A Thesis

entitled

The Impact of Comparative Thought: Exploring the Similarities and Differences between Social and Counterfactual Comparisons

by

Elizabeth Farrell Gallinari

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Master of Arts Degree in Experimental Psychology

_____
Jason Rose, PhD, Committee Chair

_____
Jon Elhai, PhD, Committee Member

_____
JD Jasper, PhD, Committee Member

_____
Barry W. Scheuermann, PhD, Interim Dean
College of Graduate Studies

The University of Toledo

December 2020

An Abstract of

The Impact of Comparative Thought: Exploring the Similarities and Differences between Social and Counterfactual Comparisons

by

Elizabeth F. Gallinari

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Master of Arts Degree in
Experimental Psychology

The University of Toledo
December 2020

Comparison is pervasive and an important part of the human experience. People are often motivated to use comparison standards to form evaluations, regulate their emotions, make decisions, and guide behavior. Comparisons can be made to different standards, such as to another person (social) or to *what might have been* if some aspect of the past had been different (counterfactual). Often, researchers dichotomize comparisons directionally, such that they can take an upward or downward focus. Upward comparisons involve a better off target, either another person or an alternate reality depending on whether it's a social or counterfactual source respectfully, while downward comparisons focus on worse off targets. The direction of comparison can influence a variety of downstream reactions, such as self-evaluations, motivations, affect and future performance. Despite ample evidence of both socially- and counterfactually- based comparison processes, few studies have directly examined these sources under the same framework. The current research aimed to fill this gap in the literature by directly comparing social and counterfactual comparative processes and outcomes in terms of their impact on cognition, affect and behavior within the same study. In three experiments, participants were prompted to

consider either social or counterfactual information that was upward or downward in academic/performance related contexts following 2 × 2 between-participants designs. In Experiment 1, participants read comparative-based scenarios where the target person made either upward or downward comparisons to social or counterfactual sources on things like exam performance. The findings suggested that those reading upward comparative scenarios reported poorer self-evaluations and greater behavioral motivation compared to those in the downward conditions, while those who read social comparative scenarios reported lower self-evaluations compared to the counterfactual scenarios. To extend the findings of Experiment 1, participants in Experiments 2 and 3 engaged in performance-based tasks, received upward or downward feedback involving social or counterfactual sources, provided self-evaluations and affective responses to the feedback, and then completed the performance-based task again to assess behavioral change (in Experiment 3 only). The results revealed that upward comparison information led to poorer self-evaluations and greater negative affect compared to downward comparison information. However, no condition differences were found for behavioral intentions or future performance. Overall, the results revealed some consistencies and some differences in the impact of social and counterfactual comparisons across studies. Theoretical and practical implications of the results are discussed.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CFT.....Counterfactual Thought

MR Task.....Mental Rotation Task

PR Task.....Pursuit Rotor Task

REM.....Reflection and Evaluation Model

SCO.....Social Comparison Orientation

Chapter One

Introduction

Comparison is pervasive and an important part of the human experience (Festinger, 1954). Indeed, people are often motivated to use comparison standards to form evaluations, regulate their emotions, make decisions, and guide behavior (Klein, 1997; Wheeler & Miyake, 1992; Wills, 1981; Zell & Alicke, 2009a, 2010a). Importantly, there is no shortage of comparison standards available. For example, when reflecting upon a recent math exam performance, a student might consider how others in the class performed on the exam (*social comparison*), how they performed on a past math exam (*temporal comparison*), how they performed on an exam in a different subject (e.g., English; *dimensional comparison*), or how they could have performed differently (*counterfactual comparison*). Moreover, it is clear that these types of comparative-based thoughts comprise a substantial portion of our waking mental activity. Indeed, using an experience sampling methodology, Summerville and Roese (2008) found that 12% of all thoughts were comparative. While comparison processes and outcomes have been studied for decades, relatively limited research has evaluated the implications that divergent comparison sources (e.g., social comparisons, counterfactual comparisons) might have on self-evaluations, motivations, and emotions. The current research examined similarities and differences in social and counterfactual comparisons in terms of the impact of such comparisons on motivation, emotion, self-evaluation, and behavior in performance-relevant contexts (Olson, Buhrmann, & Roese, 2000).

Source of Comparison

Social Comparison

Social comparison is the process by which one looks to others to learn about the self. In original conceptualizations of this construct, social comparison was viewed as a volitional process that was engaged when a person lacked objective information about the self and wanted to self-evaluate (Festinger, 1954). Indeed, Social Comparison Theory (Festinger, 1954) posits that humans have a drive to gain accurate self-evaluations through comparison to others. Further developments in the area of social comparison have shown that people engage in social comparisons for other reasons, such as self-enhancement (Willis, 1981), self-improvement and inspiration (Lockwood & Kunda, 1997), predicting future behavior (Blanton et al., 1999) and motivating behavior change (Markman & McMullen, 2003). Moreover, while social comparison is often consciously and purposefully initiated (Taylor et al., 1996), social comparison can also be unintentional or forced (Gibbons & Buunk, 1999) and, depending on the context, activated spontaneously and unconsciously (Mussweiler et al., 2004a). Importantly, social comparison is most often initiated when someone is feeling uncertain (Gibbons & Buunk, 1999; Lee, 2014), negative (Gibbons, 1986; Willis, 1981, 1987), or threatened (Aspinwall & Taylor, 1993; Willis, 1981).

Counterfactual Comparison

While social comparisons focus on another *person*, counterfactual comparisons focus on what *might have been* if some aspect of the past had been different (Goodman, 1947; Roese, 1997). In other words, whereas social comparisons focus on the actual or predicted performance, behavior, ability, or trait of another person, counterfactual comparisons involve a hypothetical, imagined reality for the self (Olson et al., 2000). Like social comparisons, counterfactual thoughts are typically activated by negative or surprising experiences that could have been

avoided, such as failing an exam or crashing a car (Roese, 1997). Counterfactual thoughts compare the experienced (actual) reality to hypothetical realities and facilitate thoughts about what could have been done to avoid the negative experience (Epstude & Roese, 2008; Markman et al., 1993; Roese et al., 1999). However, it is important to note that positive outcomes can also elicit counterfactual thoughts, such as when people made it through something stressful unscathed (e.g., just made it on a flight or avoiding a car crash).

Comparison Direction and the Impact on Affect and Cognition

When describing comparisons, it is instructive to discuss different "directions" of comparison. Most researchers dichotomize comparisons along the vertical dimension such that comparisons can take an upward or downward directionality. First, in social comparison, upward comparisons occur when looking to a superior target considered better than the self (Buunk & Gibbons, 2007; Suls et al., 2002). For example, if a person took an exam and failed, they may compare their score to a friend who did better on the exam. Contrariwise, downward comparisons occur when looking to an inferior target, such as if a person had done well on the exam, they may compare the score to a friend who did worse (Suls et al., 2002; Willis, 1981). Second, and similar to social comparisons, counterfactual comparisons can also be directional (Markman et al.,1993; Rim & Summerville, 2014; Roese, 1994; Summerville & Roese, 2008). Upward counterfactuals involve comparisons that are better than the past, experienced reality. For example, if a person was involved in a car crash, she may think "If only I had not run the red light, I would not have been in an accident". On the other hand, downward counterfactuals involve comparisons that are worse than the past, experienced reality. For example, if a person

avoided a bad car crash, he may think "If I hadn't left a couple minutes late this morning, I would have been killed".

The direction of the comparison can influence a variety of downstream reactions. Evidence for the impact of upward and downward comparisons, both social and counterfactual, on things like self-evaluations, behavioral intentions, and emotions is varied and often depends upon factors such as motivation, current mood state, and similarity to the comparison target (Buunk & Gibbons, 2007; Epstude & Roese, 2008; Roese, 1994; Roese & Morrison, 2009; Suls et al., 2002; Wheeler & Miyake, 1992; Wood, 1989; Willis, 1981). However, the typical trend in the literature, at least when it comes to the impact on immediate self-appraisals, self-evaluations, and affect following comparison, is that upward (downward) comparisons are associated with poorer (better) self-appraisals, self-evaluations, and affect (Gilvovich & Medvec, 1995; Roese & Morrison, 2009; Suls et al., 2002). Indeed, for instance, when looking for ways to self-enhance and increase positive mood, one may compare to someone who is less fortunate or think about how a situation could have been worse to boost well-being (Roese, 1994; Suls et al., 2002; Wheeler & Miyake, 1992; Willis, 1981).

Importantly, counterfactual and social comparisons have been shown to be mediated by similar processes (Markman & McMullen, 2003) and produce similar outcomes (Olson et al., 2000). Indeed, to provide a theoretical framework for understanding the processes and outcomes that arise following comparisons, Markman and McMullen (2003) developed the Reflection and Evaluation Model (REM). The REM states that the same basic processes of mental simulation are common to any type of comparative thinking, including social and counterfactual comparisons. In particular, the model asserts that, when prompted with comparative information,

a person can engage in one of two distinct modes. First, a person can enter a reflection mode to determine whether the information from the comparison target is true to oneself. Second, and more commonly, a person can engage in an evaluative mode wherein one uses the comparison target as a reference point against which to evaluate their present standing. Engagement in reflection or evaluation impacts whether a person's cognitions, judgments, and emotions will assimilate to (go in the direction of) or contrast away from (go in the opposite direction of) the comparison standard. For example, imagine that Emma and another girl join a high school basketball team at the same time, and it is clear that the other girl is superior to Emma. If Emma views herself as highly similar or close to the teammate or that her teammate's status is attainable (Aspinwall, 1997; Lockwood & Kunda, 1997, 1999; Mussweiler, 2001, 2003; Mussweiler et al., 2004b), she might evaluate her basketball ability and feelings about basketball positively through assimilation to her teammate's impressive ability and performance. Contrariwise, if Emma viewed herself as distinct from the teammate or that her teammate's status was difficult to achieve, she would contrast away from the teammate and have negative thoughts and feelings about her basketball ability. Likewise, this same type of process occurs with directionally-based counterfactual comparisons (Markman et al., 1993; Markman et al., 2008; McMullen & Markman, 2002; Medvec et al., 1995; Smallman & Summerville, 2018; Roese, 1994). For example, imagine that Jennifer stayed out late with friends the night before an exam and she did not do as well as she would have liked. Assimilation to an upward comparison about how she could have done better might actually make her feel good (e.g., hopeful about what might be possible in the future), whereas contrast would make her feel bad (e.g., disappointed, regretful).

In summary, most comparison outcomes result from engagement of an evaluative mode of processing and contrast-like effects where upward comparisons reduce self-evaluations and increase negative feelings. While past models have discussed social and counterfactual comparisons under the same framework, there are actually very few studies that directly examine both comparisons within the same study. Interestingly, although research shows that both upward (vs. downward) social and counterfactual comparisons increase (decrease) negative affect and decrease (increase) self-evaluations, there has been some recent evidence to suggest that social comparisons sometimes produce a stronger impact on such variables than other types of comparisons (e.g., temporal, dimensional; Edmonds et al,m 2020; Müller-Kalthoff et al., 2017; Strickhouser & Zell, 2015). However, none of this recent research has specifically examined counterfactual and social comparisons together.

Comparison Direction and its Impact on Goal-Directed Cognition and Behavior

An important question for the current research is what impact different types of comparisons have on goal-directed behavior. While upward comparisons can temporarily reduce well-being, self-appraisals, and affect, they may also offer environments for positive goal-directed behavior and improvement of one's future outcomes. Specifically, upward counterfactuals create mental environments for self-improvement motives and upward social comparisons can provide role models for paths to self-improvement (Buunk & Gibbons, 2007; Epstude & Roese, 2008; Wood, 1989). In the example of Emma and her superior basketball teammate, the negative self-appraisals and feelings resulting from upward social comparison could prompt Emma to try harder, practice more, and push herself to improve in order to catch up with her teammate (see Blanton et al., 1999; Johnson & Stapel, 2010; Lockwood & Kunda,

1997). In counterfactual comparisons, when Jennifer feels disappointed after considering how she could have done better on the exam (upward counterfactual comparison), she may be prompted to make changes to her routine to improve her next exam performance (e.g., get more sleep, study more; see Epstude & Roese, 2008; Markman et al., 1993; Roese, 1994). Indeed, past research supports the functionality of both social and counterfactual comparative thoughts to solve problems, prepare for the future, and work harder while simultaneously making people feel better about their current situation (Martin et al., 2002; Roese, 1994; Roese & Olson, 1993). That said, some recent research suggests that upward social comparisons can sometimes be less beneficial for future motivation and behavior than other types of comparisons (e.g., dimensional) due to their threatening nature (Edmonds et al., 2020). However, no studies have attempted to directly examine this possibility in social vs. counterfactual comparisons.

Affect as a Mediator in the Relationship between Comparisons and Behavior

As noted above, both social and counterfactual comparisons have been theorized (and sometimes demonstrated) to guide goal-directed future behavior. Typically, the result is that upward comparisons promote adaptive behavioral changes, such as trying harder on a future exam following exposure to a superior comparison (Roese, 1994; van de Ven, 2017). Beyond looking at behavioral outcomes, it is important to examine what psychological mechanisms prompt such behavioral changes. As alluded to above, here we highlight the relevance of affective states in prompting behavioral changes (Epstude & Roese, 2008). Given that comparisons tend to elicit emotional responses (as noted above) and given that emotions are often viewed as functional guides to adaptive behavioral change (van de Ven et al., 2011), it is surmised that emotions could be responsible for prompting such behavior change. Indeed, prior

research on counterfactual thought has found that the negative affect created from counterfactual comparisons may be an important element that motivates behavior change in similar future circumstances (Markman & McMullen, 2003; Markman et al., 2006; McMullen & Markman, 2000). Moreover, the negative affect experienced following an upward social comparison can, in turn, also induce motivation for future improvement (Collins, 1996; Park & Park, 2017; van de Ven, 2017). Overall, counterfactual and social comparisons tend to be similar in that the experience of negative affect (generally) prompts adaptive behavioral change. This idea is consistent with prior research on emotion theory suggesting that the experience of negative affect acts as a cue to the organism that something is wrong, prompting attention and a change in response patterns. On the other hand, the experience of positive affect is a cue that everything is fine and that an organism can coast and maintain the status quo (Easterbrook, 1959; Huntsinger et al., 2014). In summary, general affective experiences can be viewed as a mechanism for explaining the impact of comparison (regardless of source) on adaptive behaviors.

Although counterfactual and social comparisons share a default response pattern in terms of reactions to valence—with upward (downward) social and counterfactual comparisons producing general negative (positive) affect (Roese, 1997; Wheeler & Miyake 1992)—there may also be subtle differences in the discrete negative emotions that people experience in reaction to social vs. counterfactual comparisons. For example, counterfactual comparisons tend to result in feelings of regret (Roese, 1994), guilt (Niedenthal et al., 1994), pride (Tangney, 1990) or gratitude/luck (Weiner, 1985), whereas social comparisons tend to result in emotions like envy (Smith & Kim, 2007), shame (Wehrens et al., 2010), inspiration (Lockwood & Kunda, 1997), and optimism (Wehrens et al., 2010). It could be the case that the specific emotions prompted by

upward social and counterfactual comparisons differ and that a consideration of these discrete emotions (vs. more general positive or negative affectivity) can better account for the relationship between comparisons and goal-directed motivation and behavior.

Focus of the Current Study

While various models related to both social and counterfactual comparisons have been offered (Markman & McMullen, 2003), very few studies have directly compared processes and outcomes related to these two types of comparison sources in terms of their impact on cognition, affect, and behavior. Moreover, relatively few studies have examined the impact of comparisons on goal-directed behavior and no studies to our knowledge have contrasted the impact of social and counterfactual comparison sources on such behavioral outcomes (Olson et al., 2000). Finally, no studies to our knowledge have examined the mediating roles of general affectivity and discrete emotions in these relationships. The current research aimed to fill these gaps in the literature. In this research, we examined the impact of different sources of upward and downward comparisons on goal-directed motivations and behaviors, and whether such effects were mediated by general affect and specific discrete emotions (e.g., envy, regret). We explored our research questions in academic- and performance- related domains, such as exam outcomes and performances on tasks. In such paradigms, we explored whether comparative-based prompts (upward or downward, social or counterfactual) impacted self-evaluations, affect and emotions, motivations (e.g., intentions to perform better or try harder), and changes in future behaviors themselves (e.g., actual performances).

More specifically, we describe three experiments wherein participants were prompted to consider social or counterfactual information that was either upward or downward in academic/

performance-related contexts (hence 2 × 2 between-participants designs were used). After considering the comparative-based information, participants answered questions related to their hypothetical or actual performance or outcome. In Experiment 1, participants read comparison-based scenarios where the target person makes upward or downward comparisons to social or counterfactual sources on things like exam performances before providing self-evaluations and behavioral intentions for the relevant contexts. In Experiments 2 and 3, participants engaged in performance-based tasks, received upward or downward comparison feedback/prompts involving social or counterfactual sources, provided self-evaluations and affect ratings in response to the feedback/prompts, indicated their future motivations/intentions, and completed the performance-based task again to assess goal-directed behavioral change. Moreover, in Experiment 3 (but not Experiment 2), we measured actual performances.

Based on prior research, we developed the following core hypotheses:

Main Hypotheses

H1: There will be a main effect of comparison direction, such that upward (vs. downward) comparisons will produce more negative affect, poorer self-evaluations, more adaptive behavioral intentions and motivations, and improved future performances (Experiment 3; Collins, 1996; Epstude & Roese, 2008; Lockwood & Kunda, 1997).

H2: There will be an interaction, such that the effect of comparison direction described in H1 was expected to be stronger for affect and self-evaluations when making social comparisons (Müller-Kalthoff et al., 2017; Strickhouser & Zell, 2015; Zell & Alicke, 2009a, 2010a), but stronger for motivations and actual performances when making counterfactual comparisons (Edmonds et al., 2020). As noted above, although research

has shown that social comparisons have a more robust impact than other comparisons on affect and self-evaluations, such social comparisons can simultaneously be threatening, which can decrease motivation and, potentially, goal-directed behavioral effort. As such, other comparison sources such as counterfactual comparisons, might be expected to have a weaker impact on affect, but a stronger impact on motivation and behavior due to such internal (as opposed to external) sources being less threatening (Edmonds et al., 2020).

H3: Changes in affect will statistically mediate the relationship between comparison direction and performance, such that upward comparisons that lead to more negative affect should facilitate better performances (Experiment 3). However, we further surmised that the specific type of discrete emotion most critical for demonstrating mediation might depend upon exposure to counterfactual or social comparison information, such that specific counterfactual-relevant negative emotions like regret would be most relevant for upward counterfactual comparison feedback (Gilovich & Medvec, 1995; Kahneman & Miller, 1986), whereas specific social comparison-relevant negative emotions like envy would be more important for upward social comparison feedback (Alicke & Zell, 2008; Park & Baek, 2018; Smith & Kim, 2007).

Chapter Two

Method

Experiment 1

To provide preliminary evidence for the impact of social and counterfactual comparisons on self-evaluations and intentions for goal-directed behavioral change, a pilot study was conducted involving academic/performance-based scenarios. Participants in the study were randomly assigned to read different versions of hypothetical scenarios wherein we manipulated the outcome for the target person to involve upward or downward comparisons (direction) to social or counterfactual standards (source). Participants then answered questions about how they would feel and think after making such comparisons, as well as how they might change their behavior in the future if they were the person in the scenario.

Participants and Design

Participants were 150 (107 Women) undergraduate students who were recruited from psychology courses and received course credit for participating. The median age was 19 years old (M =20.06, SD = 3.41). Participants were randomly assigned to one cell in a 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) between-participants design. A G*power analysis using an alpha of .05 and an effect size (η^2) of .06 for the interaction suggested 128 participants for the study would be required to achieve a power of .80 (Faul et al., 2007).

Procedure and Measures

Participants signed up for a study about perceptions of others and the self on SONA Systems and came to the lab in groups of 1-4. Upon arrival into the lab and after providing

informed consent, participants were seated at individual computers and provided general instructions by the experimenter. All portions of the study were completed using Qualtrics.

Participants were asked to read scenarios while imagining they were the person described. Each participant saw a total of two scenarios in which the direction of comparison (upward or downward) and the type of comparative prompt (social or counterfactual) differed depending on the condition. In the GPA scenario, participants imagined they were in high school applying to colleges and that they had a 3.2 GPA. In the exam scenario, participants imagined they were a junior in college and that they correctly answered 21 questions out of 30 on their Principles of Psychology exam. Critically, participants viewed different endings to the scenarios depending upon their experimental condition. First, comparison direction was manipulated by having participants learn that their GPA or exam score was either lower (upward comparison) or higher (downward comparison) than a particular source. Second, comparison source was manipulated by altering the source of comparison to be another friend or the class average (social) vs. an imagined outcome (counterfactual; see Appendix A for full scenarios).

After reading each scenario, participants provided self-evaluations (e.g., "If you were this person, how would you rate your performance on the Principles of Psychology Exam?") and future behavioral intentions (e.g., "If you were this person, how likely is it that you would try or study harder when taking classes next semester") using 5-point Likert scales (1=*very poor/not at all likely/not at all motivated*; 5 = *very good/very likely/very motivated*). Responses to the self-evaluation items (αs > .72) and behavioral intentions items (αs > .78) were aggregated separately for each scenario set. Participants also answered questions about perceived similarity to the

person described in each scenario (1=*not at all*; 5=*extremely*). After completing all questions, participants were debriefed, thanked, credited and dismissed.

Results and Discussion

To examine the impact of our variables, we submitted self-evaluation indices and behavioral intentions indices to separate 2 (scenario: GPA or exam) × 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) mixed-model ANOVAs, with the first factor as a within-participants factor.

Self-Evaluations

The within subject factor of scenario type did not have any main or interactive impact on the results (Fs < 2.43, ps > .34), suggesting scenario type was not relevant to the self-evaluation ratings. Thus, all findings reported below are collapsed across the two scenarios. Regarding the more critical between-subjects factors, there was a significant main effect of direction, $F(1, 146) = 3.92$, $p = .05$, $\eta_p^2 = .026$, such that self-evaluations in the upward condition ($M = 3.17$, $SD = .65$) were significantly lower compared to the downward comparison condition ($M = 3.36$, $SD = .57$, *Cohen's d* = .31. This result was consistent with H1. There was also an unexpected, significant main effect of comparison source, $F(1, 146) = 4.09$, $p = .045$, $\eta_p^2 = .027$, such that self-evaluations in the social comparison condition ($M = 3.17$, $SD = .60$) were significantly lower compared to the counterfactual comparison condition ($M = 3.37$, $SD = .62$), *Cohen's d* = .33. Finally, and as predicted in H2, there was a significant interaction between comparison direction and source, $F(1, 146) = 5.06$, $p = .026$, $\eta_p^2 = .033$. Specifically, the nature of this effect was such that the impact of comparison direction was stronger when exposed to social comparison scenarios than counterfactual comparison scenarios (See Figure 1). As predicted, results from a

follow up test indicated that participants reported significantly lower self-evaluations when in the upward social conditions (M = 3.16, SD = .68) compared to the downward social conditions (M = 3.57, SD = .48), $t(72) = -3.01$, p = .004, d = .70. For counterfactual comparisons, there were no significant differences found for participants in the upward counterfactual condition (M = 3.18, SD = .62) relative to the downward counterfactual comparison condition (M = 3.16, SD = .59), $t(74) = .190$, p =.85, d = .03.

Behavioral Intentions

The within subject factor of scenario type did not have any main or interactive impact on comparison direction and source (Fs <.099, ps > .75), so we again collapsed across scenarios. Regarding the more critical between-subjects factors, there was a significant main effect of direction, $F(1, 146) = 9.09$, p = .003, η_p^2 = .059, such that participants in the upward comparison condition (M = 4.57, SD = .47) reported greater behavioral intentions for future improvement compared to those in the downward comparison condition (M = 4.29, SD = .64), *Cohen's d* =.50. This result was consistent with H1. The main effect of comparison source was not significant, $F(1, 146) = .24$, p = .623. Finally, unlike for self-evaluations, and partially consistent with H2, the comparison source × comparison direction interaction was not significant, $F(1, 146) = .005$, p = .95. This result suggests that behavioral intentions are impacted similarly by both counterfactual and social sources.

Experiments 2 and 3

Results from Experiment 1 revealed that there are similarities and differences between social and counterfactual comparisons regarding self-evaluations. First, hypothesis 1 was supported. Specifically, a main effect of comparison direction was found such that participants in

the upward condition reported poorer self-evaluations and greater behavioral motivation compared to those in the downward condition. Hypothesis 2 was partially supported in that there was some divergence in the impact of social vs. counterfactual comparison depending upon whether we considered self-evaluations or intentions. Indeed, the impact of upward comparison in reducing self-evaluations was stronger for those who read social comparison than counterfactual comparison scenarios. Thus, the portion of H2 indicating that self-evaluations would be more strongly impacted by social than counterfactual comparisons was supported. However, the impact of comparison direction on behavioral intentions did not differ between comparison sources. Thus, the portion of H2 indicating that intentions would be more strongly impacted by counterfactual than social comparisons was not supported. As noted previously, we argued that people might be threatened by social comparison feedback, which could impair motivation and intentions. One possible reason for the lack of a threat impact here could be that participants were asked to read and imagine themselves in hypothetical scenarios. As such, any emotional reactions or experiences of threat may have been dulled using this type of methodology. Real comparison feedback may be more meaningful and may, in turn, increase feelings of threat and affect.

There were several limitations to Experiment 1 that motivated conducting Experiments 2 and 3. First, as noted above, participants read hypothetical scenarios in Experiment 1. Beyond potentially mitigating (or completely altering) the impact of the comparison information, the use of hypothetical scenarios did not allow us to examine actual goal-directed behavior in response to comparison information. Moreover, prior research in a variety of fields reveals that sometimes

responses to hypothetical scenarios do not align with responses to real situations (Bostyn et al., 2018). Second, Experiment 1 did not examine the impact of comparisons on affect and emotion.

To expand upon the findings of Experiment 1 and address these core limitations, participants in Experiments 2 and 3 engaged in a performance-based target task and received comparative feedback about their performance. Participants provided self-evaluations, affect ratings, and motivations after receiving feedback and were given the opportunity to perform the task and provide (some) ratings again post-feedback. This design allowed us to more fully examine the relationship between comparative information and critical outcome measures, including the mediating impact of general affect and discrete emotions. Moreover, to increase generalizability, an online adult sample was recruited from Amazon Mechanical Turk (MTurk; Litman et al., 2016). Previous research has shown that collecting data online using MTurk is as reliable and valid as lab-based experiments and college student samples, and also increases the diversity in demographic characteristics relative to college student samples (Buhrmester et al., 2011; Hauser & Schwarz, 2015).

To ease exposition and because the methodologies, core variables, and analyses were nearly identical, we have partially combined descriptions of Experiments 2 and 3 below. The primary difference between Experiments 2 and 3 was the use of different target performance tasks (of which only the task used in Experiment 3 tracked actual performance data). The use of two distinct tasks across experiments allowed us to rule out that something about the specific task itself was critical for the results. More specifically, for Experiment 2, we used an online version of the Pursuit Rotor Task (PR Task) task wherein participants followed a rotating ball around a circle with their computer mouse. However, performance was not actually being

recorded as the participants were only viewing a video of the task. For Experiment 3, we used an online version of a Mental Rotation Task (MR Task) where participants indicated (as quickly as possible within a short time frame) which of a pair of rotated objects matched a target object. For participants in Experiment 3, we recorded the accuracy of responses in the MR Task.

Participants and Design

Within each experiment, participants were randomly assigned to one of five conditions in a 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) between-participants design with a floating no-comparison control condition. An *a priori* G*power (Faul et al., 2007) analysis using an alpha of .05 and an effect size (η^2) of .06 for the interaction suggested 128 participants for each experiment would be required to achieve a power of .80.

For Experiment 2, participants were 131 adults (68 Women) recruited through MTurk and received $3 for participation. The median age was 37 (M = 39.37, SD = 11.84) and the racial make-up of the sample was 79.4% White, 8.4% Black, 6.9% Asian, 3.8% mixed-race.8% Native American, and .8% unknown. A total of 129 adults (57 Women, 2 Non-Binary) were recruited for Experiment 3, with a median age of 36 (*M*=38.94, *SD*=11.06) and the racial makeup of the sample was 82.2% White, 8.5% Black, 6.2% Asian, 2.3% mixed-race, and .8% Native Hawaiian. Participants who failed the suspicion probe question(s) described below (Exp. 2 *n*= 25; Exp. 3 *n*=19) were excluded from analysis.

Procedure and Materials

Experiments 2 and 3 were conducted in successive weeks on MTurk and posted on CloudResearch (cloudresearch.com). MTurk is an online marketplace where participants

complete research studies for monetary compensation. CloudResearch is a crowdsourcing platform that operates in conjunction with MTurk, allowing for researchers to use selection criteria and other features (Chandler et al., 2019). Participants had to be U.S. residents with a 99% approval rating for prior work in MTurk and must have completed at least 5,000 Human Intelligence Tasks. Participants signed up online for each experiment through CloudResearch. Measures were in place to ensure participants only completed one of the experiments. All portions were completed using Qualtrics. Participants were told that we were interested in examining factors that influence whether people will be successful in life (e.g., relationships, careers). They were further told that performance on specific types of lab-based tasks have been shown to be indicators of future success in life. This type of cover story has been used by other comparison researchers to bolster motivation for success and accuracy, and to make the task and feedback more meaningful (Buckingham & Alicke, 2002; Bruchmann, 2017). Participants were further told that feedback may be provided at some point during the study and that this feedback would allow them to assess their progress and evaluate their achievement on the tasks. Participants were also told that they would have the opportunity to repeat at least one of the tasks in order to assess their ability to improve and learn from feedback/mistakes, which they were told was another important predictor of future success in life.

After these basic instructions, participants answered preliminary questions about themselves, including baseline measures of mood, demographics, and questions about experience with puzzles and other cognitive tasks. Next, participants engaged in a series of four computer-based cognitive tasks (described below). Three of these tasks were non-target tasks (alternative uses task, anagram task, and visual search task) and always came before the target task. These

non-target tasks were included to bolster the cover story that we were interested in how participants perform on a battery of tasks. Specifically, for the Anagram Task (Ammons, & Ammons, 1959), participants were given a list of words and asked to rearrange letters to create different words, such as turning 'cat' to 'act' or 'meat' to 'team'. For the Alternative Uses Task (Guilford, 1967), participants were given an item (e.g., stapler) and asked to list several alternative uses for that item (e.g. paper weight, hammer, weapon, etc.). Finally, for the Visual Search Task (Williams, 1996), participants looked for a specific object in a field of distractors, such as an upright orange T hidden among other letters.

Experiment 2 Task

As noted above, the main distinction between Experiments 2 and 3 was the target task. For Experiment 2, the target task was based on the Pursuit Rotor Task (PR task; Ammons et al., 1955), where participants attempt to keep their computer mouse on a red target as it moves quickly around a circle on the screen (see Appendix C for visual example). Though participants believed their performance was recorded, they were actually viewing a video of the PR task[1]. The PR task was chosen as the target task because 1) performance success on the task is ambiguous and thus the bogus feedback should appear authentic and believable and 2) it has been proceduralized to improve performance, thus indicating that improvement across trials (even if not captured here) is feasible (Mueller, 2010). Prior testing of the PR Task in previous research (Piper, 2011) revealed that the default speed (2 full circular rotations/15 seconds) was

[1] The PR task was originally designed as an in-person lab task, but the COVID-19 pandemic caused the study to move to an online format in the middle of data collection. The PR task was retained given the previous approval by the thesis committee yet there were challenges moving the task to an online environment and as such, we were unable to collect performance metrics on this task.

moderately challenging yet avoided both ceiling and floor effects. To validate the ambiguity of the task for the current study, pilot testing was conducted. Participants (n=10) engaged in five rounds of the task, each at a different speed, and were asked to report their perceived accuracy for time on target. The scores for each speed were averaged to determine the ideal speed to support the bogus feedback, a 73% time on target accuracy score (see Table 1). The results also supported the ambiguity of the task and found that believability of the feedback was best achieved when the target moved 2 full circular rotations every 12 seconds, which falls slightly above the default speed.

Experiment 3 Task

For Experiment 3, the target task was the Mental Rotation Task (MR Task; Vandenberg, & Kuse, 1978; Stoet, 2017). For each trial of the MR task, participants had five seconds to compare two 2D rotated objects to determine which one matched the original, non-rotated object (See Appendix D for visual example). Similar to the PR Task, performance on the MR task is ambiguous as participants were instructed to answer relatively quickly. Pilot testing of the MR task (n = 5) suggested the task is moderately challenging yet improvement on the task is feasible. Unlike the PR Task, participants' performance in the MR Task was recorded and used to assess if performance changed pre- and post- feedback. Specifically, we counted the number of correct responses for both the pre- and post- feedback trials of the MR Task.

Upon completion of the set of tasks, participants received one of the four types of feedback, ostensibly based on their performance on the target task (in actuality, the feedback was pre-determined). More specifically, participants in the feedback conditions received one of four combinations of feedback (see Table 2 for PR Task and Table 3 for MR Task). Critically, two

key details of each comparative feedback prompt differed: 1) the source of comparative prompt (social or counterfactual) and 2) the direction of the prompt (upward or downward).

Experiment 2 Feedback

In Experiment 2, participants in all conditions received an accuracy score on the PR task of 73%, meaning they stayed on target within the acceptable error range 73% of the time (value determined through evaluation of prior comparison research; Alicke & Zell, 2008; Bruchmann, 2017; Strickhouser & Zell, 2015). First, participants in the *upward counterfactual condition* were asked to imagine that they had performed better on the task than they actually did (i.e., stayed on target 84% of the time). Second, participants in the *downward counterfactual condition* were asked to imagine that they had performed worse than they actually did (i.e., stayed on target 62% of the time). Third, participants in the *upward social comparison condition* were informed that the last MTurk participant to complete the task had performed better than them (i.e., stayed on target 84% of the time). Fourth, participants in the *downward social comparison condition* were told that the last MTurk participant had performed worse (i.e., received an accuracy score of 62%). The last participant was used, rather than an average score, because research has shown comparing to an individual's score, instead of an average score, has a greater impact on self-evaluations (Buckingham & Alicke, 2002; Zell & Alicke, 2009b, 2010b). Participants in the control condition received information about their accuracy score of 73% but were not provided with any comparison context. Finally, as an additional part of the feedback process, participants were informed that they would have a chance to complete the target task again to see if they could improve their score by evaluating their ability and learning from their mistakes.

Experiment 3 Feedback

In Experiment 3, participants in all conditions received an accuracy score of 73%, meaning 73% of the time they chose the correct rotated image. As in Experiment 2, participants in the feedback conditions received one of four combinations of feedback, again with two key details differing: 1) the source and 2) the direction. Participants in the control condition were provided an accuracy score but were not given any comparison information. It is important to note that the feedback provided was identical to what was provided in Experiment 2 in terms of the values.

Immediately after receiving the prompt/feedback, all participants were asked to free recall their score and, if applicable, the comparison score (social or counterfactual) as a manipulation check. Participants were also asked to consider their strategy and approach (and feedback if applicable) upon completing another round of the task ("Please write a few sentences about your strategy when you complete the circle/mental rotation task again in a few minutes. Will you do anything differently? How do you feel you will do? Was the feedback useful to you [if applicable]?"). They also answered questions related to their self-evaluations, motivations for performing well, and affect before completing the target task again. Once participants completed the second round of the target task, they were asked a final set of questions related to their mood and self-evaluations. Participants were probed for suspicion using a funnel debriefing procedure before exiting the survey.

Measures

Baseline Measures

First, participants completed several baseline measures. Participants were asked to indicate the extent to which they currently feel several emotions linked to social and

counterfactual comparisons (see: Park & Baek; 2018, Smith, 2000; e.g., sympathy, worry, optimism/hope, inspiration, depression/sadness, envy/jealousy, pride, schadenfreude, gratitude, relief, regret) on a 7-point Likert scale (*1 = not at all; 7 = extremely*) and indicated their general emotional valence (0 = *negative*; 100 = *positive*) and arousal (0 = *very calm*; 100 = *very aroused*). Participants also answered questions about their experience with puzzles and other cognitive type tasks ("How good are you at solving puzzles?", "How precise is your hand-eye coordination?") and their experience with computers and video games ("How often do you use a computer with a mouse (not a track pad)?", "How often do you play video games?") using 5-point scales (*1= very poor/ not at all accurate/never, 5 = very good/very accurate/very often*).

Post-Task Measures—After Time 1

Second, post-feedback, participants provided self-evaluations about their performance on the target task (e.g. "How would you rate your performance on the circle/mental rotation task?" and "How would you rate your general ability in tasks involving hand-eye coordination?"; r = .43, $p < .01$) and motivation for performing well on a future trial (e.g. "How likely is it that you will try harder during the second round of the circle/mental rotation task?" and "How motivated are you to do well on the next round of the circle/mental rotation task?"; $r = .71$, $p < .01$) using 5-point Likert scales (1 = *very poor/not likely/not motivated at all*; 5 = *very good/very likely/very motivated*). Participants also indicated how they felt at the current moment on an assortment of discrete emotions (e.g. sympathy, worry, optimism/hope, inspiration, depression/sadness, envy/jealousy, pride, schadenfreude, gratitude, relief, regret) on a 7-point Likert scale (1 = *not at all*; 7 = *extremely*) and indicated their general affective state on 100-point slider scales reflecting both valence (0 = *negative*; 100 = *positive*) and arousal (0 = *very calm*; 100 = *very aroused*).

Finally, participants were asked about perceived closeness of the comparison score to one's own (Social: "How close do you see your actual score on the circle/mental rotation task to the other person's score that you were informed about?"; Counterfactual: "How close do you see your actual score on the circle/mental rotation task to the score you imagined getting?") and attainability of the comparison score (Social: "When you complete the circle/mental rotation task again, what's the likelihood of receiving a score close to the other person's score you were informed about?; Counterfactual: "When you complete the circle/mental rotation task again, what's the likelihood of receiving a score close to the one you imagined getting?) using 5-point Likert scales (1 = *not at all/very unlikely*; 5 = *extremely/very likely*).

Post-Task Measures—After Time 2

Third, after completing the second round of the target task, participants answered questions pertaining to their own perception of improvement ("Do you think you did better on the second round compared to the first on the circle/mental rotation task?" *Yes or No* and "How much do you think you improved on the circle/mental rotation task?" 1 = *Performance was much worse the second time*; 5 = *Performance was much better the second time*"). Participants answered the same questions about their performance on the task (same measures as described above; self-evaluation: $r = .57$, $p < .01$; motivation for an additional future performance: $r = .80$, $p < .01$). Participants also indicated how they felt at the current moment on an assortment of emotions and indicated their general affect and arousal (same measures as described above).

Performance Measure in Experiment 3

During each trial of the MR Task, participants selected one of the two 2D rotated objects. If participants selected the correct object, they received a score of 1 for that trial; however, if the

participant selected the incorrect object, they received a score of 0. A total performance score was calculated for each participant for each round of the MR task. This resulted in two performance scores for Time 1 and Time 2.

Exploratory Individual Differences

Three exploratory individual difference measures were included that relate to general tendencies to make social comparisons and counterfactual comparisons in everyday life. The main reason for inclusion of these variables was to examine whether such individual differences might moderate the core patterns of interest described above. For example, people high in tendencies to make frequent social comparisons (counterfactual thinking) might show stronger reactions to the social comparison (counterfactual) feedback. Likewise, handedness strength (mixed or strong handed) has been shown to play a role counterfactual thought production and social comparison, such that mixed-handed individuals who have greater access to right-hemisphere based processes have been shown to generate more counterfactual thoughts (Jasper et al., 2008) and to weight information about other people more heavily in comparative judgments (Rose et al., 2012). These results suggest that mixed handers might be more sensitive to comparative information overall (regardless of source).

First, to assess individual differences in social comparison orientation (SCO), participants completed the Iowa-Netherlands Comparison Orientation Measure (INCOM; Gibbons & Buunk, 1999). Participants indicated their agreement with 11 statements on a 5-point Likert scale (*1 = strongly agree*; 5 = *strongly disagree*). Sample items included "I always pay a lot of attention to how I do things compared with how others do things" and "If I want to find out how well I've

done something, I compare what I have done with how others have done" (α = .92, *M* = 3.20, *SD* = .93). Higher scores represented higher SCO.

Second, to assess different aspects of counterfactual thinking, participants completed the Counterfactual Thinking for Negative Events Scale (CNTES; Rye et al., 2008). Participants viewed 16 statements and rated how often they experience each one on a 5-point Likert scale (*1 = never; 5 = very often*). The 16 items assessed four domains of counterfactual thought: non-referent downward (e.g., "I think about how much worse things could have been", α = .83, *M* = 3.00, *SD* = .86), other-referent upward (e.g., "If only another person had not been so selfish, this whole mess could have been avoided", α = .91, *M* = 2.55, *SD* =.96) , self-referent upward (e.g., "I think about how much better things would have been if I had acted differently, α = .83, *M* = 2.85, *SD* = .90), and non-referent upward (e.g., "I think about how much better things could have been", α = .83, *M* = 2.88, *SD* = .90*)*. For purposes of this study, we collapsed all items into an overall counterfactual thought (CFT) composite (α = .76), where higher scores represented a greater extent of making counterfactual thoughts.

Third, to determine handedness strength score, participants completed the Edinburgh Handedness Inventory (EHI; Oldfield, 1971; α = .95). Participants rated the frequency with which they perform each of ten common tasks, such as writing, drawing, or using a spoon, with their right or left hand on the following scale: always left, sometimes left, no preference, sometimes right or always right. Responses are then summed across the ten tasks and range from strongly left-handed (-100) to strongly right-handed (+100). Following the practice of past research, the median absolute score on the handedness inventory, which was 90, was used to define the cut-off point between mixed versus strong handed participants. A median split was

performed on the scores such that those scoring 90 (absolute value) and above were categorized as "strong-handed" (n = 126) while those scoring below 90 where categorized as "mixed-handed" (n = 134).

Results

Preliminary Analysis

To begin, two steps were taken to confirm that participants believed and attended to the feedback. First, participants who indicated they did not believe the feedback to be accurate and/or the target task to be believable when asked if anything in the study made them feel suspicious (total n= 44) were excluded from analysis, leaving the combined total of 260 participants described previously. Second, as a manipulation check, we examined whether (at the group level) participants could generally recall the comparative value. First, a t-test confirmed that participants in the upward social condition reported a greater score for the other participant (M = 83.63, SD = 6.76) compared to the downward social comparison condition (M = 63.79, SD = 5.38), $t(111) = 17.32$, $p < .001$, d =3.25. Likewise, for the counterfactual condition, participants in the upward CFT condition reported being asked to consider achieving higher scores (M = 83.00, SD = 4.84) compared to the downward CFT condition (M = 64.67, SD = 5.87, $t(89) = 16.34$, $p < .001$, d = 3.45. Thus, overall, it appears that our manipulation of comparison direction was successful overall.

Core Analyses Involving Impact of Comparison Direction and Source (H1 and H2)

Next, we examined mean differences in self-evaluations, motivations/behavioral intentions and emotion/affect, and performance across our core independent variables. Given the overlapping goals and methods across Experiments 2 and 3 and to maximize on statistical power,

analyses were conducted involving data from both studies simultaneously. Also, note that analyses involving self-evaluations, behavioral intentions, and emotion/affect excluded participants in the control condition as they did not receive feedback (n = 51; final N = 209) and analyses examining performance excluded those in the control condition in Experiment 3 as well as all participants in Experiment 2 (n = 154; final N = 106).

The means involving self-evaluation and behavioral intentions were submitted to separate 2 (experiment/task: PR or MR task) × 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) ANOVAs. For general affect and discrete emotions, means were submitted to separate 2 (experiment/task: PR or MR task) × 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) × 2 (Time: 1 or 2) mixed-model ANOVAs, with the last factor as a within-participants factor. In relation to the latter factor (time), recall that general affect and discrete emotions were reported prior to the first round of the target task (Baseline) and then again after participants received the comparative feedback (Time 1). Finally, means for performance were submitted to a 2 (comparison source: social or counterfactual) × 2 (comparison direction: upward or downward) × 2 (Time: 1 or 2) mixed-model ANOVAs, with the last factor as a within-participants factor. For the latter factor (time), performance was taken after Time 1 performance and again after Time 2 performance.

Self-Evaluations. First, in terms of the main effects, there was a statistically significant main effect of direction, $F(1, 201) = 6.62$, $p = .011$, $\eta_p^2 = .032$. Overall, the results replicated those of Experiment 1, where participants reported higher self-evaluations in the downward comparison condition (M = 3.62; SD = .61) than the upward comparison condition (M = 3.37; SD

= .76), d = 36. The main effect of source was not significant, $F(1, 201) = .257$, $p = .613$, $\eta_p^2 = .001$. Second, in terms of the interactions, there were no statistically significant two-way interactions (Fs < .332, ps > .595) nor a three-way interaction between source, direction, and task, $F(1, 201) = .637$, $p = .426$, $\eta_p^2 = .003$.

Behavioral Intentions. First, in terms of the main effects, the effects of comparison source, comparison direction, and task were not significant (Fs < .984, ps > .322). Second, there were also no significant interactions (Fs < .931, ps > .336).

General Affect. Unlike with the self-evaluation analyses, the main effects here are less critical than the interactions due to the inclusion of the time factor. First, in terms of the main effects, there was a statistically significant main effect of source, $F(1, 201) = 4.01$, $p = .031$, $\eta_p^2 = .023$, such that participants in the social condition (M = 72.93, SD = 23.34) reported more positive affect overall compared to those in the counterfactual condition (M = 66.08, SD = 22.09). The main effect of time was not significant, $F(1, 201) = 2.76$, $p = .098$, $\eta_p^2 = .014$, nor was the main effect of direction, $F(1, 201) = .749$, $p = .388$, $\eta_p^2 = .004$. Second, and more importantly, in terms of the interactions, there was a significant two-way interaction between time and comparison direction, $F(1, 201) = 10.42$, $p = .001$, $\eta_p^2 = .049$. Specifically, follow up tests indicate no significant differences between participants in the upward (M = 70.36, SD = 24.16) and downward (M = 70.42, SD = 21.58*)* conditions for baseline affect ratings, $t(207) = -.018$, $p = .986$, $d = .003$. However, post-feedback, participants in the upward condition reported directionally less positive affect (M = 65.95, SD = 24.84) compared to participants in the downward condition (M = 71.75, SD = 22.51), $t(207) = -1.76$, $p = .079$, $d = .24$. See Figure 2. There were no other significant two-way interactions (Fs < 2.51, ps > .115), three-way

interactions (Fs < .1.32, ps > .251), nor a statistically significant four-way interaction, $F(1, 201) = .572$, $p = .450$, $\eta_p^2 = .003$.

Discrete Emotions. Four discrete emotions (relief, regret, envy, pride) were selected for analysis based on their relevance to social and counterfactual comparative processes (both upward and downward) and were submitted separately for analysis.

Relief. First, in terms of main effects, the effects of comparison source, comparison direction, task and time were not significant (Fs < .972, ps > .325). Second, in terms of the interactions, there was a significant two-way interaction between time and comparison source, $F(1, 201) = 6.11$, $p = .014$, $\eta_p^2 = .030$. Specifically, the nature of this effect was such that participants in the counterfactual condition reported somewhat higher feelings of relief post-feedback ($t = -1.83$, $p = .07$, $d = .15$) whereas participants in the social condition reported somewhat lower feelings of relief post-feedback ($t = 1.66$, $p = .10$, $d = .14$; See Figure 3). However, the effects were small and not theoretically meaningful. There were no other significant two-way interactions (Fs < 3.77, ps > .054), three-way interactions (Fs < .024, ps > .665), nor a statistically significant four-way interaction, $F(1, 201) = 1.32$, $p = .253$, $\eta_p^2 = .006$.

Regret. First, in terms of main effects, the effects of comparison source, comparison direction, task and time were not significant (Fs < 1.59, ps > .209). Second, in terms of interactions, there was a significant two-way interaction between time and comparison source, $F(1, 201) = 8.75$, $p = .003$, $\eta_p^2 = .042$. Specifically, participants in the counterfactual condition reported somewhat higher feelings of regret post-feedback ($t = -1.42$, $p = .158$, $d = .11$) whereas participants in the social condition reported somewhat lower feelings of regret post-feedback ($t = 2.94$, $p = .004$, $d = .23$; See Figure 4).

A significant two-way interaction between task and comparison source was also found. Follow up tests indicate that, for participants that engaged in the MR Task, a significant difference was found, such that participants in the counterfactual condition (M = 2.00, SD = 1.25) reported less regret compared to participants in the social condition (M = 2.61, SD = 1.79), $t(104) = -2.03$, $p = .045$, $d = .40$. For participants that engaged in the PR task, no difference was found between ratings of regret for the counterfactual condition (M = 2.46, SD =1.59) and social condition (M = 2.09, SD = 1.47), $t(101) = 1.21$, $p = .23$, $d = .24$. However, this result was not theoretically meaningful without consideration of direction or time factors. There were no other significant two-way interactions found (Fs < 1.31, ps > .254).

In terms of the three-way interactions, there was a statistically significant three-way interaction between time, comparison direction, and task, $F(1, 201) = 4.61$, $p = .033$, $\eta_p^2 = .022$. Overall, the nature of this interaction was such that the significant time × comparison direction interaction noted above appeared to be stronger when considering the PR task. Indeed, when looking at the PR task, there was a significant time and comparison direction interaction, $F(1, 99) = 6.07$, $p = .015$, $\eta_p^2 = .058$. Follow-up tests indicated that a significant difference was found in baseline ratings such that participants in the upward condition ($M = 1.93$, $SD = 1.50$) reported less regret compared to participants in the downward condition ($M = 2.75$, $SD = 1.95$), $t(101) = -2.14$, $p = .018$, $d = .47$. However, there was not a significant difference in regret ratings post-feedback between participants in the upward condition ($M = 2.22$, $SD = 1.77$) and participants in the downward condition ($M = 2.25$, $SD = 1.66$), $t(101) = -.094$, $p = .926$, $d = .02$. Second, when looking at the MR task, there was not a significant time × comparison interaction found, $F(1, 102) = .079$, $p = .779$, $\eta_p^2 = .001$. Finally, there were no other significant three-way interactions,

(Fs < .643, ps > .424) nor a statistically significant four-way interaction, $F(1, 201) = .149$, $p = .700$, $\eta_p^2 = .001$.

Envy. First, in terms of main effects, the effects of comparison source, comparison direction, task and time were not significant (Fs < 2.85, ps > .093). Second, in terms of interactions, there were no significant two-way interactions (Fs < 2.86, ps > .05), three-way interactions, (Fs < .973, ps > .235), nor a four-way interaction, $F(1, 201) = .298$, $p = .586$, $\eta_p^2 = .001$.

Pride. First, in terms of the main effects, there was a significant main effect of source, $F(1, 201) = 6.44$, $p = .012$, $\eta_p^2 = .031$, such that participants in the counterfactual condition ($M = 3.01$, $SD = 1.86$) reported less pride compared to those in the social condition ($M = 3.69$, $SD = 1.92$). The main effect of time was not significant, $F(1, 201) = 2.46$, $p = .119$, $\eta_p^2 = .012$, nor was the main effect of direction, $F(1, 201) = .798$, $p = .373$, $\eta_p^2 = .004$. Second, there were no interactions (Fs < .2.37, ps > .125).

Performance Analyses. First, in terms of the main effects, there was a significant main effect of time, $F(1, 102) = 21.04$, $p < .001$, $\eta_p^2 = .171$, such that participants did significantly better on the MR task during the second round ($M = 6.67$, $SD = 1.74$) compared to the first round ($M = 5.75$, $SD = 1.67$), indicating a practice effect. The main effect of source was not significant, $F(1, 201) = 2.46$, $p = .119$, $\eta_p^2 = .012$, nor was the main effect of direction, $F(1, 201) = .798$, $p = .373$, $\eta_p^2 = .004$. Second, there were no significant interactions (Fs < 2.45, ps > .121).

Control Conditions for the Core Measures

While the measures from the control condition were not included in the primary analyses due to the fact that we used a floating control condition that could not be folded into the factorial

ANOVA, it is worth noting where the means fell in this condition relative to the other conditions. Descriptive statistics for all conditions appear in Table 4. In general, looking at the trends for self-evaluations, behavioral intentions, and general affect, it appears that the scores from participants in the downward social condition were nominally greater than those in the control condition, while those in the and upward social condition and both counterfactual conditions were nominally lower than those in the control condition (i.e., the control condition tended to fall in-between). However, when looking at discrete emotion change scores, the results were more mixed. For instance, for relief, participants in both social conditions reported a decrease in relief, whereas participants in both counterfactual conditions reported an increase in relief, and participants in the control condition reported no change. For regret, participants in the upward counterfactual condition reported an increase in feelings of regret after receiving comparison feedback, whereas participants in the downward counterfactual, both social conditions, and the control conditions reported decreases in feelings of regret (with the control condition reporting the nominally greatest decrease). For envy, participants in the upward social condition reported a slight increase in envy post-feedback, whereas participants in all other conditions reported decreased envy. Lastly, for pride, participants in all conditions reported a decrease in pride after receiving the comparative feedback.

Mediation Analyses Involving Affect and Discrete Emotion (H3)

Recall that one of our core hypotheses involved the mediating impact of affect in accounting for the relationship between our manipulations and task performances. However, given that there were no significant main or interactive effects involving our critical independent variables on task performance, we could not test for mediation.

Exploratory Analyses Involving Moderating Roles of Individual Difference Variables

To assess our exploratory hypotheses involving individual differences, a series of moderation analyses using PROCESS v. 3.3 model 3 (Hayes, 2018) were conducted, where comparison direction, comparison source, and the relevant individual difference variables (SCO, CFT, and handedness) were entered (after centering) as predictors of the core variables noted above (self-evaluations, behavioral intentions, general emotions, discrete emotions, and performance). Notably, because of the limitations of PROCESS, our main dependent variables for general affect, discrete emotion, and performance involved a difference score. The main effects and all two-way and three-way interactions were tested for. PROCESS (Hayes, 2018) was also used to probe significant interactions (plotted at standard values of -1 SD, mean, +1 SD of the moderator variable of interest).

Self-Evaluations. First, the overall model for SCO was not significant, $F(7, 201) = 1.45$, $p = .19$, $R^2 = .05$. More critically for our purposes, all interaction effects involving SCO were not significant (ts < |1.33|, ps >.18). Second, the overall model for CFT was not significant, $F(7, 201) = 1.35$, $p = .23$, $R^2 = .05$. Critically, all interaction effects involving CFT were not significant (ts < |1.25|, ps >.21). Third, the overall model for handedness was significant, $F(7, 201) = 2.28$, $p = .03$, $R^2 = .07$. Notably, there was a significant three-way interaction between comparison source, comparison direction, and handedness, $b = .77$, $t(205) = 2.03$, $p = .04$. The results suggest that strong handers were sensitive to social comparisons ($b = .56$, $t = 3.16$, $p = .0018$, LCI=.21, UCI=.91), but not counterfactual comparisons ($b = .13$, $t = .67$, $p = .5017$, LCI=-.25, UCI=.52), such that they reported higher self-evaluations after receiving downward social comparison feedback compared to upward social comparison feedback. Mixed handers were somewhat, but not significantly, more sensitive to counterfactual comparisons ($b = .27$, $t = 1.41$, $p = .1612$,

LCI=-.11, UCI=.64), but not social comparisons (b = -.08, t = -.41, p = .6836, LCI=-.48, UCI=.31), such that higher self-evaluations were reported after receiving downward counterfactual comparison feedback compared to upward counterfactual comparison feedback. All other interaction effects involving handedness were not significant (ts < |1.61|, ps > .11).

Behavioral Intentions. First, the overall model for SCO was significant, $F(7, 201)$ = 2.48, p = .02, R^2 =.08. Critically, all interactions involving SCO were not significant (ts < |.71|, ps > .48). Second, the overall model for CFT was not significant, $F(7, 201)$ = 1.08, p = .38, R^2 =.04. More critically for our purposes, all interaction effects involving CFT were not significant (ts < |1.49|, ps > .14). Third, the overall model for handedness was not significant, $F(7, 201)$ = .71, p = .67, R^2 =.02. Critical for purposes of these analyses, all interaction effects involving handedness were not significant (ts < |1.14|, ps > .25).

General Affect. First, the overall model for SCO was significant, $F(7, 201)$ = 2.03, p = .05, R^2 =.07. Critically, all interaction effects involving SCO were not significant (ts < |1.46|, ps >.15). Second, the overall model for CFT was significant, $F(7, 201)$ = 2.10, p = .05, R^2 =.07. Critical for purposes of these analyses, all interaction effects involving CFT were not significant (ts < |1.72|, ps >.09). Third, the overall model for handedness was not significant, $F(7, 201)$ = 1.71, p = .11, R^2 =.06. More critically for our purposes, all interaction effects involving handedness were not significant (ts < |.74|, ps > .46).

Discrete Emotions. Again, four discrete emotions (relief, regret, envy, pride) were selected and submitted separately for analysis.

Relief. First, the overall model of SCO was not significant, $F(7, 201)$ = 1.11, p = .36, R^2 =.04. Critical for purposes of these analyses, all interaction effects involving SCO were not significant (ts < |1.14|, ps >.26). Second, the overall model of CFT was not significant, $F(7, 201)$

= 1.55, p = .15, R^2 =.05. Critically, all interaction effects involving CFT were not significant (ts < |1.33|, ps >.19). Third, the overall model of handedness was not significant, F(7, 201) = 1.38, p = .22, R^2 =.05. More critically for our purposes, all interaction effects involving handedness were not significant (ts < |1.27|, ps > .20).

Regret. First, the overall model of SCO was significant, F(7, 201) = 2.78, p = .0087, R^2 =.09. Notably, a source × SCO interaction was found, b = .44, t(205) = 2.06, p = .04. Conditional effects revealed that the impact of comparison source was stronger for participants high in SCO (b = 1.00, t(205) = 3.95, p = .006) than those low in SCO (b = .16, t(205) = .52, p = .61). This result is consistent with exploratory hypotheses, as the effect is stronger for those that are more likely to engage in social comparison. All other interactions involving SCO were not significant (ts < |1.29|, ps > .20). Second, the overall model of CFT was significant, F(7, 201) = 2.08, p = .05, R^2 =.07, though, all critical interaction effects involving CFT were not significant (ts < |1.15|, ps >.25). Third, the overall model of handedness was significant, F(7, 201) = 2.44, p = .02, R^2 =.08, though all critical interaction effects involving handedness were not significant (ts < |1.30|, ps > .19).

Envy. First, the overall model for SCO was not significant, F(7, 201) = .85, p = .54, R^2 =.03. Critically, all interaction effects involving SCO were not significant (ts < |1.82|, ps > .07). Second, the overall model for CFT was not significant, F(7, 201) = 1.09, p = .37, R^2 =.04. More critically for our purposes, all interaction effects involving CFT were not significant (ts < |1.85|, ps > .07). Third, the overall model for handedness was not significant, F(7, 201) = .79, p = .60, R^2 =.03. Critical for purposes of these analyses, all interaction effects involving handedness were not significant (ts < |1.85|, ps > .07).

Pride. First, the overall model for SCO was not significant, $F(7, 201) = .1.42$, $p = .20$, $R^2 = .05$. However, there was a significant three-way interaction between comparison direction, comparison source, and SCO, $b = -1.11$, $t(205) = -2.78$, $p = .006$. These results suggest that participants low in SCO showed the most sensitivity to social comparison feedback ($b = .95$, $t = 2.40$, $p = .0175$, LCI=.17, UCI=1.73), but that SCO did not predict sensitivity to counterfactual feedback ($b = -.54$, $t = -1.29$, $p = .1969$, LCI=-1.37, UCI=.28), such that receiving upward social comparison feedback resulted in greater negative change in pride compared to downward social comparisons. Interestingly, participants high in SCO were not significantly affected by the social comparison feedback, ($b = -.38$, $t = -.99$, $p = .3242$, LCI=-1.15, UCI=.38), nor the counterfactual feedback, ($b = .44$, $t = 1.19$, $p = .2356$, LCI=-.29, UCI=1.16). This result contradicts what would be expected theoretically, suggesting that participants low in SCO were more impacted by the comparison information. All other interactions involving SCO were not significant (ts < |.77|, ps > .44). Second, the overall model for CFT was not significant, $F(7, 201) = .65$, $p = .71$, $R^2 = .02$. More critically for our purposes, all interaction effects involving CFT were not significant (ts < |1.39|, ps > .17). Third, the overall model for handedness was not significant, $F(7, 201) = .60$, $p = .76$, $R^2 = .04$. Critical for these analyses, all interaction effects involving handedness were not significant (ts < |1.14|, ps > .25).

Performance. First, the overall model for SCO was not significant, $F(7, 98) = .68$, $p = .69$, $R^2 = .05$. Critically, all interaction effects involving SCO were not significant (ts < |.91|, ps > .37). Second, the overall model for CFT was not significant, $F(7, 98) = .89$, $p = .52$, $R^2 = .06$. More critically for our purposes, all interaction effects involving CFT were not significant (ts < |1.30|, ps > .19). Third, the overall model for handedness was not significant, $F(7, 98) = 1.07$, $p =$

.39, R^2 =.07. Critical for purposes of these analyses, all interaction effects involving handedness were not significant (ts < |1.54|, ps > .13).

Chapter Three

Discussion

A wealth of prior research has examined the influence of social and counterfactual comparisons on a variety of outcomes (e.g., self-evaluations, emotions, behavior). However, few studies have attempted to directly compare the direction and magnitude of influence across these two distinct comparison sources. The current research was among the first to directly compare social and counterfactual comparative processes and outcomes and did so in performance-relevant contexts. Across 3 experiments using scenario-based and performance-based methodologies, we tested several core hypotheses. First, we tested the hypothesis that upward comparisons (regardless of source) would produce more negative affect and poorer self-evaluations but more adaptive behavioral intentions and motivations and improved future performances (H1). Second, we tested the hypothesis that there would be an interaction between comparison source and direction, such that the effect of direction was expected to be stronger for affect and self-evaluations when making social comparisons, but stronger for motivations and actual performances when making counterfactual comparisons (H2). Finally, we tested a mediational hypothesis that changes in affect would statistically account for the impact of comparison direction (and interaction with source) on task performance (H3; specific to Experiment 3).

First, we found partial support for Hypothesis 1, and whether we achieved supportive results depended upon the study and dependent variable. In all three experiments, we found that participants who were exposed to upward comparison information reported lower self-evaluations compared to participants exposed to downward comparison information (regardless of source). We also found that participants who received upward comparative feedback reported

greater negative affect compared to participants who received downward comparative feedback in Experiments 2 and 3. These results involving self-evaluations and affect support H1 and are consistent with prior research showing that upward comparisons tend to produce more negative self-evaluations (Collins, 1996; Markman et al., 1993; Marsh & Gergen, 1970) and affective reactions (Roese, 1997; Roese & Morrison, 2009; Wheeler & Miyake, 1992). However, the results regarding the other variables (behavioral intentions/motivations and performance) were less consistent. For instance, while we found that participants who read upward comparison scenarios reported greater behavioral intentions for future improvement compared to participants who read downward comparison scenarios in Experiment 1, no differences were found in Experiments 2 and 3. Moreover, no evidence was found to support that comparative feedback impact future performances (Experiment 3).

Second, we found partial support for Hypothesis 2, albeit only for one specific measure in one experiment. Specifically, in Experiment 1, there was a significant interaction between comparison direction and source, such that the effect of comparison direction described above for self-evaluations was stronger when exposed to social comparison information than counterfactual comparison information. This result supports H2 and prior research revealing that social comparisons often have a stronger impact relative to other comparative sources (Müller-Kalthoff et al., 2017; Strickhouser & Zell, 2015), but extends this to counterfactual sources. However, as noted above this result pattern only emerged in Experiment 1 for self-evaluations. Results from other dependent measures (e.g., behavioral intentions/motivations, affect/emotions, and performance) and self-evaluations from Experiments 2 and 3 did not reveal the predicted pattern of results.

Third and finally, we did not find support for Hypothesis 3. As indicated above, there were no significant main or interactive effects of our main independent variables on performance. This result rendered an examination of mediation unjustifiable and unnecessary. While previous research has shown that comparisons often evoke an emotional response that can be used as a guide for behavioral change (Epstude & Roese, 2008; van de Ven et al., 2011) and that the negative emotions felt after engaging in upward comparisons can result in an increase in motivation (Park & Park, 2017; van de Ven, 2017), it was clearly not the case here that emotions/affect led to changes in behavior/performance.

In addition to examining our core hypotheses (outlined above), we also tested several exploratory hypotheses involving the moderating impact of several theoretically-relevant individual differences, including Social Comparison Orientation (SCO), Counterfactual Thought and Handedness. Overall, the vast majority of results did not reveal a moderating (interacting) influence of these individual difference variables on our core analyses, with a few notable exceptions. For instance, we found that self-evaluations of strong handers were more sensitive to social comparison information than counterfactual information (whereas mixed handers were equally influenced by both types). This finding is inconsistent with previous research which suggests mixed handers are more attentive to social information in the comparative judgment process (Rose et al., 2012), although notably the design and goals of these prior studies and the current ones are quite different. Additionally, when examining discrete emotions, significant interactions were found between SCO and regret and pride. With regret, the results indicated that impact of comparison information was stronger for participants high in SCO than those low in SCO. This finding was expected and makes sense theoretically, given that people who are more

likely to engage in social comparisons are more affected by comparative feedback. With pride, the results suggested that participants low in SCO were more sensitive to social comparison feedback whereas participants high in SCO were not affected; this result was the opposite of what would be expected theoretically.

In summary, the results across the three experiments revealed both consistencies and inconsistences with one another and with the broader literature. For instance, the results consistently revealed that upward comparisons (regardless of source) resulted in lower self-evaluations and greater negative affect than downward comparisons. However, the impact of comparison direction (and its interaction with source) across discrete emotions, behavioral intentions/motivations, and performance was less clear and either produced inconsistent results across experiments and across measures or did not reveal any impact at all. The inconsistencies and/or lack of effects along these lines could reflect that comparison source and direction lack a theoretical connection with these outcome variables. However, we suspect this is not the case due to the ample prior research and theory on the impact of both social and counterfactual information and due to the fact that self-evaluations and negative affect were influenced in the current studies. Rather, as we discuss in more detail below, we suspect that the lack of effects on the other variables might have more to do with the methodologies chosen and the specific tasks used.

Limitations and Future Directions

There were several limitations of note in the current research. First, although there were some consistencies across experiments, there were also some notable inconsistencies. These inconsistencies could have been due to a mismatch across experiments in terms of the

methodologies (e.g., scenario-based vs. performance-based), broader contextual backdrop (e.g., GPA/exam vs. online cognitive tasks), familiarity of feedback (e.g., exam scores vs. percentage correct/on target for online task), or participant characteristics (e.g., college students vs. MTurk participants). For instance, the use of a scenario in Experiment 1 created a hypothetical setting for the participant, whereas in Experiments 2 and 3 participants actually engaged in the tasks. Research has shown that providing real versus hypothetical information to participants can impact the decision-making process and result in different outcomes (Vlaev, 2012; Xu et al., 2019). Another methodological difference may have been the familiarity of the contexts and feedback. Specifically, the use of exam and GPA scenarios in Experiment 1 were likely more familiar and known than the use of novel, online cognitive performance tasks in Experiments 2 and 3.

Second, as noted above, the performance-based tasks used for Experiments 2 and 3 involved relatively unfamiliar, online cognitive tasks of which participants likely had minimal experience or interest. Given that the tasks were relatively unfamiliar, unimportant, and unengaging, participants might not have felt motivated to improve performance nor to pay much attention to the comparison feedback. Indeed, previous research has shown that the motivation, as well as the interaction between comparison direction and affective response, can be meaningfully influenced by the type of task in which participants engage (Hirt et al., 1996; Kruglanski & Mayseless, 1990; Markman & McMullen, 2003). Moreover, although research has shown using online samples produces data just as reliable as data collected from lab-based studies (Buhrmester et al., 2011; Hauser & Schwarz, 2015), it is possible that having an experimenter monitor participants as they engage in tasks would force them to take the tasks, and

subsequent feedback, more seriously. Thus, future research that examines our research questions using in-person studies and more important, engaging, and familiar performance-based task is needed.

Third, it is inherently challenging to design studies that simultaneously examine social and counterfactual paradigms. For instance, one difficulty is selecting a paradigm that can sensibly and easily incorporate both social and counterfactual feedback/information. Another issue is that it is often difficult to provide counterfactual feedback without also including other comparison information, particularly in a performance context. For example, prior research has often blended social and counterfactual comparisons together, such as research on comparisons in athletic performances (Kahneman & Varey, 1990; Medvec et al., 1995). Moreover, when contemplating alternate outcomes in real life, people often consider situations and information that involve other people and/or contain other comparison elements. While it can be challenging to design studies that cleanly examine these issues, future research should continue to explore the similarities and differences between comparison types in order to fully understand their respective impact on goal-directed behavior, emotions, and self-evaluations.

Fourth, this research was conducted only in restrictive performance-relevant contexts. Future research should extend this work to broader performance (and non-performance) contexts, notably into more ecologically-relevant settings (e.g., school/work, health/athletics). Given that comparative information is often provided across a variety of settings and has been shown to both enhance and inhibit performance (Blanton et al., 1999; Petrocelli et al., 2012), gaining a deeper understanding of the possible implications of these comparison types is important. Furthermore, when faced with undesirable information, students often consider counterfactual

outcomes which can add or hinder academic improvement (Petrocelli et al., 2012; Roese, 1997). Additionally, considering comparisons between employees can have implications on behavior (Vidyarthi et al., 2010) and possibly affect performance reviews, further exploration of the consequences of such comparisons should be examined further.

Conclusions and Implications

The current research has both theoretical and applied implications. First, from a theoretical perspective, this research contributes to our knowledge of both the social comparison and counterfactual research fields. For instance, the current findings confirmed that upward comparisons, regardless of source, can have a negative impact on self-evaluations and affective state. Moreover, this research is also relevant to several comparison-based process models that describe how people use comparison information to inform judgment, decisions, and behaviors. For example, the REM model suggests two distinct modes of process arise during comparative thought: reflection and evaluation (Markmen & McMullen, 2003). First, the reflective mode involves a determination of whether the comparative information is relevant to the self; second, the evaluative mode uses the comparative information as a reference point to evaluate the self. Here, we provide evidence that participants reflected on the comparative feedback and experienced contrast effects, such as lower self-evaluations and greater negative affect, for both social and counterfactual upward targets. Previous research has examined social and counterfactual comparison sources separately and found the same contrast-like evaluations (McMullen & Markman, 2002; Mussweiler et al., 2004b; Smallman & Summerville, 2018), and by looking at both types within the same study, the current research builds upon previous findings. In sum, our research provides direct evidence that social and counterfactual

comparisons should generally be viewed as falling under the same model in terms of how such sources impact the comparison process and outcomes.

Second, from an applied perspective, this research is relevant to a variety of situations wherein people perform tasks (e.g., academic or work settings) and receive or consider comparison-based feedback. For example, in academic settings, previous research has shown that comparative information plays a role in the development of students' self-concepts in connection to their motivation and academic achievement (Schurtz et al., 2014; Wolff et al., 2018). Given that upward social comparisons often lead to lower self-evaluations and, in turn, lower self-concepts, students who receive information about more successful peers could experience negative repercussions, such as poorer self-esteem, lower self-evaluations, and loss of motivation. For example, in workplace settings, employees may feel they are not experiencing the same level of success as others, creating an upward social comparison which could lower performance evaluations and negative affective states (Greenberg et al., 2006; Sterling & Labianca, 2015) and a loss of motivation for future performance. Alternatively, educators may want to focus on the potential value of internal comparisons (e.g., temporal, counterfactual) rather than external comparisons (i.e., social comparisons), as they could offset the unpleasant consequences of the comparison process by being less threatening. However, evidence for these possibilities is only partially supported here and elsewhere, particularly with regard to the value of counterfactual over social comparison information. Thus, more research is needed to understand the potential value of comparison-based information for interventions in performance contexts. The current research can be viewed as an important initial step toward understanding the theoretical and applied implications of comparison sources.

Table 1
Pilot Pursuit Rotor Task Speed

PRT Speed	**Estimated Accuracy Percentage**		
***N* = 10**	Mean	Std Dev	Median
2 full circular rotations/17 seconds	85.2	11.36	87
2 full circular rotations/15 seconds (default speed)	74.6	16.17	77.5
2 full circular rotations/12 seconds	75.8	10.79	74.5
2 full circular rotations/10 seconds	62	17.27	60
2 full circular rotations/8 seconds	49.9	22.82	50

Table 2
Comparative Feedback Prompts (Pursuit Rotor Task)

Upward Counterfactual Prompt	For the circle task, you received an accuracy score of 73%. That is, 73% of the time you stayed on target and within the acceptable error range. Sometimes it is useful to think about outcomes that didn't happen. Take a moment to vividly imagine if you had performed better than you did on this task. For example, imagine that you scored 85% accuracy on the circle task rather than 73%. If you performed at that level, what kinds of strategies would you have used? What would you have been doing differently to get that score? Please take a few minutes to think and write about this.
Downward Counterfactual Prompt	For the circle task, you received an accuracy score of 73%. That is, 73% of the time you stayed on target and within the acceptable error range. Sometimes it is useful to think about outcomes that didn't happen. Take a moment to vividly imagine if you had performed worse than you did on this task. For example, imagine that you scored 62% accuracy on the circle task rather than 73%. If you performed at that level, what kinds of strategies would you have used? What would you have been doing differently to get that score? Please take a few minutes to think and write about this.
Upward Social Comparison Prompt	For the circle task, you received an accuracy score of 73%. That is, 73% of the time you stayed on target and within the acceptable error range. Sometimes it is useful to think about how your performance compares to others. The last student who completed this task on this computer received an accuracy score of 85%. Thus, the last participant performed <u>better</u> than you. What kinds of strategies do you think this other student used? What might they have done differently to get that score? Please take a few minutes to think and write about this.
Downward Social Comparison Prompt	For the circle task, you received an accuracy score of 73%. That is, 73% of the time you stayed on target and within the acceptable error range. Sometimes it is useful to think about how your performance compares to others. The last student who completed this task on this computer received an accuracy score of 62%. Thus, the last participant performed <u>worse</u> than you. What kinds of strategies do you think this other student used? What might they have done differently to get that score? Please take a few minutes to think and write about this.
Control Prompt	For the circle task, you received an accuracy score of 73%. That is, 73% of the time you stayed on target and within the acceptable error range.

Table 3
Comparative Prompts (Mental Rotation Task)

Upward Counterfactual Prompt	For the mental rotation task, you received an accuracy score of 73%. That is, 73% of the time you chose the correct rotated image. Sometimes it is useful to think about outcomes that didn't happen. Take a moment to vividly imagine if you had performed better than you did on this task. For example, imagine that you scored 85% accuracy on the mental rotation task rather than 73%. If you performed at that level, what kinds of strategies would you have used? What would you have been doing differently to get that score? Please take a few minutes to think and write about this.
Downward Counterfactual Prompt	For the mental rotation task, you received an accuracy score of 73%. That is, 73% of the time you chose the correct rotated image. Sometimes it is useful to think about outcomes that didn't happen. Take a moment to vividly imagine if you had performed worse than you did on this task. For example, imagine that you scored 62% accuracy on the mental rotation task rather than 73%. If you performed at that level, what kinds of strategies would you have used? What would you have been doing differently to get that score? Please take a few minutes to think and write about this.
Upward Social Comparison Prompt	For the mental rotation task, you received an accuracy score of 73%. That is, 73% of the time you chose the correct rotated image. Sometimes it is useful to think about how your performance compares to others. The last student who completed this task on this computer received an accuracy score of 85%. Thus, the last participant performed <u>better</u> than you. What kinds of strategies do you think this other student used? What might they have done differently to get that score? Please take a few minutes to think and write about this.
Downward Social Comparison Prompt	For the mental rotation task, you received an accuracy score of 73%. That is, 73% of the time you chose the correct rotated image. Sometimes it is useful to think about how your performance compares to others. The last student who completed this task on this computer received an accuracy score of 62%. Thus, the last participant performed <u>worse</u> than you. What kinds of strategies do you think this other student used? What might they have done differently to get that score? Please take a few minutes to think and write about this.
Control Prompt	For the mental rotation task, you received an accuracy score of 73%. That is, 73% of the time you chose the correct rotated image.

Table 4

Main Dependent Measures in Experiments 2 and 3

	Counterfactual Upward		Counterfactual Downward		Social Upward		Social Downward		Control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Self-Evaluations (Time 1)	3.42	.64	3.50	.71	3.41	.86	3.72	.54	3.61	.67
Self-Evaluations (Time 2)	3.33	.89	3.57	.92	3.65	.83	3.71	.74	3.70	.73
Behavioral Intentions (Time 1)	3.86	.90	3.68	.83	3.74	1.23	3.91	.92	3.83	1.00
Behavioral Intentions (Time 2)	3.32	1.31	3.26	1.25	3.41	1.29	3.54	1.06	3.59	1.28
General Affect (Baseline)	66.81	22.18	66.73	23.48	73.5	25.69	73.83	19.26	70.90	19.51
General Affect (Time 1)	63.30	22.21	67.63	24.44	68.56	27.13	75.55	20.06	71.20	18.46
Pride (Baseline)	2.96	1.95	3.18	2.01	3.70	2.05	3.83	1.99	3.31	1.88
Pride (After Feedback)	2.85	1.88	3.04	2.00	3.41	2.04	3.81	2.15	3.08	1.87
Relief (Baseline)	2.96	1.73	3.02	1.98	3.26	2.05	3.64	1.81	3.16	1.88
Relief (After Feedback)	3.28	1.89	3.27	1.85	2.94	2.09	3.40	1.98	3.16	1.96
Regret (Baseline)	2.00	1.52	2.27	1.58	2.35	1.86	2.77	1.93	2.41	1.84
Regret (After Feedback)	2.43	1.84	2.20	1.43	2.04	1.78	2.26	1.64	1.86	1.37
Envy (Baseline)	1.83	1.54	2.02	1.60	2.17	1.79	2.17	1.44	2.20	1.74
Envy (After Feedback)	1.77	1.60	1.82	1.45	2.26	1.85	1.77	1.24	1.73	1.37
MR Task Score (Time 1)*	5.80	1.47	6.04	1.58	5.81	1.59	5.37	1.98	5.70	1.46
MR Task Score (Time 2)*	6.52	1.78	7.04	1.68	6.30	1.73	6.81	1.80	6.09	1.98
MR Task Difference Score*	.72	1.84	1.00	2.13	.48	1.77	1.44	2.38	.39	1.75

Note. Self -evaluation and behavioral intention items were rated on 5-point scales, where higher numbers indicate higher ratings. General emotion was rated on a 100-point sliding scale, where 0 is negative and 100 is positive. All discrete emotions (pride, relief, regret, envy) were rated on 7-point scales, where 1 is not at all and 7 is extremely. The MR Task score is calculated by adding up the number of correct answers, with a perfect score equaling 10. The MR Task Difference score was calculated by subtracting the Time 1 MR Task score from the Time 2 MR Task score. * indicates items from Experiment 3 only.

Figure 1. Interaction Results for Self-Evaluations in Experiment 1

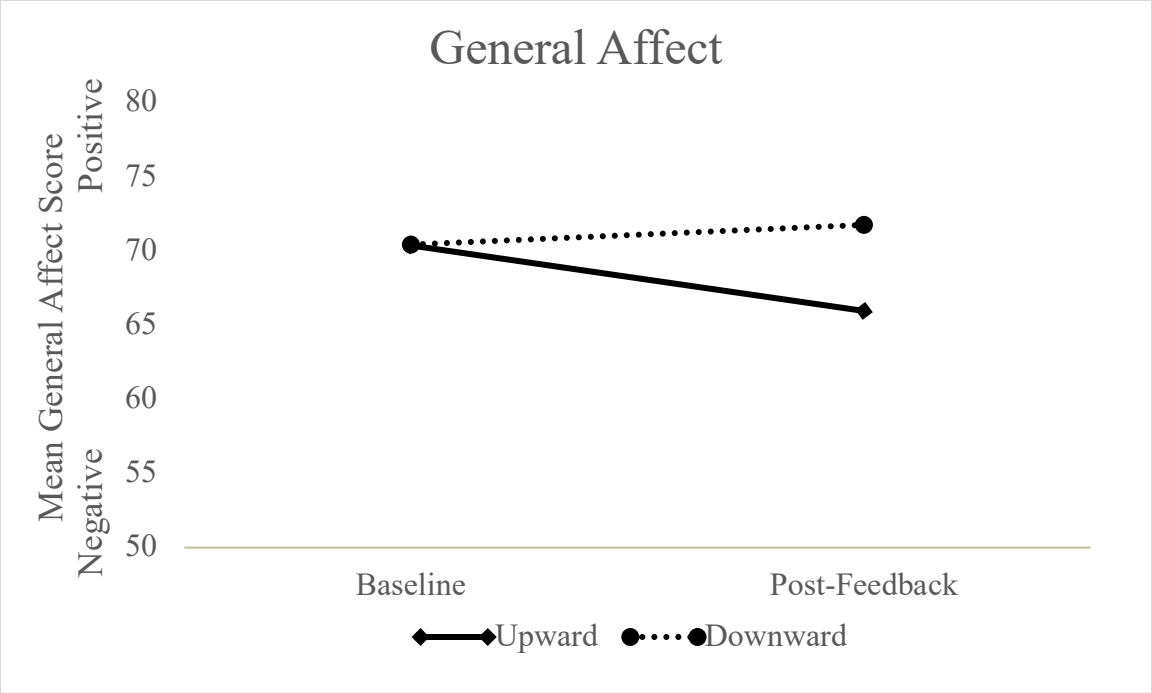

Figure 2. Interaction Results for General Affect in Experiments 2 and 3

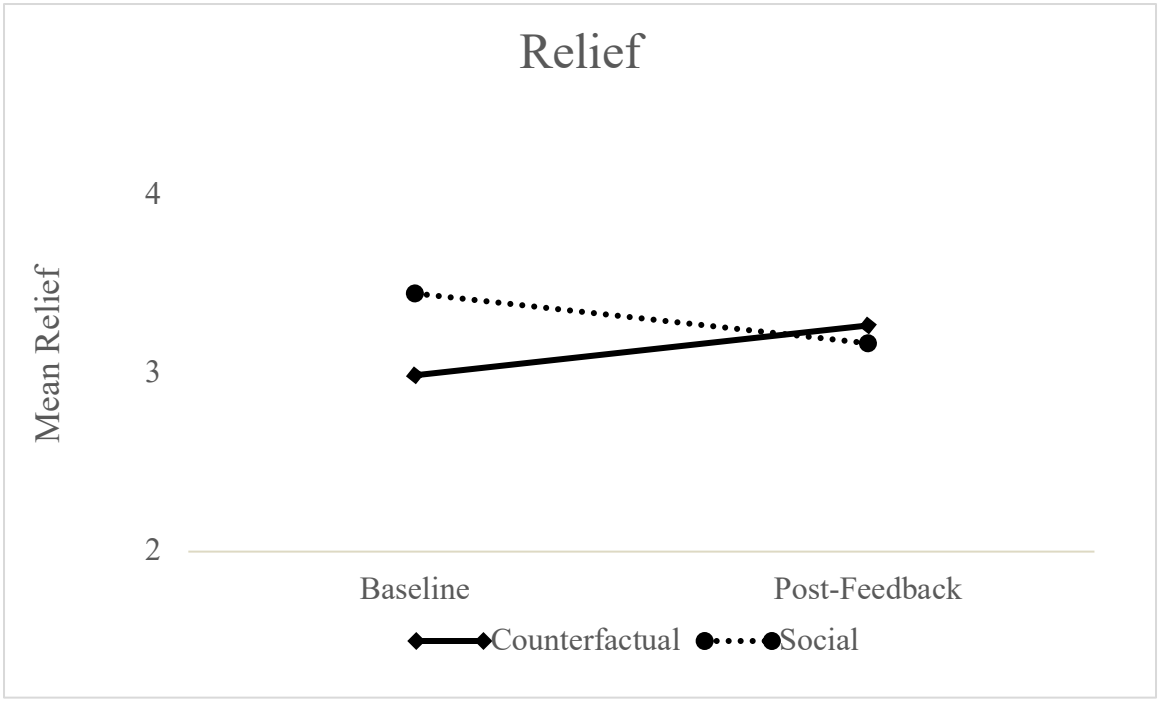

Figure 3. Interaction Results for Relief in Experiments 2 and 3

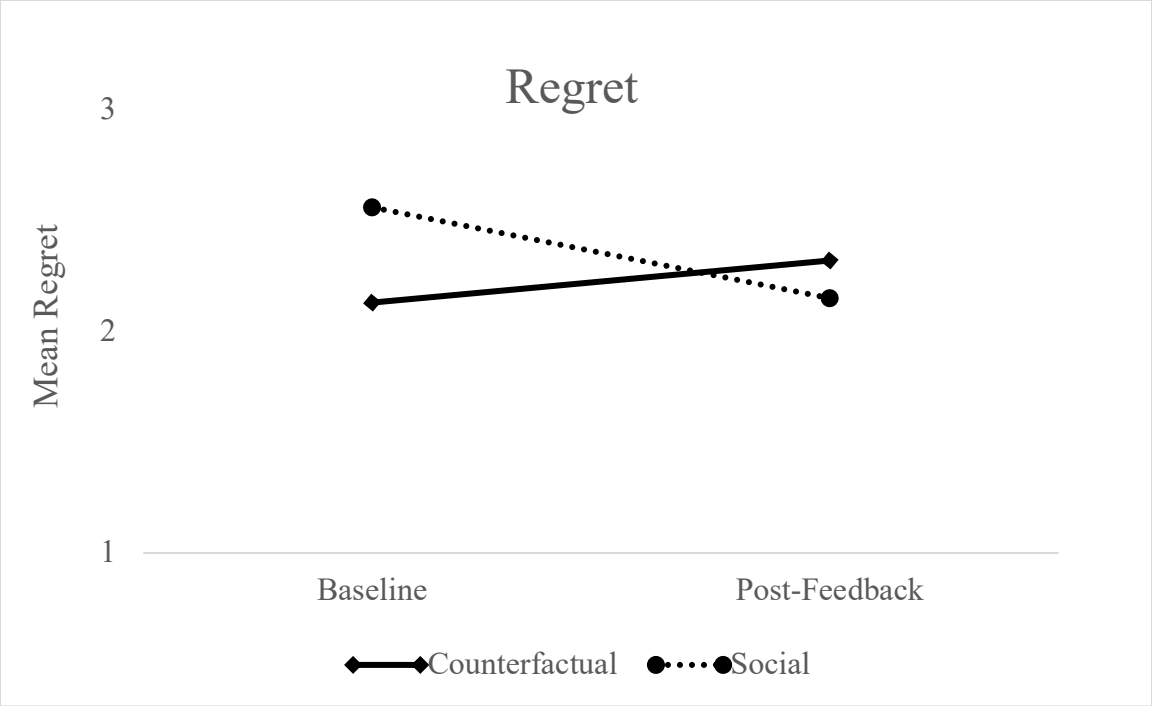

Figure 4. Interaction Results for Regret in Experiments 2 and 3

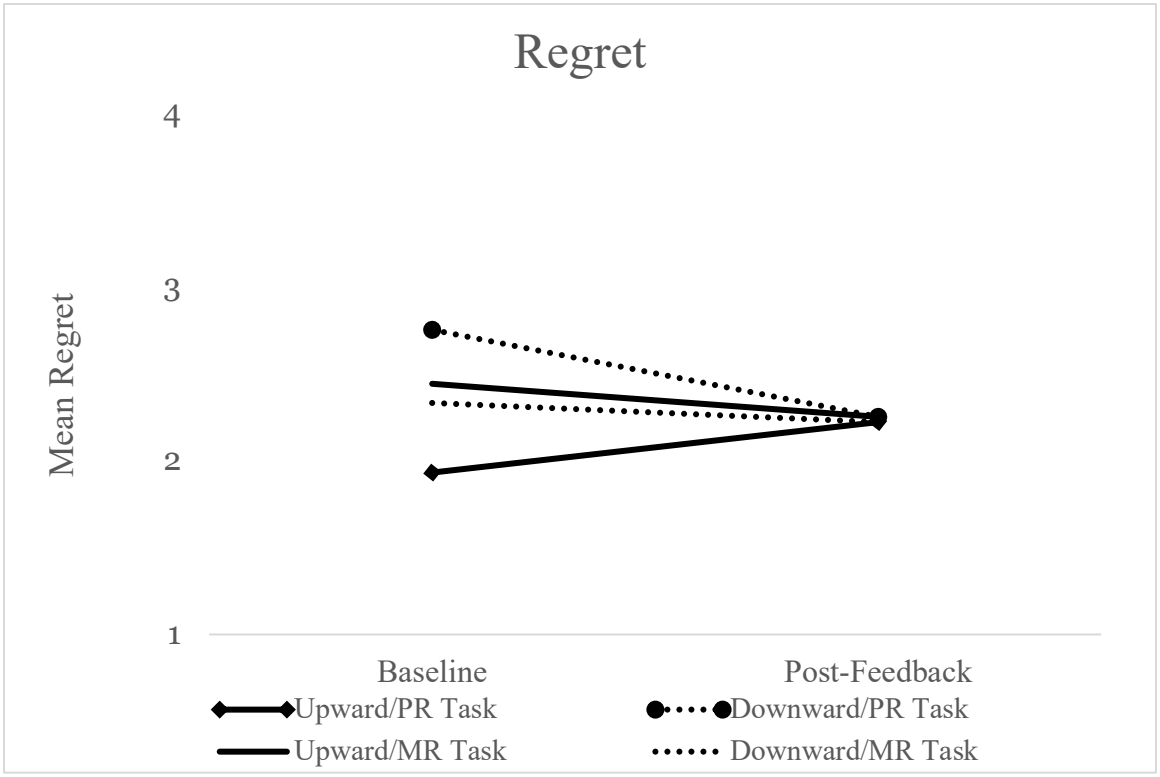

Figure 5. Interaction Results for Regret in Experiments 2 and 3

References

Alicke, M. D., & Zell, E. (2008). Social comparison and envy. In R. H. Smith (Ed.), *Series in affective science. Envy: Theory and research* (pp. 73-93). New York, NY, US: Oxford University Press.

Ammons, R. B., Alprin, S. I., & Ammons, C. H. (1955). Rotary pursuit performance as related to sex and age of pre-adult subjects. *Journal of Experimental Psychology*, *49*, 127.

Ammons, R. B., & Ammons, C. H. (1959). A Standard Anagram Task. *Psychological Reports*, *5*, 654–656.

Aspinwall, L. G. (1997). Future-oriented aspects of social comparisons. In B. P. Buunk & F. X. Gibbons (Eds*.), Health, coping, and well- being: Perspectives from social comparison theory* (pp. 125–166). Mahwah, NJ: Lawrence Erlbaum Associates.

Aspinwall, L. G. & Taylor, S. E. (1993). Effects of social comparison direction, threat, and self-esteem on affect, self-evaluation, and expected success. *Journal of Personality and Social Psychology, 64,* 708-722.

Blanton, H., Gibbons, F. X., Buunk, B. P., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology, 76*, 420-430.

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgement versus real-life behavior in trolley-style moral dilemmas. *Psychological Science, 29*, 1084-1093.

Buckingham J. T. & Alicke, M. D. (2002). The influence of individual versus aggregate social comparison and the presence of others on self-evaluations. *Journal of Personality and Social Psychology, 83*, 1117-1130.

Bruchmann, K. (2017). Compared to what? The importance of control groups in social comparison research. *Basic and Applied Social Psychology, 39*, 91-100.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3-5.

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

Chandler, J., Rozenweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods, 51*, 2022-2038.

Collins, R. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychology Bulletin, 199*, 69-69.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66,* 183-201.

Edmonds, K. A., Beeler, C. N., Rose, J. P. & Zell, E. (2020, February). *Expanding our understanding of the internal/external frame of reference model: dimensional (and social) comparisons in a physical and mental fitness context.* Poster accepted at the annual meeting of the Society for Personality and Social Psychology, New Orleans, LA.

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review, 12,* 168-192.

Faul, F. Erdfelder, E. Lang, AG & Buchner, A. (2007) G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39* 175-191.

Goodman, N. (1947). The problem of counterfactual conditions. *Journals of Philosophy, 44*, 113-128.

Gibbons, F. X. (1986). Social comparison and depression: Company's effect on misery. *Journal of Personality and Social Psychology, 51,* 140-148.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology, 76,* 129-142.

Gilovich, T. & Medvec, V. H. (1995). The experience of regret: What, when and why. *Psychological Review, 102*, 379-395.

Greenberg, J., Ashton-James, C. E., & Ashkanasy, N. M. (2006). Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes, 102,* 22-41.

Hauser, D. J. & Schwarz, N. (2015). Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods, 48,* 400-407.

Guilford, J.P. (1967). *The nature of human intelligence.* New York, NY, US: McGraw-Hill.

Hayes, A. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York, NY: The Guilford Press.

Hirt, E. R., Melton, R. J., McDonald, H. E., & Harackiewics, J. M. (1996). Processing goals, task interest, and the mood-performance relationship: A mediational analysis. *Journal of Personality and Social Psychology, 71,* 245-261.

Huntsinger, J. R., Isbell, L. M., & Clore, G. L. (2014). The affective control of thought: Malleable, not fixed. *Psychological Review, 121,* 600-618.

Jasper, J.D., Barry, K., & Christman, S.D. (2008). Individual differences in counterfactual production. *Personality and Individual Differences, 45,* 488-492.

Johnson, C. S. & Stapel, D. A. (2010). Harnessing social comparisons: When and how upward comparisons influence goal pursuit. *Basic and Applied Social Psychology, 23,* 234-242.

Kahneman, D. & Miller, D. T. (1986). Norm theory: comparing reality to its alternatives. *Psychological Review, 93*, 136-153.

Kahneman, D. & Varey, C. A. (1990). Propensitites and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology, 59,* 1101-1110.

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

Kruglanski, A. W. & Mayseless, O. (1990). Classic and current social comparison research: Expanding the perspective. *Psychological Bulletin, 108,* 195-208.

Lee, S. Y. (2014). How do people compare themselves with others on social network sites: The case of Facebook. *Computers in Human Behavior, 32*, 253-260.

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 1 - 10.

Lockwood, P. & Kunda Z. (1999). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology 73*, 91-103.

Martin, R., Suls, J., & Wheeler, L. (2002). Ability evaluation by proxy: Role of maximal performance and related attributes in social comparison. *Journal of Personality and Social Psychology*, *82*, 781-791.

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology, 29*, 87–109.

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality and Social Psychology Review, 7*, 244-267.

Markman, K. D., McMullen, M. N. & Elizaga, R. A. (2008). Counterfactual thinking, persistence, and performance: A test of the reflection and evaluation model. *Journal of Experimental Social Psychology, 44*, 421-428.

Markman, K. D., McMullen, M. N., Elizaga, R. A., & Mizoguchi, N. (2006). Counterfactual thinking and regulatory fit. *Judgment and Decision Making, 1*, 98-107.

McMullen, M. N. & Markman, K.D. (2000). Downward counterfactuals and motivation: The wake-up call and the Pangloss effect. *Personality and Social Psychology Bulletin, 26*, 575-584.

McMullen, M. N. & Markman, K. D. (2002). Affective impact of close counterfactuals: Implications of possible futures for possible pasts. *Journal of Experimental Social Psychology, 38*, 64-70.

Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology, 69,* 603-610.

More, S. & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology. 16,* 148-156.

Mueller, S. (2010, April 24). The pursuit rotor task. [Blog post]. Retrieved from http://peblblog.blogspot.com/2010/04/pursuit-rotor-task.html.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods, 222*, 250-259.

Müller-Kalthoff, H., Helm, F., & Möller, J. (2017). The big three of comparative judgment: On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Social Psychology of Education: An International Journal, 20*, 849-873.

Mussweiler, T. (2001). 'Seek and ye shall find': Antecedents of assimilation and contrast in social comparison. *European Journal of Social Psychology, 31*, 499-509.

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

Mussweiler, T., Rüter, K., & Epstude, K. (2004a). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. *Journal of Experimental Social Psychology, 40*, 689–696.

Mussweiler, T., Rüter, K., & Epstude, K. (2004b). The ups and downs of social comparison: Mechanisms of assimilation and contrast, *Journal of Personality and Social Psychology, 87*, 832-844.

Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). "If only I weren't" versus "if only I hadn't": Distinguishing shame and guilt in counterfactual thinking. *Journal of Personality and Social Psychology, 67*, 585-595.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Olson, J. M., Buhrmann, O., & Roese, N. J. (2000). Comparing comparisons: An integrative perspective on social comparison and counterfactual thinking. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 379-398). New York: Kluwer Academic/Plenum.

Park S. Y. & Baek Y. M. (2018). Two faces of social comparison on Facebook: The interplay between social comparison orientation, emotions, and psychological well-being. *Computers in Human Behavior, 79*, 83-93.

Park Y. & Park S. W. (2017). Goal orientations and social comparison: The role of difference motivations in affiliation preferences. *Motivation and Emotion, 41,* 617-627.

Petrocelli, J. V., Seta, C. E., Seta, J. J., & Prince, L. B. (2012). “If only I could stop generating counterfactual thoughts”: When counterfactual thinking interferes with academic performance. *Journal of Experimental Social Psychology, 28,* 1117-1123.

Piper, B. J. (2011). Age, handedness, and sex contribute to fine motor behavior in children. *Journal of Neuroscience Methods, 195*, 88–91.

Rim, S. & Summerville, A. (2014). How far to the road not taken? The effect of psychological distance on counterfactual direction. *Psychology Bulletin, 40*, 391-401.

Roese, N.J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology, 66*, 805-818

Roese, N.J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*, 133-148.

Roese, N.J. & Olson, J.M. (1993). The structure of counterfactual thought. *Personality and Social Psychology Bulletin, 19*, 312-319.

Roese N. J. & Morrison, M. (2009). The psychology of counterfactual thinking. *Historical Social Research, 34,* 16-26.

Roese, N. J., Hur, T., & Pennington, G. L. (1999). Counterfactual thinking and regulatory focus: Implications for action versus inaction and sufficiency versus necessity. *Journal of Personality and Social Psychology, 77*, 1109-1120.

Rose, J. P., Jasper, J. D., & Corser, R. (2012). Interhemispheric interaction and egocentrism: The role of handedness in social comparative judgment. *British Journal of Social Psychology, 51*, 111-129.

Rye, M. S., Cahoon, M. B., Ali, R. S., & Daftary, T. (2008). Development and validation of the counterfactual thinking for negative events scale. *Journal of Personality Assessment, 90*, 261-269.

Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction, 34,* 32-41.

Smallman, R. & Summerville, A. (2018). Counterfactual thought in reasoning and performance. *Social and Personality Psychology Compass, 12.*

Smith, R. H. (2000). Assimilative and contrastive emotional reactions to upward and downward social comparisons. In J. Suls, & L. Wheeler (Eds.), *Handbook of social comparison* (pp. 173-200). New York: Plenum.

Smith, R. H. & Kim, S. H. (2007). Comprehending Envy. *Psychological Bulletin, 133*, 46-64.

Sterling C. M. & Labianca, G. (2015). Costly comparisons: Managing envy in the workplace. *Organizational Dynamics, 44,* 296-305.

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31.

Strickhouser, J. E., & Zell, E. (2015). Self-evaluative effects of dimensional and social comparison. *Journal of Experimental Social Psychology, 59,* 60-66.

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

Summerville, A. & Roese, N. J. (2008). Dare to compare: Fact-based versus simulation-based comparison in everyday life. *Journal of Experimental Social Psychology, 44*, 664-671.

Tangney, J. P. (1990). Assessing individual differences in proneness to shame and guilt: Development of the self-conscious affect and attribution inventory. *Journal of Personality and Social Psychology, 59,* 102-111.

Taylor, S. E., Wayment, H. A., & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (pp. 3–27). New York: Guilford Press.

Vlaev, I. (2012). How different are real and hypothetical decisions? Overestimation, contrast, and assimilation. *Journal of Economic Psychology, 33,* 963-972.

van de Ven, N. (2017). Envy and admiration: Emotion and motivation following upward social comparison. *Cognition and Emotion*, 31, 193-200.

van de Ven, N., Zeelenberg, M., & Pieters, R. (2011). Why envy outperforms admiration. *Personality and Social Psychology Bulletin, 27,* 784-795.

Vandenberg S. G., & Kuse, A. R. (1978). Mental Rotations, a Group Test of Three-Dimensional Spatial Visualization. *Perceptual and Motor Skills, 47*, 599-604.

Vidyarthi, P. R., Liden, R. C., Anand, S., Erdogan, B., & Ghosh, S. (2010). Where do I stand? Examining the effects of leader-member exchange social comparison on employee work behaviors. *Journal of Applied Psychology, 95,* 849-861.

Wehrens, M. J. P. W., Buunk, A. P., Lubbers, M. J., Dijkstra, P., Kuyper, H. & van der Werf, G. P. C. (2010). The relationship between affective response to social comparison and academic performance in high school, *Contemporary Educational, Psychology, 32*, 203-214.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548-573.

Wheeler, L., & Miyake, K. (1992). Social comparisons in everyday life. *Journal of Personality and Social Psychology, 62*, 760–773.

Williams, L. G. (1996). The effect of target specification on objects fixated during visual search. *Perception and Psychophysics, 1*, 315-318.

Wills, T.A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90,* 245–271.

Wills, T. A. (1987). Downward comparison as a coping mechanism. In C. R. Snyder & C. Ford (Eds.), (1991). *Coping with negative life events: Clinical and social-psychological perspectives.* New York: Plenum Press.

Wolff, F., Helm, F., Zimmermann, F., Nagy, G., & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology, 110,* 1005-1025.

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

Xu, S., Xiao, Z., & Roa, H. (2019). Hypothetical versus real monetary reward decreased the behavioral and affect effects in the balloon analogue risk task. *Experimental Psychology, 66,* 221-230.

Zell, E. & Alicke, M. D. (2009a). Self-evaluative effects of temporal and social comparison. *Journal of Experimental Social Psychology, 45,* 223-227.

Zell, E. & Alicke, M. D. (2009b). Contextual neglect, self-evaluation, and the Frog-Pond effect. *Journal of Personality and Social Psychology, 97,* 467-482.

Zell, E. & Alicke, M. D. (2010a). Comparisons over time: Temporal trajectories, social comparison, and self-evaluation. *European Journal of Social Psychology*, *40,* 375-382.

Zell, E. & Alicke, M. D. (2010b). The local dominance effect in self-evaluation: Evidence and explanations. *Personality and Social Psychology Review, 14*, 368-384.

Appendix A

Experiment 1 Scenarios

Exam Scenarios

Counterfactual

"Imagine you're a junior in college and it's approaching break. You just took your second exam in your Principles of Psychology class and wonder how you did. You log into your Blackboard account and discover you answered 21 out of 30 questions correctly. You then consider how things could have gone differently for you to a get a higher [lower] score."

Social

"Imagine you're a junior in college and it's approaching break. You just took your second exam in your Principles of Psychology class and wonder how you did. You log into your Blackboard account and discover you answered 21 out of 30 questions correctly. You then notice the average score for other students was 26 [16] out of 30 points."

GPA Scenarios

Counterfactual

"Imagine you're in high school and starting to apply to colleges. have a 3.2 GPA. You think about how you could have earned a higher [lower] GPA."

Social

"Imagine you're in high school and starting to apply to colleges. You have a 3.2 GPA and your friend has a 3.7 [2.7] GPA."

Appendix B

Experiment 1 Measures

Exam

If you were this person, how would you rate your performance on the Principles of Psychology Exam? (1 = *very poor*; 5 = *very good*)

If you were this person, how would you rate your general knowledge in psychology? (1 = *very poor*; 5 = *very good*)

If you were this person, how likely is it that you would try or study harder for your next Principles of Psychology exam? (1 = *not likely*; 5 = *very likely*)

If you were this person, how motivated would you be to do well on your next Principles of Psychology exam? (1 = *not at all motivated*; 5 = *very motivated*)

How similar do you see yourself to the college Junior in the Principles of Psychology class who just received their exam feedback? (1 = *not at all*; 5 = *extremely*)

GPA

If you were this person, how would you rate your GPA? (1 = *very poor*; 5 = *very good*)

If you were this person, how would you rate your general knowledge? (1 = *very poor*; 5 = *very good*)

If you were this person, how likely is it that you would try or study harder when taking classes next semester? (1 = *not likely*; 5 = *very likely*)

If you were this person, how motivated would you be to do well in your classes next semester? (1 = *not at all motivated*; 5 = *very motivated*)

How similar do you see yourself to the high school student preparing for college? (1 = *not at all*; 5 = *extremely*)

Appendix C

Pursuit Rotor Task

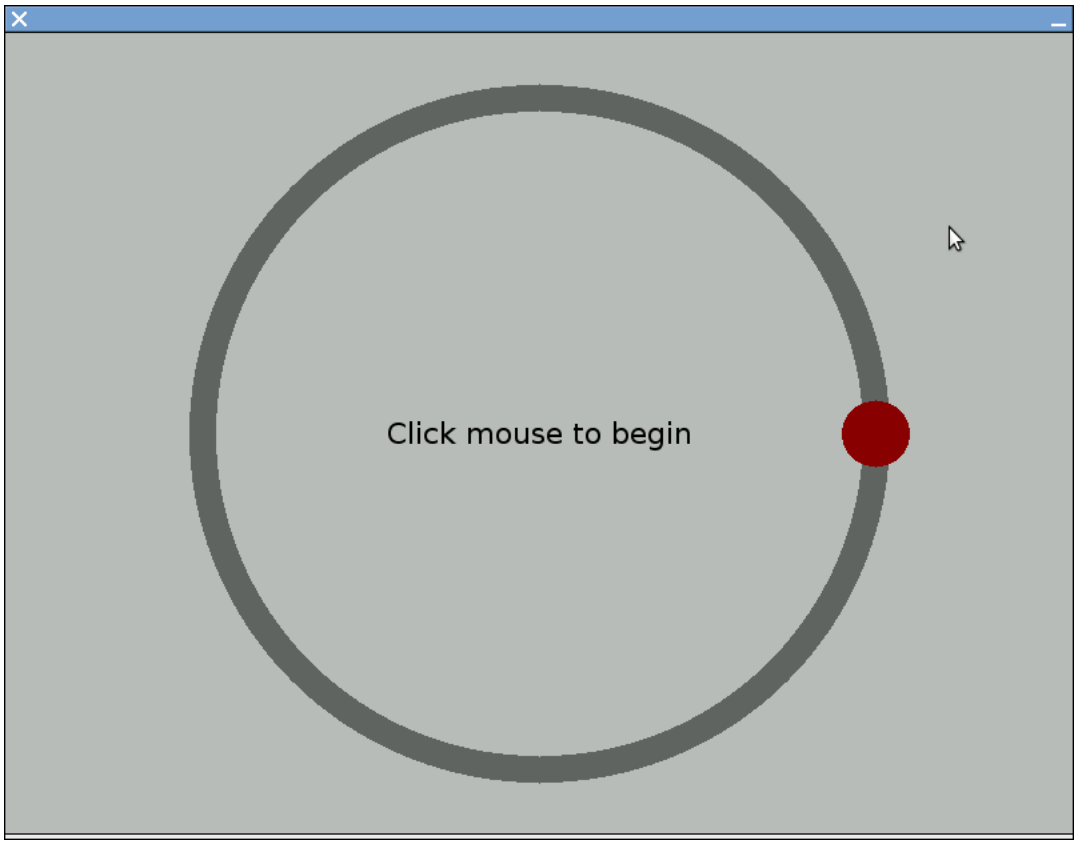

Figure C1. Image of Pursuit Rotor Task

Appendix D

Mental Rotation Task

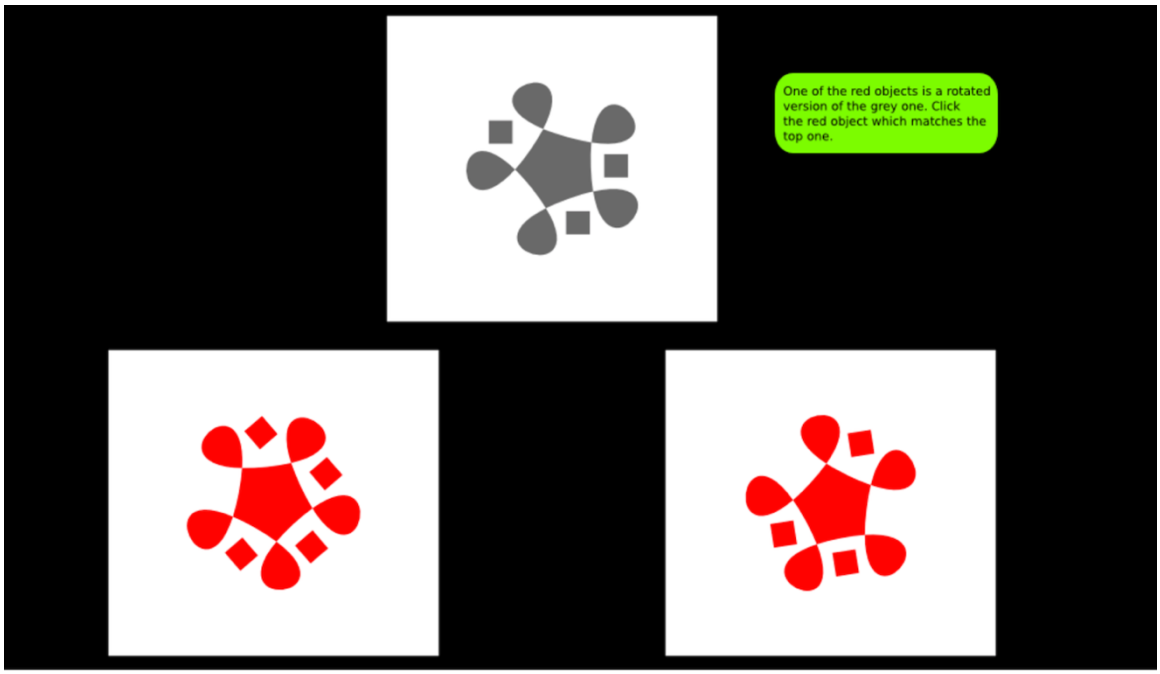

Figure D1. Image of Mental Rotation Task