Future research can induce interdependent versus independent mindsets to bicultural participants and then examine the effects of social comparisons on various outcomes, keeping the language and wording of the study materials constant (Hong et al., 2000).

Lastly, although the materials in Studies 3 to 4 were written so as to clearly manipulate the different types of comparisons, participants in collectivistic versus individualistic cultures may have inferred certain types of comparisons based on the cultural norms within their societies (D. R. Ames, 2004). For example, Chinese participants in the goal comparison condition may have thought that their evaluations compared their performance to their coworkers’ performance. More specifically, when Chinese participants were told that they failed to attain their performance goals, they might have inferred that they were falling behind their coworkers. Although researchers have examined the prevalence of self-initiated social comparisons across different cultures (Chung & Mallery, 1999; White & Lehman, 2005), there is no research on whether social comparisons work as the default mode of interpreting performance evaluations. Therefore, it would be valuable for future research to examine how much social comparisons are
used in people’ interpretation of their performance evaluations and how this tendency is shaped by their cultures.

**Conclusion**

Although the literature on social comparisons has a long history, there has been scant attention on the psychological effects of social comparisons in the workplace context. In the present research, I attempted to fill this gap in two ways. First, I examined to what extent managers emphasized social comparisons to evaluate their employees across different cultures and why. Second, I investigated how employees reacted to social comparison evaluations across different cultures and why. The results revealed that managers in collectivistic cultures emphasized social comparisons to a stronger extent than managers in individualistic cultures to evaluate employees’ performance. Moreover, employees in collectivistic cultures perceived higher procedural fairness when they received social comparison evaluations as compared to employees in individualistic cultures. These two differences in collectivistic and individualistic cultures were mediated by managers’ and employees’ domain-general attitudes toward using social comparisons, which in turn were explained by their perceived domain-general norms of social comparisons. My findings allude to the competitive interpersonal relationships that may exist in collectivistic cultures where people continuously draw on social information to assess themselves and others.
References


Barankay, I. (2012). *Rank Incentives: Evidence from a Randomized Workplace Experiment*. University of Pennsylvania. Retrieved from [http://repository.upenn.edu/bepp_papers/75](http://repository.upenn.edu/bepp_papers/75)


Liu, S. S., Morris, M. W., & Yang, Q. (2018). *Beneath Eastern Harmony: Covert Competition*
and Ingroup Suspicion. Paper presented at the 2018 Annual Meeting of Society for Personality and Social Psychology, Atlanta, GA.


Table 1

*Descriptive Statistics and Intercorrelations from Study 1*

<table>
<thead>
<tr>
<th>Variables</th>
<th>$M$</th>
<th>$SD$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Peer comparisons</td>
<td>2.29</td>
<td>1.35</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Ranking comparisons</td>
<td>3.14</td>
<td>1.56</td>
<td>.46</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>3. Goal comparisons</td>
<td>3.88</td>
<td>1.20</td>
<td>.01</td>
<td>.25</td>
<td>--</td>
</tr>
<tr>
<td>4. Temporal comparisons</td>
<td>3.35</td>
<td>1.35</td>
<td>.10</td>
<td>.05</td>
<td>.31</td>
</tr>
</tbody>
</table>

*Note.* $n = 452$. $r$s greater than .09 are significant at .05 level.
Table 2

*Different Types of Comparisons from Study 1*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Collectivistic</th>
<th>Individualistic</th>
<th>p-value for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer comparisons</td>
<td>2.80 (1.34)</td>
<td>2.21 (1.34)</td>
<td>.002</td>
</tr>
<tr>
<td>Ranking comparisons</td>
<td>3.52 (1.30)</td>
<td>3.08 (1.59)</td>
<td>.042</td>
</tr>
<tr>
<td>Goal comparisons</td>
<td>3.72 (1.15)</td>
<td>3.91 (1.21)</td>
<td>.257</td>
</tr>
<tr>
<td>Temporal comparisons</td>
<td>3.20 (1.30)</td>
<td>3.37 (1.36)</td>
<td>.366</td>
</tr>
</tbody>
</table>

*Note. n = 452. Values in parentheses are standard deviations.*
Table 3

*Descriptive Statistics and Intercorrelations from Study 2*

<table>
<thead>
<tr>
<th>Variables</th>
<th>$M$</th>
<th>$SD$</th>
<th>$\alpha$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Social comparisons</td>
<td>4.92</td>
<td>1.32</td>
<td>.91</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Absolute levels</td>
<td>5.21</td>
<td>.95</td>
<td>.73</td>
<td>.41</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>3. Domain-general attitude</td>
<td>3.51</td>
<td>.78</td>
<td>.86</td>
<td>.36</td>
<td>.08</td>
<td>--</td>
</tr>
<tr>
<td>4. Perceived domain-general descriptive norms</td>
<td>4.02</td>
<td>.60</td>
<td>.75</td>
<td>.20</td>
<td>-.02</td>
<td>.48</td>
</tr>
<tr>
<td>5. Perceived domain-general injunctive norms</td>
<td>3.97</td>
<td>.64</td>
<td>.79</td>
<td>.19</td>
<td>.00</td>
<td>.54</td>
</tr>
</tbody>
</table>

*Note: $n = 247$. Rs greater than .12 are significant at .05 level*
Table 4

*Social Comparisons and Absolute Levels from Study 2*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Social comparisons</th>
<th>Absolute levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>China</td>
<td>United States</td>
</tr>
<tr>
<td>Positive evaluation</td>
<td>5.33 (1.14)</td>
<td>4.08 (1.77)</td>
</tr>
<tr>
<td>Negative evaluation</td>
<td>4.52 (1.37)</td>
<td>4.45 (1.60)</td>
</tr>
<tr>
<td>Mean values</td>
<td>4.92 (1.32)</td>
<td>4.27 (1.69)</td>
</tr>
</tbody>
</table>

*p-value for difference*  
< .001                  .390

*Note.* $n = 247$. Values in parentheses are standard deviations.
Table 5

The Hypothesized Mediators from Study 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Domain-general attitudes</th>
<th>Perceived domain-general norms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>China</td>
<td>United States</td>
</tr>
<tr>
<td>Positive evaluation</td>
<td>3.54 (.78)</td>
<td>2.45 (.99)</td>
</tr>
<tr>
<td>Negative evaluation</td>
<td>3.49 (.79)</td>
<td>2.80 (1.11)</td>
</tr>
<tr>
<td>Mean values</td>
<td>3.51 (.78)</td>
<td>2.63 (1.06)</td>
</tr>
</tbody>
</table>

*p-value for difference*< .001  < .001

*Note. n = 247. Values in parentheses are standard deviations.*
Table 6

*Descriptive Statistics and Intercorrelations from Study 3*

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>α</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Procedural fairness</td>
<td>3.63</td>
<td>.97</td>
<td>.85</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>2. Positive valence</td>
<td>3.36</td>
<td>1.26</td>
<td>--</td>
<td>.51</td>
<td>--</td>
</tr>
<tr>
<td>3. Negative valence</td>
<td>3.20</td>
<td>1.27</td>
<td>--</td>
<td>-.29</td>
<td>-.70</td>
</tr>
</tbody>
</table>

*Note: n = 562. All correlation coefficients were significant at the .05 level.*
### Table 7

*Procedural Fairness from Study 3*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Social comparison</th>
<th>Goal comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>China</td>
<td>United States</td>
</tr>
<tr>
<td>Positive evaluation</td>
<td>3.84 (.82)</td>
<td>3.20 (.95)</td>
</tr>
<tr>
<td>Negative evaluation</td>
<td>3.53 (.91)</td>
<td>2.68 (1.08)</td>
</tr>
<tr>
<td>Mean values</td>
<td>3.69 (.88)</td>
<td>2.96 (1.04)</td>
</tr>
</tbody>
</table>

*p*-value for difference: \(< .001, .522\)

*Note. n = 562. Values in parentheses are standard deviations.*
Table 8

*Descriptive Statistics and Intercorrelations from Study 4*

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>α</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Procedural fairness</td>
<td>3.57</td>
<td>.97</td>
<td>.87</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Interpersonal fairness</td>
<td>3.60</td>
<td>1.14</td>
<td>.94</td>
<td>.79</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Positive valence</td>
<td>3.14</td>
<td>1.30</td>
<td>--</td>
<td>.54</td>
<td>.62</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Negative valence</td>
<td>3.31</td>
<td>1.29</td>
<td>--</td>
<td>-.41</td>
<td>-.48</td>
<td>-.78</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Perceived domain-general descriptive norms</td>
<td>3.91</td>
<td>.77</td>
<td>.76</td>
<td>.06</td>
<td>.01</td>
<td>.03</td>
<td>-.06</td>
<td>.44</td>
<td>--</td>
</tr>
<tr>
<td>7. Perceived domain-general injunctive norms</td>
<td>3.84</td>
<td>.83</td>
<td>.81</td>
<td>.06</td>
<td>.00</td>
<td>.04</td>
<td>-.07</td>
<td>.51</td>
<td>.82</td>
</tr>
</tbody>
</table>

*Note: n = 494. rs greater than .08 are significant at .05 level*
Table 9

*Procedural Fairness from Study 4*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Social comparison</th>
<th>Goal comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>China</td>
<td>United States</td>
</tr>
<tr>
<td>Positive evaluation</td>
<td>3.96 (.74)</td>
<td>3.73 (.83)</td>
</tr>
<tr>
<td>Negative evaluation</td>
<td>3.18 (.95)</td>
<td>2.72 (1.05)</td>
</tr>
<tr>
<td>Mean values</td>
<td>3.55 (.94)</td>
<td>3.20 (1.08)</td>
</tr>
</tbody>
</table>

*p*-value for difference

.003 .889

*Note. n = 494. Values in parentheses are standard deviations.*
Table 10

The Hypothesized Mediators from Study 4

<table>
<thead>
<tr>
<th>Condition</th>
<th>Domain-general attitudes</th>
<th>Perceived domain-general norms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>China</td>
<td>United States</td>
</tr>
<tr>
<td>Positive evaluation</td>
<td>3.60 (.76)</td>
<td>2.83 (.97)</td>
</tr>
<tr>
<td>Negative evaluation</td>
<td>3.54 (.73)</td>
<td>2.40 (.97)</td>
</tr>
<tr>
<td>Mean values</td>
<td>3.57 (.74)</td>
<td>2.62 (.99)</td>
</tr>
</tbody>
</table>

*p-value for difference*  < .001  < .001

*Note. n = 247. Values in parentheses are standard deviations.*
Figure 1. Different types of comparisons from Study 1

Note. Error bars represent standard errors.
<table>
<thead>
<tr>
<th>Category</th>
<th>Score</th>
<th>Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity</td>
<td>107</td>
<td>Productivity</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Customer satisfaction</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Timeliness</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Test score</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total score</td>
<td>387</td>
</tr>
</tbody>
</table>

**Figure 2.** Performance reports used in Study 2 (English version)
Figure 3. Social comparisons and absolute levels from Study 2

Note. Error bars represent standard errors.
Figure 4. Procedural fairness from Study 3

Note. Error bars represent standard errors.
Figure 5. Procedural fairness from Study 4

Note. Error bars represent standard errors.
Appendix

Items used to measure perceived domain-general descriptive norms of social comparisons, domain-general injunctive norms of social comparisons, and domain-general attitudes toward using social comparisons (all items were measured using a scale from 1 “not at all” to 5 “very much,” with “n/a” as an option)

To what extent do you agree with the following statements?

**Employees**

- In my society, it’s common that employees are evaluated in comparison to others.
- In my society, people approve when employees are evaluated in comparison to others.
- I personally believe that employees have to be evaluated in comparison to others.

**Students**

- In my society, it’s common that students are told how they rank in their class.
- In my society, people approve when students are told how they rank in their class.
- I personally believe that students have to be told how they rank in their class.

**Children**

- In my society, it’s common that parents motivate their children by comparing them to other children.
- In my society, people approve when parents motivate their children by comparing them to other children.
I personally believe that parents have to motivate their children by comparing them to other children.

**Young athletes**

In my society, it’s common that coaches in sports programs tell young athletes how good they are relative to their peers.

In my society, people approve when coaches in sports programs tell young athletes how good they are relative to their peers.

I personally believe that coaches in sports programs have to tell young athletes how good they are relative to their peers.

**Music lessons**

In my society, it’s common that a student’s performance in music lessons is compared others who are their age.

In my society, people approve when a student’s performance in music lessons is compared others who are their age.

I personally believe that student’s performance in music lessons has to be compared others who are their age.

**Dance classes**

In my society, it’s common that students’ talent in dance classes is assessed relative to the other students.
In my society, people approve when students’ talent in dance classes is assessed relative to the other students.

I personally believe that students’ talent in dance classes has to be assessed relative to the other students.
CHAPTER 2. SOCIAL AND TEMPORAL COMPARISONS ON FAIRNESS
Introduction

Performance evaluations involve comparison to standards, and the characteristics of the standards determine the consequences of the evaluations (Locke & Latham, 2002). For example, other people’s performance is frequently used as a standard to assess the focal individual’s achievement (Festinger, 1954; Wood, 1989). This process of social comparison has been shown to exert significant effects on the thoughts, feelings, and behaviors of the individuals receiving the evaluations (Tesser, 1988). In fact, social comparison information may even outweigh the effects of objective performance information in determining individuals’ reactions (Klein, 1997).

Noting this possibility, researchers have examined the effects of social comparisons in the context of performance evaluations. For example, upward social comparisons (e.g., “You’re doing worse than your peers”) have been shown to reduce affective trust in peers whereas downward social comparisons (e.g., “You’re doing better than your peers”) reduce cognitive trust in peers (Dunn et al., 2012). In negotiation context, it has been shown that the type of social standards (the counterparts of the negotiations vs. other negotiators in similar situations) can have significant effects on negotiators’ satisfaction (Novemsky & Schweitzer, 2004). Specifically, comparisons with the counterparts’ outcomes were more likely to decrease negotiators’ satisfaction by focusing their attention on the portion they failed to claim. These findings suggest that to whom they are compared can significantly influence people’s experiences of and reactions to performance evaluations.

Another important type of performance standards is how well an individual has done in the past. When one’s own past performance is used as a standard, the nature of the evaluation is temporal (i.e., me now vs. me in the past) rather than social (i.e., me vs. others). Since Albert’s (1977) cogent analysis of temporal comparisons, researchers have tended to consider temporal
and social comparisons as alternatives to one another. For example, Wilson and Ross (2000) examined how frequently temporal and social comparisons were used as a basis of self-evaluations. They showed that individuals’ desires to obtain enhancing (i.e., positive) versus accurate information about themselves served as the fundamental motives underlying the use of temporal versus social comparisons, respectively.

An important aspect of previous research that compared the effects of temporal and social comparisons is that it has largely been confined to situations in which people were evaluating themselves (Robinson-Whelen & Kiecolt-Glaser, 1997; Suls, Marco, & Tobin, 1991; Wilson & Ross, 2000). In contrast, there has been a lack of research on how temporal and social comparisons differentially influence people when those comparisons are used by others to evaluate them. This omission is unfortunate because people often receive evaluations from others in a variety of settings, such as schools and workplaces (Ilgen et al., 1979).

There are two noteworthy exceptions. First, Levine and Green (1984) examined the interactive effect of temporal and social comparison information (from others) on children’s attention to their peers. Their findings suggested that children reduced their attention to their peers when they were told that their own performance was decreasing, especially when they were also being outperformed by their peers. Second, Zell and Alicke (2009b) demonstrated that individuals’ self-perceptions of competencies were affected by how their performance changed over time (i.e., temporal comparisons) whereas external observers’ perceptions of competencies were largely shaped by how individuals fared against others (i.e., social comparisons).

Despite these findings, much needs to be learned about how people respond to performance evaluations made by others as a function of temporal and social comparisons. First, prior research (e.g., Levine & Green, 1984; Zell & Alicke, 2009) has mainly focused on the
valence of comparisons (i.e., whether people are doing better or worse than temporal versus social standards). In contrast, relatively little is known about how the utilization of temporal versus social comparisons in and of itself influences those who are evaluated. This is important to examine because individuals who are receiving the evaluations may make different inferences about how they are being treated depending on the type of comparisons (Lind & Tyler, 1988). For example, people may consider a temporal evaluation to be more individualized because it focuses only on them, whereas a social evaluation includes information about other people. As I hypothesize below, being treated in a more individualized fashion may induce people to consider that they are being dealt with in a more dignified and respectful way. Furthermore, believing that the evaluators have taken into account detailed information about their performance, employees who receive temporal evaluations may also consider the evaluation processes as more accurate and unbiased.

I investigate how temporal versus social comparisons shape employees’ judgments related to the evaluation process. More specifically, I examine whether temporal versus social comparisons lead to different perceptions of procedural and interpersonal fairness (Colquitt, 2001; Colquitt et al., 2005). I further explore the psychological mechanism that explains the differential effects of temporal versus social comparisons on fairness perceptions.

It is worth examining whether the type of comparisons influences fairness perceptions because performance evaluations are likely to have more positive effects when employees receive them well (e.g., openly rather than defensively; Ilgen et al., 1979). Performance evaluations are designed to help employees assess how they are doing their jobs and ultimately, to motivate them to perform better (Mayer & Davis, 1999). For these positive effects of performance evaluations to be realized, they need to be accepted by the individuals on the
receiving end. A cardinal principle in the organizational justice literature is that people are more likely to accept information and decisions that are accompanied by a fair process (Colquitt et al., 2001; Lind & Tyler, 1988). Therefore, the positive consequences of performance evaluations are more likely to be realized when the evaluations are seen as adhering more to principles of procedural and interpersonal fairness (Leung et al., 2001).

Temporal Comparisons, Social Comparisons, and Fairness

The distinction between temporal and social comparisons has proven meaningful in a variety of literatures such as achievement goals (Elliot & Thrash, 2001), health management (Suls et al., 1991), and aging (Robinson-Whelen & Kiecolt-Glaser, 1997). For example, researchers have suggested that individuals with learning (or mastery) goals typically make temporal comparisons to evaluate their achievement whereas those with performance goals are more apt to use social comparisons (C. Ames & Ames, 1984; Elliot & Thrash, 2001). However, as noted above, prior research has mainly examined why people use temporal versus social comparisons when they are evaluating themselves, rendering less clear how individuals react to temporal versus social comparisons coming from others. This is partially due to a lack of research on the effects of temporal comparisons. As Zell and Alicke (2009b) put it, “Compared to the hundreds of studies that fly under the social comparisons banner, research on temporal comparisons is sparse” (p. 224). The present research attempts to extend previous work by examining how temporal versus social comparisons made by others differentially affect people’s perceptions along a dimension known to have ubiquitous effects on work attitudes and behaviors: fairness.

The vast literature on organizational justice (e.g., Colquitt et al., 2005) has distinguished between the fairness related to (1) the outcomes that employees receive (distributive fairness;
Adams, 1965), (2) the decision-making process associated with the outcomes (procedural fairness; Leventhal et al., 1980), and (3) the behavior of the parties who plan and implement decisions (interactional fairness; Bies, 1987). Interactional fairness has further been subdivided into interpersonal and informational fairness, the former capturing the extent to which individuals are treated in a respectful manner and the latter reflecting how effectively relevant information has been communicated (Shapiro, Buttner, & Barry, 1994). The present research focuses on how temporal versus social comparisons affect perceptions of procedural and interpersonal fairness, because I focus on employees’ perceptions related to the evaluation process, not the evaluation outcome. Next, I provide the theoretical rationale for the prediction that employees’ judgments of procedural and interpersonal fairness will be higher when they receive a performance evaluation that is based on temporal versus social comparisons.

**The Effects of Temporal and Social Comparisons on Fairness Perceptions**

A temporal comparison uses employees’ own past performance as the standard to evaluate their current performance. Therefore, in temporal comparisons, employees receiving the evaluations are the only individuals being considered during the appraisal process. This conveys to employees that the evaluations have focused on them and that the evaluators have devoted attention to the particular evaluations they are receiving (Sluss & Thompson, 2012). Thus, employees may infer that the evaluator has incorporated a plenty of information specific to their performance.

In contrast, a social comparison evaluation discusses employees’ performance relative to other people’s performance, which may elicit a perception of being viewed as another face in the crowd. When employees receive a social evaluation, it is implied that other people are also receiving the same type of evaluations. Employees may thus think that the evaluators’ effort to
conduct the evaluations has been distributed across multiple parties (self and others), giving them the impression that less attention was devoted to their own appraisal. In this situation, employees are less likely to think that the evaluators have taken into account the details of their performance.

Thus, I predict performance evaluations that emphasize temporal rather than social comparisons are more likely to lead employees to believe that their evaluators have incorporated specific details of their performance. I refer to this reaction of employees as perceptions of individualized treatment. This construct is distinct from related concepts such as leader-member exchange (LMX; Graen, Novak, & Sommerkamp, 1982; Liden & Maslyn, 1998) and feedback specificity (Goodman, Wood, & Chen, 2011). Particularly, LMX focuses on employees’ perceptions describing their relationships with the leaders, and thus it encompasses broad domains of interpersonal dynamics (e.g., “How would you characterize your working relationship with your immediate supervisor?”; Liden & Graen, 1980). Therefore, LMX does not appear to capture what employees experience specifically in the context of performance evaluations. On the other hand, feedback specificity directly deals with performance evaluation settings. However, it does not reflect the social dynamics between the evaluators and employees. Feedback specificity instead pertains to objective rules governing performance evaluations (e.g., "I was given specific feedback about my performance”; Goodman et al., 2011). Therefore, LMX and feedback specificity may not be suitable constructs to capture interpersonal dynamics in performance evaluations. Moreover, neither of them directly measures the extent to which employees believe that their evaluators have incorporated specific details related to their performance.

I focus on the construct of individualized treatment because I believe that it mediates the
relationships between the type of comparisons and fairness perceptions. I predict that perceptions of individualized treatment, which are more likely to be elicited by temporal than social comparisons, will lead employees to perceive higher levels of procedural fairness in the domains of accuracy, the suppression of personal biases, and the ethicality of evaluation processes (Colquitt, 2001; Leventhal et al., 1980; van Prooijen & Zwenk, 2009). Each employee’s performance involves many details. For example, an employee who previously demonstrated stellar performance might temporarily lose focus and struggle. Conversely, an employee who started with a low level of productivity might improve after gaining experience. Importantly, employees tend to consider such details as important pieces of information related to their performance (Zell & Alicke, 2009b). Thus, employees who think that their evaluation incorporates such specific details may think that it is based on a good deal of information related to their performance and thus is more accurate and unbiased. They may also perceive the evaluation as more ethical in that the evaluators may have invested significant efforts to take into account much information regarding their performance. Therefore, temporal comparisons and ensuing perceptions of individualized treatment may lead people to believe that the evaluators have engaged in due diligence, which makes the appraisal right not only in terms of accuracy or unbiasedness but also from a moral or ethical point of view.

I do not predict that the distinction between temporal and social comparisons applies to all elements of procedural fairness, such as perceived voice (Colquitt, 2001). Regardless of whether temporal or social comparison evaluations are provided, they are conducted by the evaluators according to the procedures determined by the evaluators, without requiring input from the parties being evaluated. Therefore, I limit the scope of my prediction to the procedural fairness elements of accuracy, bias suppression, and ethicality, those I considered to be most conceptually
linked with the distinction between temporal and social comparisons.¹⁰

*Hypothesis 5. Relative to a social comparison performance evaluation, a temporal comparison performance evaluation is perceived as adhering more to the procedural fairness principles of accuracy, bias suppression, and ethicality.*

Performance evaluations that use temporal comparisons may also be perceived to adhere to the principles of interpersonal fairness such as respectful, dignifying, and polite treatment. Interpersonal exchange in organizations has symbolic value. Feedback provided by an evaluator can have strong meaning to employees because it conveys how they are viewed in the organization (De Cremer & Mulder, 2007). When temporal comparisons signal that the evaluation is focusing on them and incorporating specific details about their performance, it may symbolize that they are valued by their evaluators (Olkkonen & Lipponen, 2006), which may lead to the employees’ sense of receiving respectful, dignifying, and polite treatment (Bies, 1987; Bies & Shapiro, 1987).

In contrast, social comparisons are less likely to prompt favorable judgments of these interpersonal fairness principles. When the evaluation does not seem to draw on specific information about their performance, employees may be less likely to think that they are being recognized on an individual basis by their evaluators. Believing that the details related to their

¹⁰ Because I am investigating a subset of the elements that comprise procedural fairness, it is more appropriate to refer to the dependent variable (i.e., perceptions of accuracy, bias suppression, and ethicality) as perceptions pertaining to some elements of procedural fairness (cf. Scott et al., 2009). However, to be more concise in the presentation I use the term “perceptions of procedural fairness” in referring to judgments of accuracy, bias suppression, and ethicality. A similar approach applies to upcoming Hypothesis 6, which focuses on elements of interpersonal fairness
performance have not been considered as much, they may infer that they are being treated as one of the masses rather than as individuals in their own right (Hofmann, Morgeson, & Gerras, 2003; Lind & Tyler, 1988), in which case the symbolic message of respectful treatment, dignity, and politeness may not shine through. I thus predict that employees’ perceptions of respectful, dignifying, and polite treatment will be higher in response to a temporal comparison evaluation than a social comparison evaluation.

Similar to procedural fairness, I do not predict that the effects of temporal versus social comparisons are relevant to all elements of interpersonal fairness. One of the items comprising interpersonal fairness, “refraining from making improper comments,” seems conceptually distant from the distinction between temporal versus social comparisons (Bies & Moag, 1986; Colquitt, 2001); both types of comparisons can be made without resorting to rude or improper comments. The following hypothesis is thus limited to the interpersonal fairness elements of respectful, dignifying, and polite treatment.

\textit{Hypothesis 6. Relative to a social comparison performance evaluation, a temporal comparison performance evaluation is perceived as adhering more to the interpersonal fairness principles of respectful, dignifying, and polite treatment.}

The culmination of the above reasoning gives rise to another prediction:

\textit{Hypothesis 7. Employees’ perceptions of how much they are receiving individualized treatment mediate the relationships between temporal versus social comparison evaluation and perceptions pertaining to procedural fairness (i.e., accuracy, bias suppression, and ethicality) and interpersonal fairness (i.e., dignity, respect, and politeness).}
Overview of studies

The present research consists of 3 empirical studies. In Study 5, using the data collected from a sample of employees, I collected data on temporal and social comparisons that participants experienced in their workplaces and examined their effects on fairness perceptions.

To enhance the internal validity of the findings from Study 5, in Study 6 I experimentally manipulated temporal versus social comparisons and examined their effects on fairness perceptions. All participants in Study 6 read scenarios in which they imagined that they were on the receiving end of a performance evaluation in an organization. I also tested in Study 6 whether the effects of temporal versus social comparisons on fairness perceptions emerged in both positive and negative evaluations. To do so, I orthogonally manipulated the type of comparisons (temporal or social) and the valence of the evaluation (positive, negative, or mixed).

In Study 7 I attempted to replicate the causal effects of temporal and social comparisons on fairness perceptions. The difference between Study 7 and Study 6 was that participants in Study 7 actually received performance evaluations from their manager (whereas in Study 6 participants reported how they would have responded in a hypothetical situation). Furthermore, in Study 7 I performed the tests of mediation set forth in Hypothesis 7.

In sum, the present studies draw on a variety of methodologies to examine whether people perceive higher levels of procedural and interpersonal fairness in response to a temporal comparison evaluation than to a social comparison evaluation. To the extent that similar results emerge across studies, I gain increased confidence in the validity of the findings. By examining the mediating mechanism, I also seek to shed light on why people perceive higher levels of procedural and interpersonal fairness in response to temporal versus social comparisons.
Study 5

In Study 5, I collected data regarding performance evaluations that employees received from their workplaces. I manipulated temporal versus social comparisons by asking participants to recall their performance evaluations that emphasized temporal versus social comparisons, respectively.

Method

Sample. I recruited 178 American online participants via Amazon.com’s Mechanical Turk platform. Only those who were currently working in organizations that provided performance evaluations on a regular basis were eligible to participate. The mean age of the participants was 32.9 (SD = 10.5) and 39% of them were female. The participants identified their race with the following frequencies: 80% White, 7% Asian, 6% Black, 3% Latino(a), and 4% two or more races. Twenty-seven of them failed to pass the attention checks and thus were excluded from the analyses.

Procedures. After eligible participants agreed to take part in the study, they were randomly assigned to a condition in which they were asked to think of a performance appraisal they received within the previous year that emphasized either temporal or social comparisons. Participants who were assigned to the temporal comparison condition were asked to, “think of the evaluation that emphasized how you were evaluated relative to your previous evaluations,” whereas those in the social comparison condition were asked to, “think of the evaluation that emphasized how you were evaluated relative to other people’s evaluation.” To heighten the salience of the recalled evaluations, participants were encouraged to think about what comments their evaluators specifically made, how the evaluators approached the issue, and how they felt about the evaluations. They were then requested to write a few sentences to describe the
performance evaluations in a blank box located below the instructions. After the descriptions, they filled out a survey measuring perceptions of temporal and social comparisons (i.e., manipulation checks) and perceptions of procedural and interpersonal fairness.

**Measures.** All items in Studies 5 to 7 used a 7-point Likert-type scale from 1 (“not at all”) to 7 (“very much”). Manipulation checks of temporal and social comparisons were respectively, “How much did the evaluation compare your current performance with your previous performance?” and “How much did the evaluation compare your current performance with other people’s performance?”

To measure perceptions of procedural fairness, I drew on all seven items from the Colquitt (2001) measure. This enabled me to capture not only the three elements of procedural fairness that I hypothesized to be shaped by temporal versus social comparisons but also other elements that are less conceptually related to the distinction between temporal and social comparisons, such as perceived voice during the evaluation process. I predicted that whereas temporal comparison evaluations may lead to perceptions of greater accuracy (“The evaluation was based on accurate information”), bias suppression (“The evaluation was free of bias”), and ethicality (“the evaluation upheld ethical and moral standards”), it may not significantly influence other elements of procedural fairness such as perceived voice (e.g., “You had influence over the evaluation arrived at by the evaluation”).

Similarly, perceptions of interpersonal fairness was measured using all four items from Colquitt (2001), but I predicted that the difference between temporal versus social comparisons may emerge in the elements of respectful (“The evaluator treated you with respect”), dignifying (“The evaluator treated you with dignity”), and polite treatment (“The evaluator treated you in a polite manner”), not in the element of refraining from improper remarks (“The evaluator
refrained from improper remarks or comments”).

**Results**

**Manipulation check.** Table 11 reports descriptive statistics and their intercorrelations of the study variables. Participants in both the temporal and social comparison conditions completed the two manipulation-check items. Hence, I conducted a 2 (Type of comparisons) X 2 (Comparison measure) analysis of variance, with the former and latter treated as between- and within-subject variables, respectively. As expected, there was a significant interaction effect ($F = 13.79, p < .001$). Participants in the temporal comparison condition perceived higher levels of temporal comparisons ($M = 5.45, SD = 1.48$) than those in the social comparison condition ($M = 4.45, SD = 1.66; t = 3.91, p < .001$). In contrast, participants in the social comparison condition perceived higher levels of social comparisons ($M = 4.68, SD = 1.92$) than those in the temporal comparison condition ($M = 4.19, SD = 2.04$), but this difference did not reach the .05 confidence level ($t = -1.52, p = .132$).

**Procedural fairness.** As predicted, participants in the temporal comparison condition reported significantly higher levels of procedural fairness for the elements of accuracy, unbiasedness, and ethicality ($M = 6.18, SD = .92$) than those in the social comparison condition ($M = 5.63, SD = 1.42; d = .45, F = 7.71, p = .006$). Then I tested whether participants in the temporal versus social comparison conditions perceived different levels of procedural fairness for the remaining elements (which I did not hypothesized to be significantly affected by temporal versus social comparisons). With respect to these elements, there was no significant difference between perceptions of procedural fairness reported by participants who were in the temporal comparison condition ($M = 4.08, SD = .94$) and participants who were in the social comparison condition ($M = 3.83, SD = 1.18; d = .23, F = 2.04, p = .156$).
**Interpersonal fairness.** Participants in the temporal comparison condition reported higher levels of interpersonal fairness for the elements of accuracy, unbiasedness, and ethicality ($M = 6.53$, $SD = .80$) than those in the social comparison condition ($M = 6.28$, $SD = 1.00$; $d = .27$), but the difference was not significant at the .05 confidence level ($F = 2.97$, $p = .087$). I further tested whether participants in the temporal versus social comparison conditions perceived different levels of interpersonal fairness for the remaining element: refraining from improper comments. With respect to this particular element, there was no significant difference between perceptions of interpersonal fairness reported by participants who were in the temporal comparison condition ($M = 5.55$, $SD = 2.16$) and participants who were in the social comparison condition ($M = 5.76$, $SD = 1.18$; $d = -.09$, $F = .32$, $p = .572$). Moreover, the different was in the opposite direction from the one observed from the other elements of interpersonal fairness.\(^{11}\)

**Discussion**

The results of Study 5 indicated that employees perceived higher levels of accuracy, unbiasedness, and ethicality in response to temporal than social comparison evaluations. In contrast, for the other elements of procedural fairness (such as perceived voice), there was no significant difference between temporal versus social comparison conditions. Employees in the temporal comparison condition also reported higher perceptions of respectful, dignifying, and...

\(^{11}\) I checked whether participants’ gender and age had any effects on the results. I ran a 2 (Type of comparisons) X 2 (Gender) between-subject analyses of covariance with participants’ age as a covariate. I observed only one significant effect of participants’ gender and age: age had a positive main effect on perceptions of evaluators’ refraining from improper comments ($F = 5.84$, $p = .017$). More importantly, the results reported in this study did not change in terms of their statistical significance when gender and age were included in the analyses.
polite treatment than employees in the social comparison condition. However, the difference was smaller than the one observed from perceptions of procedural fairness. This might be explained by a general ceiling effect; the mean perceptions of interpersonal fairness across conditions were quite high ($M = 6.41$ on a 7-point scale). For the other element of interpersonal fairness (i.e., refraining from improper remarks), the difference between temporal versus social comparisons was much smaller and was in the opposite direction from the one observed for the elements of respectful, dignifying, and polite treatment.

**Study 6**

Study 6 was designed to test the research questions examined in Study 5 with greater internal validity. Whereas participants in Study 5 were randomly assigned to recall instances in which they had received evaluations emphasizing temporal or social comparisons, it is possible that the recalled instances differed in other ways as well. For example, the recalled instances may have differed in terms of how positive or negative the evaluations were. Given that the valence of performance evaluations has been shown to affect employees’ attitudes and behaviors (e.g., Podsakoff & Farh, 1989), in Study 6 I manipulated evaluation valence and examined whether the effects of temporal versus social comparisons on fairness perceptions emerged in both positive and negative evaluations.

All participants in Study 6 indicated how much fairness they perceived in response to a hypothetical performance evaluation. Both the type of comparisons (temporal versus social) and the valence of the evaluation (positive, negative, or mixed) were experimentally manipulated in a 2 X 3 factorial design. Because Study 5 results revealed that the distinction between temporal and social comparisons was relevant only to the elements of accuracy, unbiased, and ethical processes for procedural fairness and respectful, dignifying, and polite treatment for
interpersonal fairness, only these elements of fairness were measured in Study 6.

**Method**

**Sample.** I recruited 401 American online participants via Amazon.com’s Mechanical Turk (Buhrmester et al., 2011). The mean age of the participants was 36.0 ($SD = 13.0$) and 47% of them were female. The participants identified their race with the following frequencies: 76% White, 9% Black, 5% Asian, 3% Latino(a), 1% Native American, and 5% two or more races. Thirty-four individuals failed to pass the attention checks and thus were excluded from the analyses. The inclusion of these participants did not change the statistical significance of my results.

**Procedures.** After participants agreed to take part in the study, they were asked to imagine that they were employees receiving a performance evaluation in a mid-sized firm located in the United States. It was further explained that the performance evaluation in their company took place every six months, and that they were receiving the evaluation from their boss based on how they had performed during the past six months. Then participants were randomly led to believe that the evaluator was providing either a temporal or social comparison evaluation. Participants in the temporal comparison condition read the following: “The evaluation focuses on how you performed during the last six months. However, your boss is specifically emphasizing *how your current performance fares against your previous performance (i.e., six months ago).* In other words, your boss is *comparing your current performance with your previous performance.*” In contrast, participants in the social comparison condition read the following: “The evaluation focuses on how you performed during the last six months. However, your boss is specifically emphasizing *how your current performance fares against other people’s performance (i.e., your colleagues).* In other words, your boss is *comparing your current performance with your*
Next, to vary the valence of the evaluation, participants were randomly given one of three evaluative statements (positive, negative, or mixed). The statement in the positive valence condition was, “Generally, your performance is showing that you are doing a very good job. There are many aspects in which you are demonstrating strengths. Of course, you might need some minor adjustments for the things you are struggling with. But, on balance, good going!” The statement in the negative valence condition was, “Generally, your performance is showing that you are not doing a very good job. There are several aspects in which you are demonstrating weaknesses. Of course, there are certain areas where you are doing okay. But, on balance, you need to improve.” The statement in the mixed valence condition was, “There are many things that you are doing well, and there are also many things that need some corrections. Some aspects of your performance are showing that you have done a commendable job, and some are requiring additional effort. Whereas you are performing decently, I’d like to encourage you to aim higher.” Importantly, the evaluative statements were identical across the temporal versus social comparison conditions.

After reading the performance evaluation scenarios, participants completed manipulation checks pertaining to the independent variables as well as the items measuring perceptions of procedural and interpersonal fairness.

**Measures.** Manipulation checks of temporal and social comparisons, perceptions of procedural fairness (accuracy, unbiasedness, ethicality), and perceptions of interpersonal fairness (respectful, unbiased, and polite treatment) were measured using the same items as in Study 5. Manipulation checks of evaluation valence were, “How much did the evaluation consider the positive side of your performance?” (positive valence) and “How much did the evaluation
consider the negative side of your performance?” (negative valence).

Results

Manipulation check: Type of comparison. Table 12 reports descriptive statistics and intercorrelations of the study variables. I conducted a 2 (Type of comparisons) X 2 (Comparison measure) analysis of variance, with the former and latter treated as between- and within-subject variables, respectively. The results showed a significant interaction effect ($F = 456.98$, $p < .001$). Participants in the temporal comparison condition perceived higher levels of temporal comparisons ($M = 5.48$, $SD = 1.59$) than those in the social comparison condition ($M = 3.04$, $SD = 1.85$; $t = 13.59$, $p < .001$). Conversely, participants in the social comparison condition perceived higher levels of social comparisons ($M = 5.14$, $SD = 1.78$) than those in the temporal comparison condition ($M = 2.03$, $SD = 1.41$; $t = -18.63$, $p < .001$).

Manipulation check: Evaluation valence. The manipulation check items of positive and negative valence were submitted to a 3 (Evaluation valence) X 2 (Valence measure) analysis of variance, with the former and latter treated as between- and within-subject variables, respectively. The results showed a significant interaction effect ($F = 330.30$, $p < .001$). Participants in the positive valence condition perceived higher levels of positive valence ($M = 5.86$, $SD = .92$) than those in the mixed ($M = 4.52$, $SD = 1.18$) and negative valence conditions ($M = 2.79$, $SD = 1.36$; $F = 214.60$, $p < .001$). Conversely, participants in the negative valence condition perceived higher levels of negative valence ($M = 6.05$, $SD = 1.07$) than those in the mixed ($M = 4.54$, $SD = 1.17$) and positive valence conditions ($M = 3.00$, $SD = 1.16$; $F = 229.00$, $p < .001$).

It is worth noting that the manipulation of temporal versus social comparisons did not generate significant differences in perceptions of positive valence ($t = 1.38$, $p = .168$) and
negative valence ($t = -.93, p = .353$).

**Procedural and interpersonal fairness.** Perceptions of procedural and interpersonal fairness were submitted to 2 (Type of comparisons) X 3 (Evaluation valence) between-subject analyses of variance. The results demonstrated that perceptions of procedural fairness were significantly affected by both the type of comparisons ($F = 12.30, p < .001$) and evaluation valence ($F = 40.21, p < .001$). There was no significant interaction between the two independent variables ($F = .40, p = .672$). Similarly, perceptions of interpersonal fairness were significantly affected by temporal versus social comparisons ($F = 9.17, p = .003$) and evaluation valence ($F = 100.33, p < .001$), the interaction being nonsignificant ($F = .75, p = .473$).

Table 13 reports the mean levels and standard deviations of fairness perceptions and the $p$-values for the differences between the temporal and social comparison conditions (see Figure 6 for a graphical demonstration). As predicted, participants who received temporal comparison evaluations believed that they were treated with greater procedural fairness ($M = 5.16, SD = 1.29$) than their counterparts who received social comparison evaluations ($M = 4.72, SD = 1.39; d = .33, t = 3.19, p = .002, 95\% CI = [.171, .722]$). Similarly, participants who received temporal evaluations perceived higher levels of interpersonal fairness ($M = 5.38, SD = 1.42$) than those who received social evaluations ($M = 5.01, SD = 1.46; d = .25, t = 2.44, p = .015, 95\% CI = [.071, .662]$). These patterns emerged in all three levels of evaluation valence. From Table 13 it can also be seen that fairness perceptions were highest in the positive valence condition, lowest in the negative valence condition, and in between these two in the mixed valence condition.

Finally, I tested whether the differences between fairness perceptions in the temporal and social comparison conditions remained significant when the measures of positive and negative valence were included in the analyses as control variables. The type of comparisons had
significant effects on both procedural fairness \((t = 2.91, p = .004, 95\% CI = [.109, .562])\) and interpersonal fairness \((t = 2.10, p = .037, 95\% CI = [.014, .421])\).\(^{12}\)

**Discussion**

Study 6 built on the findings in Study 5 by offering greater internal validity. The experimental manipulations of temporal versus social comparisons and the valence of the evaluation had significant and orthogonal effects on perceptions of procedural and interpersonal fairness. Temporal comparisons led to higher perceptions of fairness than social comparisons, and these patterns emerged in all three levels of evaluation valence. Finally, the type of comparisons had significant effects on fairness perceptions even when I controlled for the measures of positive and negative valence, suggesting that the effects of temporal versus social comparisons existed over and above how positive and negative the evaluations were perceived to be.

**Study 7**

Although the Study 6 findings were internally valid, participants did not actually experience the events of Study 6 but rather reported how they would have responded if they had received the evaluations. In Study 7, I moved toward examining how people reacted to

---

\(^{12}\) I checked whether participants’ gender and age had any effects on the results. I ran a 2 (Type of comparisons) X 3 (Evaluation valence) X 2 (Participants’ gender) between-subject analyses of covariance with participants’ age as a covariate. Gender and age did not demonstrate any significant main or interaction effects on procedural or interpersonal fairness. Furthermore, the main findings reported in this study were not affected in terms of their statistical significance when gender and age were included in the analyses.
performance evaluations that they actually experienced, as a function of the type of comparisons and the valence of the evaluation. The dependent variables consisted of the same measures of procedural and interpersonal fairness examined in Study 6.

Another goal of Study 7 was to measure the hypothesized mediator, perceptions of individualized treatment, and to examine its effect. Moreover, in Study 7 I measured another dimension of fairness that captured the characteristics of evaluation processes: informational fairness (Bies & Moag, 1986; Colquitt, 2001; Shapiro et al., 1994).

On an exploratory basis, I also varied the way in which participants experienced temporal versus social comparisons. Half of the participants were given performance evaluations that emphasized one type of comparisons with nothing said about the other (e.g., high temporal comparisons; Version 1). The other half were given evaluations that emphasized one type of comparisons and also minimized the other type (e.g., high temporal comparisons AND low social comparisons; Version 2). Thus, Study 7 consisted of a 2 (temporal vs. social) X 2 (positive vs. negative) X 2 (Version 1 vs. Version 2) between-subject design. As I describe below, the results were consistent across the two different versions.

Method

Sample. I recruited 404 American online participants via Amazon.com’s Mechanical Turk platform. The mean age of the participants was 37.0 (SD = 11.4) and 45% of them were female. The participants identified their race with the following frequencies: 76% White, 7% Black, 6% Latino(a), 6% Asian, 0.5% Native American, and 5% two or more races. Four individuals raised suspicion at the end of the study and thus were excluded from the analyses.13

13 The exclusion of these participants did not change the statistical significance of the results
Procedures. After participants agreed to take part in the study, they were notified that two other people (named Jessie and Casey) would be working with them simultaneously. In reality, there were no such people. Participants were further informed that one of the three people would assume the role of a manager and the other two would be employees. Their roles were then decided in an ostensibly random manner. Participants and one other person (Jessie) were assigned to the role of employees whereas the other person (Casey) was given the role of a manager. Participants were told that the two employees (themselves and Jessie) would work on a task and the manager (Casey) would provide an evaluation of their performance. After the role assignment, participants worked on a task. The task was referred to as the “Social Insight Test.” In this task, participants were given a brief description of someone’s experiences at work and were asked to predict whether the individual would have remained in the company 12 months after the described incident. To heighten participants’ engagement with the task, they were informed that Social Insight is an important skill, and many studies had demonstrated that people high in Social Insight achieved better career outcomes and enjoyed more rewarding interpersonal relationships. After participants worked on a sample question, they started Round One of the Social Insight Test.

Round One consisted of five descriptions from five different people. To maintain ambiguity about how well they were performing and thereby preserve the credibility of the evaluations, participants were not given any information regarding whether they solved the questions correctly or not. After Round One, participants were told that Round Two would have the same format (five descriptions from five different people who did not overlap with those in reported in this study.
Round One) and the manager (Casey) would evaluate their performance after Round Two. After participants finished Round Two (again, without being informed whether they solved the questions correctly or not), they were told that the manager sent them a brief message before the actual evaluation, and this message served as the first part of the manipulation.

As noted, the temporal and social comparison manipulation consisted of two different versions to which participants were randomly assigned. The evaluative statements unitalicized below are from Version 1 (which only emphasized one type of comparison). Version 2 (which emphasized one type of comparison and also minimized the other) had the same contents as Version 1 with some additional phrases. The phrases that were added to Version 2 are presented in italics.

The message in the temporal comparison conditions was, “You have completed both rounds. I will evaluate your Social Insight ability and give you feedback. What is really important is how you did in Round Two relative to how you did in Round One, because it signals whether you are on the right track regardless of how you did in the two rounds relative to Jessie. So, I will focus on how you performed in Round Two compared to how you performed in Round One.” The message in the social comparison conditions was, “You have completed both rounds. I will evaluate your Social Insight ability and give you feedback. What is really important is how you did in the two rounds relative to how Jessie did in the two rounds, because it signals whether you are on the right track regardless of how you did in Round Two relative to Round One. So, I will focus on how you did in the two rounds relative to how Jessie did in the two rounds.”

After this brief message from the manager (i.e., the first part of the manipulation), participants were asked to wait for about two minutes so that the manager could evaluate their performance. After 90 seconds, participants were notified that the manager finished the
evaluation. They proceeded to check the contents of their evaluations, which served as the second part of the manipulation. Again, the evaluative statements unitalicized below are from Version 1. The contents of Version 2 were identical to those of Version 1 plus additional phrases, which are presented in italics.

The evaluation in the positive (negative) temporal conditions was, “Like I said, I’ve focused on how you performed in Round Two compared to Round One of the Social Insight Test, because that’s what really matters regardless of how you did in the two rounds compared to Jessie. Your performance in Round Two was higher (lower) than your performance in Round One, suggesting you are (are not) doing better than before. So, I’d say you did a good job (I’d say you could have done better).”

The evaluation in the positive (negative) social conditions was, “Like I said, I’ve focused on how you performed compared to Jessie in the two rounds of the Social Insight Test, because that’s what really matters regardless of how you did in Round Two compared to Round One. Your performance was higher (lower) than Jessie’s performance, suggesting you are (are not) doing better than Jessie. So, I’d say you did a good job (I’d say you could have done better).”

After the evaluations, participants answered manipulation checks pertaining to the independent variables (the type of comparisons and evaluation valence) along with the items measuring fairness perceptions, and the hypothesized mediator.

**Measures.** Perceptions of temporal comparisons, social comparisons, positive valence, negative valence, procedural fairness, and interpersonal fairness were measured using the same items as in Study 6. Perceptions of informational fairness were measured with the five items from Colquitt (2001), such as, “Your manager’s explanations regarding the feedback were reasonable.”
As noted, perceptions of individualized treatment in performance evaluations have not received direct attention from past research. Therefore, I created items to measure this construct. I adhered to three principles in developing the items. First, they tapped specifically into the context of performance evaluations. Second, they reflected participants’ perceptions of what the evaluator did instead of the objective rules governing the evaluation process. Third, they directly captured my definition of the construct (i.e., the extent to which employees believe that their evaluators have incorporated specific details of their performance). The three items that I developed based on these principles were, “Your manager incorporated details about your performance,” “Your manager took into account specific information regarding your performance,” and “Your manager provided individualized feedback to you.”

Results

The effects of temporal versus social comparisons described in this section were consistent across Versions 1 and 2. Therefore, I collapsed the analyses across the Version variable.

Manipulation check: Type of comparisons. Table 14 reports descriptive statistics and intercorrelations of the study variables. I conducted a 2 (Type of comparisons) X 2 (Comparison measure) analysis of variance, the former and latter being treated as between- and within-subject variables, respectively. I found a significant interaction effect ($F = 1045.05, p < .001$).

Participants in the temporal comparison condition perceived higher levels of temporal comparisons ($M = 6.06, SD = 1.39$) than those in the social comparison condition ($M = 2.28, SD = 1.75; t = 23.86, p < .001$). Conversely, participants in the social comparison condition perceived higher levels of social comparisons ($M = 6.30, SD = 1.30$) than those in the temporal comparison condition ($M = 1.94, SD = 1.55; t = -30.46, p < .001$).
**Manipulation check: Evaluation valence.** The manipulation-check items of positive and negative valence were submitted to a 2 (Evaluation valence) X 2 (Valence measure) analysis of variance, the former and latter being treated as between- and within-subject variables, respectively. The results showed a significant interaction effect ($F = 1244.45, p < .001$). Participants in the positive valence condition perceived higher levels of positive valence ($M = 6.00, SD = 1.05$) than those in the negative valence condition ($M = 2.30, SD = 1.36; t = 30.66, p < .001$). Conversely, participants in the negative valence condition perceived higher levels of negative valence ($M = 5.91, SD = 1.14$) than those in the positive valence condition ($M = 2.41, SD = 1.41; t = -27.27, p < .001$). As in Study 6, the manipulation of temporal and social comparisons did not generate significant differences in perceptions of positive ($t = 1.74, p = .083$) and negative valence ($t = .13, p = .900$).

**Procedural and interpersonal fairness.** Perceptions of procedural and interpersonal fairness were submitted to 2 (Type of comparisons) X 2 (Evaluation valence) between-subject analyses of variance. Perceptions of procedural fairness were significantly affected by both the type of comparisons ($F = 15.47, p < .001$) and evaluation valence ($F = 169.22, p < .001$). There was no significant interaction ($F = 1.80, p = .181$). Similarly, perceptions of interpersonal fairness were significantly influenced by the type of comparisons ($F = 43.50, p < .001$) and evaluation valence ($F = 98.79, p < .001$), the interaction being nonsignificant ($F = 1.11, p = .293$).

Table 15 reports the mean levels and standard deviations of fairness perceptions and the $p$-values of the differences between the temporal and social comparison conditions (see Figure 7 for a graphical demonstration). I found that participants who received temporal evaluations perceived higher levels of procedural fairness ($M = 5.56, SD = 1.50$) than those who received
social evaluations ($M = 5.04, SD = 1.64; d = .33, t = 3.30, p = .001, 95% CI = [.209, .827])

Those who received temporal evaluations also perceived higher levels of interpersonal fairness
($M = 5.47, SD = 1.32$) than their counterparts who received social evaluations ($M = 4.64, SD = 1.47; d = .59, t = 5.91, p < .001, 95% CI = [.551, 1.101]$). These patterns emerged in both the positive and negative evaluation conditions, while positive evaluations leading to higher perceptions of fairness than negative evaluations. Finally, I tested the effects of temporal versus social comparisons on procedural and interpersonal fairness controlling for the measures of positive and negative valence. The type of comparisons had significant effects on both procedural fairness ($t = 2.91, p = .004, 95% CI = [.119, .617]$) and interpersonal fairness ($t = 5.69, p < .001, 95% CI = [.444, .913]$).

**Informational fairness.** Perceptions of informational fairness demonstrated similar patterns: There was a significant main effect of the type of comparisons ($F = 12.20, p < .001$) and evaluation valence ($F = 82.18, p < .001$), but no significant interaction ($F = .07, p = .792$). As can be found in Table 15, participants who received temporal evaluations perceived higher informational fairness ($M = 4.80, SD = 1.39$) than those who received social evaluations ($M = 4.37, SD = 1.34; d = .32, t = 3.19, p = .002, 95% CI = [.167, .703]$). Again, this effect was observed in both the positive and negative evaluation conditions, while positive evaluations leading to higher perceptions of fairness than negative evaluations. Finally, the effect of temporal versus social comparisons was significant controlling for positive and negative valence measures ($t = 2.50, p = .013, 95% CI = [.064, .528]$).

**Individualized treatment.** The hypothesized mediator, perceptions of individualized treatment was also analyzed in a 2 (Type of comparison) X 2 (Evaluation valence) between-subject analysis of variance. The results demonstrated similar patterns as those found on the
dependent variables. Perceptions of individualized treatment were significantly shaped by both
the type of comparisons \( (F = 117.83, p < .001) \) and evaluation valence \( (F = 39.03, p < .001) \),
with no significant interaction between them \( (F = .03, p = .865) \). In Table 15, it can be found that
participants who received temporal comparison evaluations perceived higher levels of
individualized treatment \( (M = 4.93, SD = 1.40) \) than those who received social comparison
evaluations \( (M = 3.39, SD = 1.58; d = 1.04, t = 10.38, p < .001, 95\% CI = [1.255, 1.841]) \). This
effect was observed in both the positive and negative evaluation conditions, while positive
evaluations resulting in higher perceptions of the hypothesized mediator than negative
evaluations. The effect of temporal versus social comparisons was significant controlling for
positive and negative valence measures \( (t = 10.29, p < .001, 95\% CI = [1.135, 1.671]) \).14

**Tests of mediation.** I tested the mediating effects of individualized treatment in the
relationships between the type of comparisons and fairness. I used a bootstrap analysis with
5,000 resampling to estimate the indirect effects. The results suggested that perceptions of
individualized treatment demonstrated significant indirect effects for all three dimensions of
fairness (procedural, interpersonal, and informational; see Table 16 for results).

---

14 I examined whether participants’ gender and age had any significant effects on the results. I ran
a 2 (Type of comparisons) X 2 (Evaluation valence) X 2 (Participants’ gender) between-subject analyses
of covariance with participants’ age as a covariate. I observed only one significant effect of participants’
gender or age, which was a main effect of gender on interpersonal fairness: Female participants overall
perceived lower interpersonal fairness \( (M = 4.88, SD = 1.54) \) than male participants \( (M = 5.19, SD = 1.37; \\
t = -2.14, p = .033) \). Importantly, the statistical significance of the findings reported in this study did not
change when I included participants’ gender and age in the analyses.
Discussion

In Study 7, I aimed to examine the effects of temporal versus social comparisons that participants actually experienced in an experimental setting. The findings once again indicated that temporal comparisons led to higher perceptions of procedural and interpersonal fairness than social comparisons. Informational fairness was also rated higher in response to temporal than social comparisons. These patterns emerged in both positive and negative evaluations. Moreover, the effects of temporal versus social comparisons on fairness perceptions remained significant even when I controlled for how positive and negative the evaluations appeared to be. Study 7 also provided evidence related to the mechanism that may underlie these effects. Participants who received temporal comparison evaluations perceived higher levels of individualized treatment than those who received social comparison evaluations, and the perceptions of individualized treatment in turn led to more positive perceptions of procedural, interpersonal, and informational fairness.

General Discussion

Taken together, the present studies provide evidence that temporal comparison evaluations led to higher perceptions of procedural and interpersonal fairness relative to social comparison evaluations. Drawing on actual performance evaluations that employees received from their workplaces, Study 5 showed that temporal comparison evaluations were perceived to be procedurally and interpersonally fairer than social comparison evaluations. Given the potential confound that may have existed in the findings in Study 5, I moved toward establishing a causal relationship between the type of comparisons and fairness perceptions in Study 6 by randomly assigning participants to hypothetical scenarios. The results of Study 6 showed that temporal comparison evaluations induced perceptions of higher procedural and interpersonal fairness than
social comparison evaluations. Study 6 also varied the valence of the evaluation and showed that the effects of temporal versus social comparisons on fairness perceptions emerged in both positive and negative evaluations. Moreover, these effects of temporal versus social comparisons were observed even when I controlled for how positive and negative the evaluations appeared to be. In Study 7, I had participants receive temporal versus social evaluations in an experimental context. Once again, temporal evaluations led to stronger fairness perceptions than social evaluations. Study 7 results also suggested that the differential effects of temporal and social comparisons on fairness perceptions were explained by the extent to which employees perceived individualized treatment from the evaluator.

**Theoretical Implications**

**Temporal versus social comparisons.** Whereas the distinction between temporal and social comparisons is not new (e.g., Albert, 1977), most prior research has focused on when and why people instantiate one form of comparisons instead of the other to evaluate themselves (e.g., Butler, 1998; Wilson & Ross, 2001). The present research extended the investigations on temporal versus social comparisons to fairness perceptions. By doing so, it examined the effects of temporal and social comparisons when such comparisons were used by others to evaluate the focal individuals (Levine & Green, 1984; Zell & Alicke, 2009b).

The evidence from the past and present research indicates that the effects of temporal and social comparisons may depend on the source of the comparisons. Wilson and Ross (2000) showed that people used social comparisons to evaluate themselves when they were motivated to obtain an accurate assessment of themselves. In fact, the accuracy motive was at the heart of social comparison theory that was originally developed by Festinger (1954). This perspective proposed that individuals have a strong motivation to accurately assess where they stand, and
they compare themselves with others to make such an assessment. However, I found that employees who were evaluated by others using social comparisons (relative to temporal comparisons) perceived the evaluations as less accurate (an element of procedural fairness). In other words, individuals may initiate social comparisons to gain accurate information about themselves, but they might perceive lower levels of accuracy when social instead of temporal comparisons are used by other people. Perhaps people prefer to fulfill their need for accuracy on their own (by initiating social comparisons), whereas they may want their need for individualized treatment and respect satisfied by other people (by having evaluators who use temporal comparisons).

**Organizational justice.** The present research contributes to the organizational justice literature by examining what brings about fairness (Brockner et al., 2015). Previous research that investigated the antecedents of fairness focused on when and why decision makers are more likely to enact various principles of fairness (e.g., De Cremer, 2004; Greenberg, 1987; Scott, Colquitt, & Paddock, 2009). The dependent variable in the present studies was perceived fairness rather than the enactment of fairness. I am hardly the first to examine fairness-related perceptions. In fact, much of the theorizing in the earlier justice research delineated the attributes that constitute people’s perceptions of being treated fairly. For example, Leventhal et al. (1980) proposed the principles of procedural fairness (e.g., accuracy and bias suppression), whereas Bies (1987) discussed several features of interactional fairness (e.g., being treated in respectful and dignifying manner). The present research builds on these earlier contributions by identifying a factor that influences perceptions of procedural and interpersonal fairness, namely whether the evaluation standards emphasize the self (i.e., temporal comparisons) or other people (i.e., social comparisons).
While the present research focuses on procedural and interpersonal fairness as the main dependent variables, it also provides evidence that the greater perceived fairness in response to temporal than social comparisons generalizes to informational fairness. I make this statement more tentatively, since I had only one study (Study 7) that incorporated perceptions of informational fairness. Given that informational fairness is the dimension of organizational justice that has received the least attention in previous research (Colquitt et al., 2005), my findings appear to provide useful insight into how perceptions of informational fairness are formed.

**Individualized treatment.** Study 7 results demonstrated that perceptions of individualized treatment accounted for the relationships between the two types of comparisons and fairness perceptions. Individualized treatment refers to the extent to which evaluators incorporate specific details of employees’ performance. Employees may perceive individualized treatment as emanating from the evaluators’ personal characteristics or from their relationships with the evaluators (or both). In the former case, evaluators can be seen as those who generally take into account specific details to evaluate employees’ performance (“the evaluator does this in general”). In the latter case, the evaluations may be considered to result from personal ties between the employees and the evaluators (“the evaluator does this particularly for me”). These two types of attributions may influence employees’ perceptions of the evaluators. For example, judgments of trustworthiness consist of three dimensions: ability, integrity, and benevolence (Colquitt, Scott, & LePine, 2007). If employees attribute individualized treatment to the evaluators’ personal characteristics, they may think that the evaluators have the capabilities (ability) or character (integrity) to deliver such evaluations. In contrast, a relational attribution of individualized treatment may lead employees to believe that the evaluators have favorable
intention particularly toward them, leading to perceptions of high benevolence.

Practical Implications

The present findings have two important practical implications. First, the results of all three studies suggest that one way managers can enhance their followers’ perceptions of fairness in performance evaluations is to ensure that at least some aspects of the evaluations consist of temporal comparisons (C. Ames & Ames, 1984). Second, my mediation analyses suggested that temporal evaluations were seen as fairer because they signaled to employees that the evaluations were done in an individualized manner (Bies, 1987; Bies & Shapiro, 1987). Such findings imply that when organizations treat their employees in an individualized way in context other than performance evaluations, it may help the employees develop more positive job attitudes. For example, Cable, Gino, and Staats (2013) found that new employees who were socialized in a way that enabled them to use their signature strengths were more motivated and engaged six months later. One possible explanation of these findings is that employees responded positively when specific aspects of their individual characteristics were recognized and expressed on the job.

Although the present findings suggest that temporal comparisons can be more beneficial than social comparisons as reflected in fairness perceptions, there might be situations in which managers prefer to use social comparisons to evaluate employees. For example, managers may rely on social comparisons to justify the outcomes (e.g., salary increases, promotions) that employees receive (Adams, 1965). Moreover, social comparisons can be used to increase employees’ effort by triggering their competitiveness (Ryan, Koestner, & Deci, 1991). Thus, managers need to assess the tradeoff between the benefits and costs associated with social comparisons and try to ensure that the former outweigh the latter before they initiate a process of
comparing employees relative to one another.

**Limitations and Future Research**

The present research has limitations, which set the stage for future research. For example, whereas I consistently found that temporal comparison evaluations led to higher perceptions of fairness, it is theoretically as well as practically important to examine moderating influences on the present findings. That is, under what conditions are the present findings more versus less likely to emerge?

It can be particularly valuable to examine when social comparisons can heighten perceived fairness (Dunn et al., 2012; Novemsky & Schweitzer, 2004). Perhaps social comparison evaluations undertaken in an individualized manner may lead to higher perceptions of fairness. For example, if evaluators clearly have gathered lots of information about the employees they appraise and then use such information as a basis of social comparison evaluations, the employees may find the evaluations to be fair. More generally, given the well-established consequences of employees’ fairness perceptions, it is worth considering what it takes for social comparison evaluations to lead to higher perceptions of fairness.

Study 7 raises additional questions for further research. First, although I attempted to simulate organizational hierarchy by labeling the two roles in distinct ways (“manager” versus “employee”) and having participants evaluated by the manager, the sense of hierarchy may not have been very salient. It is worth examining whether the present findings generalize to a context in which people experience hierarchy more strongly.

Gender effects might also be relevant to the findings of Study 7. Based on the stereotypical beliefs that female managers should be warm and caring, participants who believed
their evaluator to be female might have reacted more negatively to social comparison evaluations. Although participants’ own gender did not moderate the effects found in Study 7, the effects of evaluators’ gender and the potential interaction between participants’ and evaluators’ gender remain unknown.

Finally, it may be possible to improve the measure of the mediator in Study 7. Two of the items referred to the evaluator’s behavior (“Your manager incorporated details about your performance” and “Your manager took into account specific information regarding your performance”) whereas the third one reflected more of an inference regarding how employees were treated by the evaluator (“Your manager provided individualized feedback to you”). Therefore, the first two items may have captured the behavioral antecedents of the perceptions measured by the third item. It is worth noting that the three items correlated highly with one another (the alpha coefficient was .83). Nevertheless, given the important mediating role played by the construct in Study 7, future research may investigate whether the measure can be improved.

**Conclusion**

Performance evaluations not only provide information on how well employees are doing their jobs, but also signal how the employees are recognized and treated in their workplaces. When individuals in organizations think that their performance is discussed relative to what they did in the past, they tend to think that the evaluators have considered detailed information about their performance to individualize the evaluations. In contrast, those who are compared with other people in their evaluations tend to believe that they are being treated as one of the masses whose specific details have not been fully incorporated. These differences between temporal and social comparisons shape perceptions of (a) how much the evaluations adhered to the principles
of accurate, unbiased, and ethical procedures and (b) how much the evaluators are providing respectful, dignifying, and polite treatment. This set of evidence has implications for what managers can do to increase the extent to which performance evaluations are accepted by the individuals on the receiving end: use temporal comparisons.
References


organizational justice (pp. 3-56). Mahwah, NJ: Erlbaum.


Klein, W. M. (1997). Objective standards are not enough: Affective, self-evaluative, and


behavior is not intrinsically motivated. *Motivation and Emotion, 15,* 185-205. doi:10.1007/bf00995170


Table 11

*Descriptive Statistics and Intercorrelations from Study 5*

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>α</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Manipulation check: Temporal comparisons</td>
<td>4.93</td>
<td>1.65</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Manipulation check: Social comparisons</td>
<td>4.44</td>
<td>1.99</td>
<td>--</td>
<td>.01</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Procedural fairness: Accuracy, unbiasedness, ethicality</td>
<td>5.90</td>
<td>1.23</td>
<td>.85</td>
<td>.24</td>
<td>.00</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Procedural fairness: Perceived voice and consistency</td>
<td>3.95</td>
<td>1.07</td>
<td>.75</td>
<td>.22</td>
<td>-.01</td>
<td>.47</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>5. Interpersonal fairness: Respectful, dignifying, polite</td>
<td>6.40</td>
<td>.92</td>
<td>.95</td>
<td>.09</td>
<td>-.02</td>
<td>.72</td>
<td>.39</td>
<td>--</td>
</tr>
<tr>
<td>6. Interpersonal fairness: Refraining from improper comments</td>
<td>5.64</td>
<td>1.98</td>
<td>--</td>
<td>-.05</td>
<td>-.12</td>
<td>.18</td>
<td>.10</td>
<td>.20</td>
</tr>
</tbody>
</table>

*Note: n = 151. r's greater than .16 are significant at .05 level*
<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>α</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Manipulation check: Temporal comparisons</td>
<td>4.33</td>
<td>2.10</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Manipulation check: Social comparisons</td>
<td>3.49</td>
<td>2.23</td>
<td>--</td>
<td>-.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Manipulation check: Positive valence</td>
<td>4.29</td>
<td>1.74</td>
<td>--</td>
<td>.24</td>
<td>-.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Manipulation check: Negative valence</td>
<td>4.62</td>
<td>1.69</td>
<td>--</td>
<td>-.10</td>
<td>.07</td>
<td>-.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Procedural fairness</td>
<td>4.95</td>
<td>1.36</td>
<td>.88</td>
<td>.27</td>
<td>-.13</td>
<td>.58</td>
<td>-.40</td>
<td></td>
</tr>
<tr>
<td>6. Interpersonal fairness</td>
<td>5.21</td>
<td>1.45</td>
<td>.97</td>
<td>.26</td>
<td>-.10</td>
<td>.72</td>
<td>-.57</td>
<td>.75</td>
</tr>
</tbody>
</table>

*Note: n = 367. rs greater than .10 are significant at .05 level*
Table 13

*Perceptions of Fairness from Study 6*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Procedural fairness</th>
<th>Interpersonal fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Temporal</td>
<td>Social</td>
</tr>
<tr>
<td>Positive valence</td>
<td>5.80 (.98)</td>
<td>5.49 (1.13)</td>
</tr>
<tr>
<td>Mixed valence</td>
<td>5.33 (1.06)</td>
<td>4.78 (1.36)</td>
</tr>
<tr>
<td>Negative valence</td>
<td>4.42 (1.36)</td>
<td>4.12 (1.33)</td>
</tr>
<tr>
<td>Mean values</td>
<td>5.16 (1.29)</td>
<td>4.72 (1.39)</td>
</tr>
</tbody>
</table>

*p*-value for difference .002 .015

*Note. n = 367. Values in parentheses are standard deviations.*
Table 14

Descriptive Statistics and Intercorrelations from Study 7

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>α</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Manipulation check: Temporal comparisons</td>
<td>4.15</td>
<td>2.46</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Manipulation check: Social comparisons</td>
<td>4.14</td>
<td>2.60</td>
<td>--</td>
<td>-.78</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Manipulation check: Positive valence</td>
<td>4.19</td>
<td>2.21</td>
<td>--</td>
<td>.24</td>
<td>-.06</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Manipulation check: Negative valence</td>
<td>4.12</td>
<td>2.17</td>
<td>--</td>
<td>-.09</td>
<td>.02</td>
<td>-.79</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Procedural fairness</td>
<td>5.30</td>
<td>1.59</td>
<td>.98</td>
<td>.24</td>
<td>-.12</td>
<td>.61</td>
<td>-.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Interpersonal fairness</td>
<td>5.05</td>
<td>1.46</td>
<td>.89</td>
<td>.31</td>
<td>-.24</td>
<td>.54</td>
<td>-.38</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td>7. Informational fairness</td>
<td>4.58</td>
<td>1.38</td>
<td>.87</td>
<td>.25</td>
<td>-.12</td>
<td>.52</td>
<td>-.38</td>
<td>.63</td>
<td>.67</td>
</tr>
<tr>
<td>8. Individualized treatment</td>
<td>4.15</td>
<td>1.68</td>
<td>.83</td>
<td>.54</td>
<td>-.42</td>
<td>.41</td>
<td>-.24</td>
<td>.48</td>
<td>.53</td>
</tr>
</tbody>
</table>

*Note: n = 400. rs greater than .09 are significant at .05 level*
Table 15

**Perceptions of Fairness and Individualized Treatment from Study 7**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Procedural fairness</th>
<th>Interpersonal fairness</th>
<th>Informational fairness</th>
<th>Individualized treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Temporal</td>
<td>Social</td>
<td>Temporal</td>
<td>Social</td>
</tr>
<tr>
<td>Positive valence</td>
<td>6.32 (.91)</td>
<td>5.96 (1.09)</td>
<td>6.02 (.97)</td>
<td>5.31 (1.31)</td>
</tr>
<tr>
<td>Negative valence</td>
<td>4.78 (1.58)</td>
<td>4.07 (1.57)</td>
<td>4.90 (1.40)</td>
<td>3.93 (1.29)</td>
</tr>
<tr>
<td>Mean values</td>
<td>5.56 (1.50)</td>
<td>5.04 (1.64)</td>
<td>5.47 (1.32)</td>
<td>4.64 (1.47)</td>
</tr>
</tbody>
</table>

*p*-value for difference: .001 < .001 .002 < .001

*Note. n = 400. Values in parentheses are standard deviations.*
### Table 16

*Mediation Analyses from Study 7*

<table>
<thead>
<tr>
<th></th>
<th>Procedural fairness</th>
<th>Interpersonal fairness</th>
<th>Informational fairness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point estimate of direct effect</td>
<td>-.233</td>
<td>.139</td>
<td>-.518</td>
</tr>
<tr>
<td>Bias-corrected 95% CI of direct effect</td>
<td>(-.529, .057)</td>
<td>(-.139, .400)</td>
<td>(-.732, -.316)</td>
</tr>
<tr>
<td>Point estimate of indirect effect</td>
<td>.751</td>
<td>.687</td>
<td>.953</td>
</tr>
<tr>
<td>Bias-corrected 95% CI of indirect effect</td>
<td>(.563, .972)</td>
<td>(.506, .901)</td>
<td>(.749, 1.180)</td>
</tr>
</tbody>
</table>

*Note.* $n = 400$.

The number of bootstrap resampling = 5,000.
Figure 6. Perceptions of fairness from Study 6

Note. Error bars represent standard errors.
Figure 7. Perceptions of fairness and individualized treatment from Study 7

Note. Error bars represent standard error.