This is the pre-peer reviewed version of the following article: Categorizing at the group-level in response to intragroup social comparisons: A self-categorization theory integration of self-evaluation and social identity motives, which has been published in final form at

http://onlinelibrary.wiley.com/doi/10.1002/ejsp.306/abstract.

Running Head: INTRAGROUP SOCIAL COMPARISONS

Categorizing at the Group-Level in Response to Intragroup Social Comparisons:

A Self-Categorization Theory Integration of Self-Evaluation and Social Identity Motives

Michael T. Schmitt   Nyla R. Branscombe

Simon Fraser University   University of Kansas

Paul J. Silvia   Donna M. Garcia

University of North Carolina at Greensboro   University of Kansas

Russell Spears

Cardiff University

Author Note

Abstract

Two experiments examined how people respond to upward social comparisons in terms of the extent to which they categorize the self and the source of comparison within the same social group. Self-evaluation maintenance theory (SEM) suggests that upward ingroup comparisons can lead to the rejection of a shared categorization, because shared categorization makes the comparison more meaningful and threatening. In contrast, social identity theory (SIT) suggests that upward ingroup comparisons can lead to the acceptance of shared categorization because a high-performing ingroup member enhances the ingroup identity. We attempted to resolve these differing predictions using self-categorization theory, arguing that SEM applies to contexts that make salient one's personal identity, and SIT applies to contexts that make collective identity salient. Consistent with this perspective, the level of identity activated in context moderated the effect of an upward ingroup comparison on the acceptance of shared social categorization.

Categorizing at the Group-Level in Response to Intragroup Social Comparisons:

A Self-Categorization Theory Integration of Self-Evaluation and Social Identity Motives

Festinger (1954) recognized that self-evaluation can rarely occur on the basis of objective

criteria, and by necessity rests heavily on comparisons between the self and others (Wood & Wilson,

2003). We explore predictions made by two theories, both of which were heavily influenced by

Festinger's ideas about social comparison: the self-evaluation maintenance model (SEM; Tesser,

1988) and social identity theory (SIT; Tajfel & Turner, 1986). We examine the consequences of

comparing the self with an ingroup member who outperforms the self. In particular, we are

interested in the implications of such upward ingroup comparisons for the degree to which

individuals self-categorize in a way that includes the self and the other person within the same social

category. We reconcile competing predictions made by SEM and SIT by using self-categorization

theory to argue that they apply to different levels of identity activated by the social context.

Upward social comparisons and shared category membership: SEM

Building on social comparison theory and the assumption that people are motivated to

perceive the self positively, Tesser's (1988) self-evaluation maintenance (SEM) model argues that

three factors interactively influence the outcomes of social comparison processes: the self-relevance

of the comparison dimension, the other person's performance on that dimension relative to one's

own, and the degree of "closeness" (either in terms of intimacy or similarity) with the other person.

When the performance dimension is relevant to the self, a comparison process occurs in which

one's own performance is evaluated relative to the performance of others. When individuals are

outperformed on self-relevant dimensions, self-evaluation is likely to suffer; whereas performing

better than others on self-relevant dimensions can enhance self-evaluation. In contrast, when the

dimension of comparison is irrelevant to one's personal self, a reflective process is activated in which one can bask in the reflected glory of another's good performance, or suffer in the shadow of another's failure. Building on Festinger's idea that similar others are more informative social comparisons, Tesser argued that the degree of closeness between the self and the other moderates both the comparison and reflection processes. Thus, one can buffer the effects of a painful social comparison or the reflection of another's failure by distancing from the identity-threatening other. Conversely, self-evaluation can be protected or enhanced by seeking closeness to others who perform similarly or worse than the self on self-relevant dimensions (Pleban & Tesser, 1981).

Following the self-evaluation maintenance model, Mussweiler, Gabriel, and Bodenhausen (2000) suggested that individuals can avoid the negative self-evaluative implications of a threatening upward comparison by distancing from social categorizations shared by both the self and the other person. Sharing a category membership creates psychological closeness between the self and other, and increases the meaningfulness of the comparison and its implications for self-evaluation. In three studies, Mussweiler et al. (2000) presented participants with an individual who either performed better or worse than the participants. Using both natural and minimal group memberships, they found that upward comparisons led participants to focus less on social identities they shared with the source of the comparison. Supporting the idea that such shifts in social categorization are self-protective, participants who scored high on a measure of self-esteem (and presumably have more experience or skill at using self-protective strategies) were more likely to exhibit these identity shifts compared to those low in self-esteem. Furthermore, in one of the studies, participants who engaged in these self-protective identity shifts expressed a more positive mood than those who lacked an opportunity to do so.

Upward social comparisons and shared category membership: SIT

Although most social comparison research has focused on the implications of social

comparisons for the individual self, social identity theory (Tajfel & Turner, 1986) is a notable exception. SIT argues that the value of a social group membership is determined by group-level social comparisons between the ingroup and relevant outgroups. SIT also assumes that people are motivated to perceive the ingroup positively, much like SEM assumes that people prefer to perceive the individual self positively.

However, because SIT focuses on group identity, the implications of comparisons with other individual ingroup members differ from the effects predicted by SEM. According to SEM, sharing a group membership with another person increases closeness and thus makes upward comparisons on self-relevant dimension more threatening compared to when group membership is not shared. From an SIT perspective, however, ingroup members who perform well can potentially make the ingroup positively distinct from outgroups—as a result, they are identity-enhancing instead of threatening. Because group identity is what is critical in SIT, whether an ingroup member outperforms or under-performs the individual self is mostly irrelevant; the ingroup member's implications for the group identity as a whole are important. Indeed, research has found that individual ingroup members who harm the ingroup identity are derogated (Branscombe, Wann, Noel, & Coleman, 1993; Marques, Yzerbyt, & Leyens, 1988), and that this derogation protects the ingroup's identity (Castano, Paladino, Coull, & Yzerbyt, 2002).

In contrast to SEM, social identity theory suggests that upward ingroup comparisons reflect positively on the ingroup identity, and are unlikely to lead to distancing from the ingroup categorization. In fact, relative to downward ingroup comparisons, upward ingroup comparisons are likely to lead to increased acceptance of the ingroup categorization. Previous research has confirmed that people do more readily identify with ingroups with relatively high status, at least when that high status is seen as legitimate (e.g., when it is based on performance; Ellemers, 1993).

Reconciling SEM and SIT: The importance of comparative context

Thus, although SEM and SIT share common assumptions and origins within social comparison theory, they make different predictions regarding how willing people will be to adopt a social categorization they share with an upward comparison. However, these apparent contradictions arise not because the theories are theoretically incompatible, but rather because their predictions apply to different comparative contexts (Mussweiler, 2003; Weber & Brewer, 1994; Schmitt, Silvia, & Branscombe, 2000). As postulated by self-categorization theory (SCT; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), the self can be defined at different categorical levels depending on the social context, and the consequences of social comparison with another ingroup member will depend on whether the context encourages self-construal at the individual or group level. For contexts that activate self-categorization at the individual level, SCT predicts shifts in the acceptance of a group categorization that reflect the implications of that categorization for personal identity, as specified by SEM predictions. SCT predicts the opposite in contexts that lead to self-categorization at the group level. In such contexts, SCT predicts that (everything else being equal) shifts in the acceptance of a group-level categorization will be determined by the extent to which the ingroup is positively distinct from other groups.

The few empirical investigations that have compared the consequences of social comparison in interpersonal and intergroup contexts have supported the contention that SEM applies to interpersonal contexts and SIT applies to intergroup contexts. Weber and Brewer (1994) manipulated level of self-categorization by varying relative group size. Consistent with SEM, majority group members' self-evaluations were contrasted with ingroup comparisons; consistent with SIT, minority group members assimilated their self-evaluations toward ingroup comparisons. Similarly, several studies demonstrate that upward comparisons lead to more positive self evaluations when the other person is construed as part of the self, compared to when the self and other were categorized as separate individuals (Gardner, Gabriel, & Hochschild, 2002;

Kemmelmeier & Oyserman, 2001). In order to examine the self-categorization explanation that these different consequences stem from different levels of self-categorization, Schmitt et al. (2000) directly manipulated whether ingroup social comparisons occurred in an interpersonal or intergroup context. As predicted by SCT, participants in the interpersonal context condition evaluated the upward ingroup comparison more negatively than the poorly performing ingroup member. Conversely, participants in the intergroup condition evaluated the upward ingroup comparison more positively than the poorly performing ingroup member. Thus, in both conditions, participants showed a preference for the ingroup member that had positive implications for the level of identity that was relevant in context.

Although previous research demonstrates that context moderates responses to social comparisons, no research to date has explored how comparative context moderates the extent to which individuals will adopt a self-categorization they share with the source of an upward comparison. Based on SCT's assumption that the identity implications of social comparison depend on the level of self-categorization encouraged by the social context, we argue that an upward social comparison will have opposite effects on the acceptance of group-level categorization in interpersonal and intergroup contexts. Therefore, we suggest that the predictions of SEM and Mussweiler et al.'s (2000) findings apply to contexts in which the interpersonal self is salient, and SIT's predictions apply to contexts in which group-level identity is salient.

<div align="center">Overview of the Present Experiments</div>

In two studies we examined how social comparisons affect the acceptance of a group-level categorization that is shared by both the self and the comparison source. In both studies, we manipulated whether the context was interpersonal or intergroup. Female participants received negative feedback on a performance task, and then were exposed to another woman who either outperformed them or performed similarly poorly. In other words, relative to their own

performance, participants faced either an upward or lateral comparison. Our choice to use a lateral

comparison rather than a downward comparison was methodological. We designed the performance

feedback manipulations such that they would invoke in participants a sense of failure and maximize

the potential negative consequences of an upward comparison for personal identity. Because

participants received highly negative performance feedback, there were limits on the extent to which

a downward comparison could plausibly be provided. Regardless, the inclusion of downward

comparisons is not necessary for testing our hypotheses regarding the relative consequences of

upward comparisons.

After exposing participants to the ingroup comparison source, we then measured the degree

to which participants accepted a group-level categorization shared by both the participant and the

other person—their gender group. We measured acceptance of group-level categorization in

different ways. In Experiment 1, participants completed a measure of gender-group identification.

Group identification, or an individual's emotional attachment to the ingroup (Tajfel & Turner,

1986), is an obvious indicator of the extent to which a person accepts a group-level categorization.

In addition, participants wrote open-ended self-descriptions that we later coded for gender-

prototypicality. We measured the prototypicality of self-descriptions because one outcome of the

acceptance of a group-level self-categorization is to define the self in less idiosyncratic terms, and

more in terms of the ingroup prototype (Turner et al., 1987). In Experiment 2, we used a more

specifically group-level indicator of categorization. Participants were given an opportunity to

differentiate between men and women on an open-ended measure. As predicted by self-

categorization theory (Turner et al., 1987), the acceptance of group-level categorization should be

reflected in greater intergroup differentiation.

Our central prediction was that the comparative context (interpersonal or intergroup) would

interact with the direction of the ingroup comparison (upwards or lateral) to affect participants'

acceptance of shared category membership. In other words, the effect of an upward ingroup comparison relative to a lateral ingroup comparison will differ depending on the comparative context. As the context becomes more interpersonal, an outperforming person becomes more threatening to personal identity, and thus participants should be relatively less likely to adopt a social categorization that they share with the other outperforming person. The stronger the intergroup aspects of the social context, the more that an outperforming person reflects positively on the shared group membership and makes it an attractive social categorization. On the other hand, when both the self and the comparison other perform poorly in an intergroup context, the shared group membership is relatively negative, and participants should be relatively less likely to adopt the shared group identity. Thus, we predicted an interaction in which, compared to participants faced with a lateral comparison, participants faced with an upward ingroup comparison would be more likely to apply the shared category membership to the self in an intergroup context, and relatively less likely to accept the shared categorization in an interpersonal context.

Experiment 1

In Experiment 1, we followed a procedure based on that of Schmitt et al. (2000; see also Martinot, Redersdorff, Guimond, & Dif, 2002). Female participants completed a difficult performance task in an interpersonal or an intergroup context. The comparative context was manipulated by telling participants that the purpose of the research was to compare individuals (interpersonal) or to compare women and men (intergroup). Participants received their actual score on the task, which we assumed would be quite negative given the difficulty of the task. Participants were then exposed to a female target, who was ostensibly another participant in the study. This comparison target either performed very well on the task and better than participants, or performed very poorly.

Participants then completed measures intended to tap the degree to which they accepted

categorization in terms of their gender group—a category shared by the comparison target and the participant, and relevant to the comparative context invoked in the intergroup comparison condition. We measured acceptance of the gender group categorization by examining two different consequences of group-level categorization. One consequence of categorization at the group level is an increased affective attachment to the group identity. Another more cognitive consequence of group-level categorization is the perception of the individual self as more similar to the ingroup prototype. As predicted by SCT (Turner et al., 1987), categorizing the self as a group member results in increasing depersonalization of the self, with the self being seen in less idiosyncratic terms and as more similar to the ingroup prototype (e.g., Simon & Hamilton, 1994; Spears, Doosje, & Ellemers, 1997).

Participants first completed a measure of gender group identification, tapping their affective connection to the group. To measure prototypicality, participants completed an open-ended self-description measure that was later coded in terms of gender prototypicality. On both the group identification and prototypicality measures we anticipated a cross-over interaction between the manipulations of comparative context and performance of the target. We predicted that compared to a lateral ingroup comparison, an upward ingroup comparison would lead to relatively low identification and prototypicality in the interpersonal condition, but relatively high identification and prototypicality in the intergroup condition.

<div align="center">Method</div>

Participants and Design

Female undergraduates ($N = 56$) participated in the study for partial fulfilment of their research requirement for an introductory psychology course. Participants were randomly assigned to one of four conditions in a 2 (interpersonal or intergroup context) x 2 (target performance good or bad) between-subjects factorial design.

Procedure

   Cover story. Participants individually met with a female experimenter who explained that the study involved the completion of a measure of creativity. To maximize the relevance of the test, the experimenter added that the test was well-validated by years of research and highly predictive of general intelligence and academic performance.

   Comparative context manipulation. In the interpersonal comparison condition, the experimenter went on to say: "We are particularly interested in how different individuals perform on the creativity test. When the study is complete, we'll look at the individual responses and see how different aspects of individual personality relate to creativity." In the intergroup comparison condition, however, she said: "We are particularly interested in how men and women might perform differently on this test. When the study is complete, we'll put all of the men's scores in one group and all the women's scores in another group and see how gender relates to creativity."

   Task feedback. Before completing the performance task, participants answered a few questions about themselves including their gender and other personal information, such as what they liked to do in their spare time and their favourite subjects in school. These questions were included to disguise the true purpose of the study. On the bottom half of this questionnaire sheet was the creativity test, which consisted of 12 difficult "trigrams" taken from a set developed by McFarlin and Blascovich (1984). For each trigram, participants had to find a word that fit with three others (the trigram). For example, the answer to "room, snow, base" would be "ball." After giving the participant five minutes to work on the task, the experimenter stopped the participant and scored the test in front of her. The experimenter left the participant's scored test out on the table so that it was visible during the rest of the study. On average, participants answered very few trigrams correctly ($M = 2.60$, $SD = .93$).[1]

   Target performance manipulation. The experimenter then told participants "For the next part of

the study, you'll read about someone who participated in the study previously. You are going to pick one of the questionnaires out of this folder—each one was filled out by another participant in this study." To reinforce the interpersonal or intergroup comparison manipulation, the experimenter said in the interpersonal condition "Their test scores and other information will be a part of the same analysis that yours will, where we look at how different people's personal characteristics are related to creativity." In the intergroup comparison condition she said "This person's test score will be a part of the same analysis that yours will, where we compare the average score for women and the average score for men to see if men and women might differ in creativity." The participant was then instructed to take one completed questionnaire from the folder, which actually contained many identical bogus questionnaires. The questionnaire was like the one previously completed by the participant, with the questions about gender and personal characteristics at the top of the sheet, and the creativity test at the bottom. In all conditions, the answers to the self-description questions were identical, and the target was always described as a woman. In the good-performing target condition, the target answered nine of the 12 trigrams correctly. In the poor-performing target condition, the target answered two of the 12 trigrams correctly. The experimenter told the participant that she would later answer some questions about the target, and instructed her to take a minute to read over the target's self-description.

Gender group identification. We measured gender group identification using six items (e.g., "I value being a member of my gender group," and "I would feel good if described as a typical member of my gender group") from a measure used previously (Schmitt & Branscombe, 2001). Participants responded to each item using a 1 to 7 Likert-type response scale. Responses to the individual items were reverse-scored where appropriate and then averaged to create a gender identification score ($\alpha = .75$).

Gender prototypicality. We measured participants' expressed gender prototypicality using an

open-ended measure. The questionnaire instructed participants to "Tell us about yourself. Please write down all of the things about yourself that come to mind, anything at all." Two coders, unaware of the conditions, categorized each thought listed by participants as stereotypically feminine (e.g., "I'm a good listener"), stereotypically masculine (e.g., "I'm assertive"), or as irrelevant to gender stereotypes (e.g., "I like movies"). The two coders' counts of feminine self-descriptions ($r = .75$), masculine self-descriptions ($r = .80$), and gender-irrelevant self-descriptions ($r = .74$) were all highly correlated. In addition, the coders categorized the self-descriptions as positive, negative, or neutral. The counts for each of those categories were also highly correlated ($r = .84$, $r = .73$, $r = .69$). The two coders met to resolve discrepancies in their coding.

Because self-categorization theory defines prototypicality in terms of both similarity to the ingroup and difference from the outgroup (Turner et al., 1987), we operationalized ingroup prototypicality as the number of feminine self-descriptions minus the number of masculine self-descriptions. Scores on this measure ranged from -3 to 7 ($M = 2.05$, $SD = 2.26$).

Perceptions of performance. As a manipulation check, we measured participants' perceptions of their own and the target's performance, "How well did you [the person you read about] do on the creativity test?" Participants responded on a 7-point scale ranging from "Extremely Poorly" to "Extremely Well." Additionally, we asked participants to report the target's score. One participant incorrectly answered this question, and was not included in the analysis.

Debriefing. After completing the questionnaire, participants were fully debriefed regarding the true purposes of the study. The experimenter made sure she explained to participants that the test was designed to be difficult, and that the target information they received was bogus.

Results

Manipulation check. We analyzed perceptions of performance using a mixed-model ANOVA including the manipulations of the target's performance and comparative context as between-

subjects factors, and whether participants rated the performance of themselves or the target as a within-subjects factor. Main effects of the manipulation of target performance, $F(1,51) = 79.91$, $p < .001$, and whether the participants were rating themselves or the target, $F(1,51) = 341.31$, $p < .001$, were qualified by a significant interaction between theses two factors, $F(1,51) = 314.17$, $p < .001$. In the good-performing target condition, participants perceived the target's performance ($M = 6.06$, $SD = .41$) more positively than their own ($M = 1.92$, $SD = .76$), $F(1,53) = 5.87.94$, $p < .001$. In contrast, for participants in the poorly-performing target condition, perceptions of performance did not differ significantly depending on whether participants were rating the target's performance ($M = 2.33$, $SD = .92$) or their own performance ($M = 2.27$, $SD = 1.05$), $F(1,53) = .07$, $p = .671$. Thus, as intended, participants perceived themselves as being out-performed by the target in the good-performing target condition, but saw the target's performance as similar to their own in the poor-performing target condition. The mixed-model ANOVA revealed no other significant main effects or interactions.

Group identification. As predicted, an ANOVA revealed a marginal cross-over interaction between the comparative context manipulation and the target performance manipulation affecting group identification, $F(1,51) = 3.81$, $p = .057$, $\eta^2 = .069$. The main effects of comparative context and target performance were not significant, $F(1,51) = 1.00$, $\eta^2 = .019$, $F(1,51) = 0.50$, $\eta^2 = .010$. As shown in Table 1, the direction of the effect of the target performance manipulation on group identification tended to differ depending on whether the comparative context was interpersonal or intergroup. In the interpersonal condition the effect of target performance did not significantly affect group identification, although the pattern of means was as predicted, with relatively lower group identification expressed when the target performed well, $t(51) = 0.92$, $p = .362$, $d = 0.329$. In the intergroup condition, we observed the reverse pattern—an upward comparison led to marginally higher group identification relative to the condition in which the target performed poorly, $t(51) =$

1.81, $p = .077$, $d = 0.779$. Thus, although the results for group identification are marginal in terms of statistical significance, the pattern of means is consistent with the predicted interaction.

Self-described prototypicality. Because stereotypes of women and men can differ in valence (Eagly & Mladinic, 1989), we analyzed self-described prototypicality in an ANCOVA controlling for the number of positive, negative, and neutral self-descriptions. This analysis examines the effect of the manipulation on prototypicality independent of any effects on the overall valence of self-descriptions. None of these covariates were significant predictors of prototypicality.

As predicted, the comparative context and target performance manipulations significantly interacted to affect the prototypicality of self-descriptions, $F(1,48) = 8.64$, $p = .005$, $\eta^2 = .153$. Neither the main effect of comparative context or target performance was significant $F(1,48) = 1.46$, $\eta^2 = .029$, $F(1,48) = 0.80$, $\eta^2 = .016$. As shown in Figure 1, the direction of the effect of the target performance manipulation on prototypicality depended on whether the comparative context was interpersonal or intergroup. In the interpersonal condition, participants tended to describe themselves as less prototypical when the target performed well compared to poorly, although this effect did not reach statistical significance, $t(48) = 1.65$, $p = .106$, $d = 0.688$. In contrast, participants in the intergroup condition expressed greater prototypicality when the target performed well compared to poorly, $t(48) = 2.72$, $p = .009$, $d = 1.028$.

Discussion

Experiment 1 offered some support for our prediction that the comparative context would moderate the effect of social comparison on acceptance of a shared category membership. As predicted, participants were most likely to describe themselves as prototypical of their gender group (a category membership shared with the source of the social comparison) when that categorization had relatively positive implications for the contextually relevant level of identity. Thus, self-described prototypicality was highest for participants exposed to a poorly performing target in the

interpersonal condition, and for those exposed to a high-performing target in the intergroup condition. Conversely, when shared categorization had negative implications for the context-relevant identity, participants were less likely to accept the shared group categorization. Self-described prototypicality was lowest for participants who were exposed to a high-performing target in the interpersonal condition, and for those exposed to a poorly performing target in the intergroup condition. Although the predicted interaction was supported for prototypicality, the effect of target performance only reached statistical significance in the intergroup comparison condition.

The pattern of means for gender group identification closely mirrored the results for prototypicality, but the effects were weaker and marginal in terms of statistical significance. In hindsight, it is perhaps not surprising that effects were weaker for gender group identification. Unlike the fairly non-reactive measure of prototypicality (participants had no reason to suspect that we would code responses in terms of gender), the gender group identification measure is likely to make salient the meaning of gender identity across a number of contexts other than our lab situation. Therefore because the measure may tap ongoing commitment to the identity across contexts, it may be less susceptible to contextual shifts than the cognitive consequences of group-level categorization (e.g., depersonalization). We suspect that affective identification with categories that have pervasive social consequences (e.g., gender) is heavily influenced by people's understanding of that category within a broader social context. For example, Tajfel (1978) described the emergence of a minority group identity in terms of the subjectively perceived status positions of groups in society, permeability of boundaries between groups, and the contestability of the existing status relationship. Therefore, variations in the immediate context could have relatively less influence on group identification than contextual shifts in perceptions of the social structure that Tajfel outlined. Nonetheless, our results suggest that group identification is not invulnerable to changes in the social context with no apparent relation to the broader relations between the

groups—the predicted cross-over interaction between our contextual manipulations approached statistical significance.

Experiment 2

Experiment 2 used a procedure similar to Experiment 1, but extended it in two important ways. In Experiment 1, participants received their actual scores rather than bogus negative feedback. Although providing real scores makes the feedback highly believable, this procedure also introduces error because participants do not all receive the same score. We addressed this issue in Experiment 2 by providing participants with non-contingent feedback concerning their performance.

Secondly, we extended the findings of Experiment 1 by using a different measure of the acceptance of shared category membership. In Experiment 1, the gender group identification and gender prototypicality measures could be interpreted as ways of increasing (or decreasing) an *interpersonal* connection with the comparison target, rather than as indicators of group-level categorization per se. Thus, we employed a different measure of the acceptance of a group-level categorization that applies more directly to intergroup perceptions—an open-ended measure of intergroup differentiation. We assumed that participants were more likely to emphasize differences between the ingroup and outgroup the more that they accepted a group level categorization. As predicted by SCT (Turner et al., 1987), one result of self-categorization at the group level is an accentuation of intergroup differences. Intergroup differentiation is a theoretically important outcome of categorization, and also qualitatively different from the dependent measures in Experiment 1. Although both group identification and prototypicality of self-descriptions are intragroup measures (i.e., concerning the individual's relationship to the group), intergroup differentiation is a more group-level variable, as it concerns the relationship between groups.

As in Experiment 1, female participants completed a performance task in an interpersonal or intergroup context. We increased the credibility of the negative task feedback by having participants

complete the task and receive feedback via a computer. After receiving the negative performance feedback, we exposed participants to a female target who ostensibly had performed better than the participant on the task, or similarly poorly. Participants then completed an open-ended measure of intergroup differentiation between women and men, as an indicator of the degree to which they accepted the shared gender category membership. As in Experiment 1, we predicted that the comparative context and the target's performance would interact to affect intergroup differentiation. Relative to a lateral ingroup comparison, we predicted that an upward ingroup comparison would lead to relatively higher intergroup differentiation in an intergroup context, and relatively lower intergroup differentiation in an interpersonal context.

Method

Participants and Design

Female undergraduates ($N = 77$) participated for partial fulfilment of their research requirement in an introductory psychology course. Participants were randomly assigned to one of four conditions in a 2 (interpersonal or intergroup context) x 2 (good or bad target performance) between-subjects factorial design. The procedure was similar to that of Experiment 1, except that participants received bogus feedback regarding their performance rather than their actual scores. Participants individually met with a female experimenter who presented the same cover story as in Experiment 1.

Procedure

Comparative context manipulation. We manipulated the comparative context as in Experiment 1. The experimenter stated that the purpose of the study was either to compare the scores and characteristics of individuals (in the interpersonal condition) or to compare the scores of women and men (in the intergroup condition).

Task feedback. As in Experiment 1, participants answered a few questions about themselves

before completing the performance task. Participants completed the creativity test on a computer, which we assumed would enhance the credibility of the false performance feedback. The trigram items were identical to those of Experiment 1. After five minutes, the experimenter stopped the participant and asked her to click on the button to submit her responses to be scored. After a brief delay, the participant saw a screen containing the false feedback. The screen contained a number of bogus statistics and, most visibly, the grade "D." Participants read that the possible scores could range from "A" (being the best possible score) to "F" (being the worst). The experimenter then printed a copy of the performance feedback screen, and left the sheet out on the table so that it was visible during the rest of the study.

    Target performance manipulation. As in Experiment 1, the experimenter told the participant that for the next part of the study she would read about a fellow participant. As in Experiment 1, the experimenter reinforced the comparative context manipulation by telling participants that the target's score would be part of the same analysis as the participant's, and by reminding the participant of the relevant comparison ostensibly of interest to the experimenters. The participant was then instructed to draw one completed questionnaire from the folder, which actually contained a number of identical bogus questionnaires. The questionnaire was like the one previously completed by the participant, with the questions about gender and personal characteristics on the top sheet, and the creativity test underneath. The personal information was on a half-sheet of paper, so that the target's score was clearly visible. In all conditions, the answers to the self-description questions were identical, and the target was always described as a woman. In the good-performing target condition, the target received an "A," and in the poor-performing target condition the target received a "D." The experimenter then told the participant that later she would answer some questions about the target, and instructed her to take a minute to read over the target's self-description.

Intergroup differentiation. We measured intergroup differentiation with an open-ended question that asked participants to "list as many possible ways that you can think of that women and men differ from each other." Two coders, who were blind to condition, counted the number of statements about either gender. The totals for the two coders were highly correlated, $r = .97$. When the counts for the two coders differed, a third coder settled the disagreement. Participants listed 0 to 11 ways in which the groups differed ($M = 3.63$; $SD = 3.61$).

Perceptions of performance. We measured perceptions of the participant's and target's performance using the same items as in Experiment 1. In addition, we asked participants to report the target's gender and grade on the creativity test. Four participants incorrectly answered the question about either the target's gender or the target's grade. These participants were not included in the analyses, and were not differentially distributed across the four experimental conditions, $\chi^2 (3) = 2.54$, $p = .468$.

Debriefing. After completing the questionnaire, participants were fully debriefed regarding the true purposes of the study and probed for suspicion regarding the false feedback. The experimenter made sure she explained to participants that the feedback they received was not representative of their actual performance, that the test was designed to be difficult, and that the target information they received was bogus. Eight participants, who were not differentially distributed across experimental conditions, expressed some suspicion in the debriefing, $X^2 (3) = 4.10$, $p = .250$. However, none of the participants came close to guessing the true purposes of the study, and thus, we conservatively retained them for the analyses. Excluding these participants does not change the pattern of results reported below.

Results and Discussion

Manipulation checks. We analyzed perceptions of performance using a mixed-model ANOVA including the manipulations of the target's performance and comparative context as between-

subjects factors, and whether participants rated their own or the target's performance within-subjects factor. Main effects of the manipulation of target performance, $F(1,69) = 93.81, p < .001$, and whether the participants rated themselves or the target, $F(1,69) = 290.70, p < .001$, were qualified by a significant interaction, $F(1,69) = 242.90, p < .001$. In the good-performing target condition, participants perceived the target's performance ($M = 6.54, SD = 1.10$) more positively than their own ($M = 1.97, SD = 1.01$), $F(1,71) = 406.21, p < .001$. In contrast, for participants in the poorly-performing target condition, perceptions of performance did not differ significantly depending on whether participants were rating the target's performance ($M = 2.56, SD = .93$) or their own performance ($M = 2.35, SD = .92$), $F(1,71) = .99, p = .323$. Thus, as intended, participants perceived themselves as being out-performed by the target in the good-performing target condition, but saw the target's performance as similar to their own in the poor-performing target condition. The mixed-model ANOVA revealed no other significant main effects or interactions.

Intergroup differentiation. As predicted, comparative context and target performance interacted to affect intergroup differentiation, $F(1,69) = 6.52, p = .013, \eta^2 = .086$. The main effects of comparative context and target performance were not significant, $F(1,69) = 0.13, \eta^2 = .002$, $F(1,69) = 0.42, \eta^2 = .006$. As shown in Figure 2, the direction of the effect of the target performance manipulation on intergroup differentiation depended on whether the comparative context was interpersonal or intergroup. In the interpersonal condition participants engaged in significantly less intergroup differentiation when the target performed well compared to poorly, $t(69) = 2.21, p = .031, d = 0.736$. In the intergroup condition, intergroup differentiation was not significantly affected by target performance, $t(69) = 1.38, p = .171, d = 0.459$. However the pattern of means was consistent with predictions--participants tended to engage in more intergroup differentiation when the target performed well compared to poorly. Thus, the pattern of the observed interaction was consistent with predictions.

Ancillary analyses across studies. Although the predicted interaction pattern observed across studies confirmed predictions, the pattern of statistically significant simple effects was not consistent across studies. In study 1 the simple effect was reliable only in the intergroup condition, and in study 2 the simple effect was only reliable in the interpersonal condition. Furthermore, of the three measures, group identification failed to produce a significant simple effect in either comparative context. However, comparing whether or not an effect reaches the level of statistical significance is a crude way to examine variability in effect sizes across studies or measures. Therefore we chose to examine the simple effects meta-analytically across studies, within each comparative context condition. Across studies and measures, participants in the interpersonal context who faced an upward comparison showed lower levels of category acceptance than did those who faced a lateral comparison, $d = -0.59$, $p = .005$. Participants in the intergroup context who faced an upward comparison showed higher levels of category acceptance than did those who faced a lateral comparison, $d = 0.71$, $p = .001$. Thus, in both interpersonal and intergroup contexts, we found evidence for the predicted effects of social comparison on the acceptance of a shared group categorization. Furthermore, the mean effect sizes in each condition fell within the medium to large range according to Cohen's (1988) standards. More importantly, results suggest that the heterogeneity in effect sizes did not reach the level of statistical significance across the interpersonal conditions, $Q\ (df = 2) = .75$, $p = .69$, or the intergroup conditions, $Q\ (df = 2) = 1.25$, $p = .53$. In other words, the strength of the simple effects does not differ significantly across studies (or measures). The results of the meta-analysis suggest that the observed effects are fairly robust, and less inconsistent than they might appear at first glance.

<div align="center">General Discussion</div>

In two studies we examined how the level of identity activated by the comparative context moderates the extent to which people will self-categorize as sharing a group membership with the

source of an upward comparison. Tesser's (1988) self-evaluation maintenance model suggests that the threat to personal identity invoked by an upward comparison can be avoided by decreasing the degree of closeness between the self and the comparison. Building on this idea, Mussweiler et al. (2000) predicted that individuals threatened by an outperforming other would decrease closeness with the other person by de-emphasizing social category memberships that they share. Our findings confirmed this prediction *in an interpersonal context*. However, we predicted and found that an upward ingroup comparison had the opposite effect in an intergroup context. According to social identity theory (Tajfel & Turner, 1986), the consequences of an individual ingroup member's performance for one's personal identity are less important than the implications of the ingroup member's performance for the ingroup identity as a whole. Relative to comparisons with a poorly performing ingroup member, an upward ingroup comparison reflects positively on the ingroup identity, making it more attractive. Thus, in contrast to SEM, SIT predicts that an upward ingroup comparison can lead people to embrace rather than reject shared category memberships with an outperforming comparison target.

These findings support our assertion that SEM and SIT make different predictions because they apply to different identity concerns that arise in different comparative contexts. SEM applies to comparative contexts that invoke personal identity; SIT applies to contexts that invoke group identity. We predicted and found that upward comparisons have opposite effects on the acceptance of a group-level categorization shared with the source of the comparison, depending on whether the context is interpersonal or intergroup. We observed this pattern of effects on three different indicators of the acceptance of a group-level social categorization: self-described prototypicality, intergroup differentiation, and to a lesser degree, gender group identification.

Integrating Individual and Group-Level Theories

These findings are consistent with self-categorization theory's (Turner et al., 1987)

assumption that the self can be defined at different levels of inclusiveness. According to SCT, motivational pressures toward a positive identity operate differently depending on what level of identity is activated by the context. Indeed, one important contribution of our work is that it demonstrates a dynamic picture of identity and self-categorization at the group-level. Although the comparative context clearly moderated the effects of social comparison, participants did not "automatically" accept the categorization invoked by the context. We did *not* find a main effect of the comparative context manipulation, such that participants accepted the group-level self-categorization in the intergroup context but rejected it in the interpersonal context. Rather, the comparative context invoked different levels of identity, and participants expressed a level of self-categorization consistent with the protection of the level of identity invoked in context. Interestingly, our findings demonstrate that the acceptance of a group categorization can be driven by either personal or group identity concerns, depending on the level of identity made relevant in context.

"Individual-level" and "group-level" theories of the self can easily be construed as competing perspectives. The present work, however, suggests that such a construal is misleading. Individual-level and group-level theories are compatible to the extent that their differing predictions apply to different contexts and different levels of self-definition. Importantly, because SCT can account for psychological phenomena in both intergroup and interpersonal contexts, it provides a more parsimonious theoretical account of our findings than SEM and SIT together. SCT can account for the differing effects across our experimental conditions because it assumes that the comparative context alters the meaning of "self" and thus shapes the meaning of social comparisons. In contexts that invoke an intergroup comparison, upward ingroup comparisons moves the group's identity toward a more positively distinct position. To put it even more definitively in self-categorization terms, the upward ingroup comparison represents a more

prototypical group identity within the normative context of the performance task and experimental setting. When the group categorization was activated by the context, participants more readily accepted that level of categorization the more than an ingroup member made the ingroup identity more positive within the context of the study.

In situations that make individual identity salient, SCT predicts that self-evaluation occurs via comparisons between one's personal identity and other individuals—with self-evaluation being contrasted from the performance of other individuals. The relevance of these comparisons is increased by a shared level of categorization. Thus, in an attempt to protect a positive view of one's individual identity, participants faced with an upward comparison were relatively less likely to invoke a more superordinate level of categorization that would make lead to a more negative (personal) self-evaluation. Conversely, a lateral comparison represents no threat to one's personal identity, and thus group memberships shared with the comparison target do not represent any additional threat to personal identity. Furthermore, a lateral comparison in an interpersonal context makes the group-level categorization more attractive because it suggests that one's personal performance is normative within the bounds of that category.

Researchers who focus on one level of self-definition to the exclusion of others run the risk of ignoring important aspects of social psychological life. For example, Mussweiler et al.'s (2000) findings could lead to a limited understanding of the effects of social comparison on group identification without examining both interpersonal and intergroup contexts. Although arguments can be made that certain social phenomena are best understood in terms of social identities or individual selves, no social phenomenon is likely to operate exclusively at one level of identity. As self-categorization theory suggests, social psychology would benefit from careful attention to how the level of self-categorization invoked in context shapes the operation of psychological processes (Onorato & Turner, 2004; Turner et al., 1994; Turner & Onorato, 1999).

Caveats and Future Directions

In an attempt to measure the acceptance of a group-level categorization in several qualitatively different ways, we measured group identification, ingroup prototypicality, and intergroup differentiation. The results of the meta-analysis of effects across these three measures suggest that the strength of the effects did not differ between measures to a significant degree. Thus, while the pattern of significant simple effects varied across conditions, interpretation of these apparent inconsistencies should be done with caution. Nonetheless, the measures we employed were fairly novel and differed from each other in future ways. Future research might consider the meaning and implications of different measures of group-categorization like the one's we employed, as well as test the generalizeability of our effects to other indicators of group-level categorization.

Our results might seem to imply that individuals in intergroup contexts will try to distance from negative social identities (i.e., when the ingroup target performs poorly) and seek closeness to positive social identities (when the ingroup target performs well). However, SIT suggests that this is not always the case. First, negative social identities can be made more positive via "social creativity"—emphasizing comparative dimensions on which the ingroup is positively distinct. If participants in the intergroup condition of our studies had been provided with an opportunity to engage in social creativity, it is possible that they would not have distanced themselves from the ingroup when faced with a poorly-performing ingroup member (Ellemers & Van Rijswijk, 1997; Jetten, Schmitt, Branscombe, & McKimmie, in press; Spears & Manstead, 1989). Furthermore, when an ingroup's low status is contestable and can be made more positive through social change, the perception of low status can *increase* group identification (Jetten, Branscombe, Schmitt, & Spears, 2001). Thus, the effects we observed are likely to be limited to contexts in which opportunities to improve the ingroup's identity are not readily available. Similar qualifications can be made with regard to the consequences of social comparisons for individual identity. If participants in the

interpersonal condition had been offered an opportunity to improve their performance, an upward comparison may have been more inspiring than threatening (Lockwood, Jordan, & Kunda, 2002; Ybema & Buunk, 1993). Likewise, if these participants had been given an opportunity to affirm another aspect of their individual identities (Steele, 1988), they might have been less likely to distance from the ingroup when faced with an upward comparison.

<div align="center">Conclusions</div>

In two studies we found that the level of identity activated in context moderated the effect of an upward ingroup comparison on the degree to which participants categorized the self in terms of a group identity shared with the source of the comparison. Following self-categorization theory, we argued and found that in interpersonal contexts, upward ingroup comparisons can lead to relatively less acceptance of a shared categorization. We predicted and found that in intergroup contexts, upward ingroup comparisons can lead to relatively more acceptance of a categorization shared with the source of the comparison. These findings suggest that the acceptance of a particular self-categorization shifts in ways that are consistent with the protection of contextually relevant levels of identity.

References

Branscombe, N. R., Wann, D. L., Noel, J. G., & Coleman, J. (1993). Ingroup or outgroup extremity: Importance of the threatened identity. *Personality and Social Psychology Bulletin, 19,* 381-388.

Castano, E., Paladino, M. P., Coull, A., Yzerbyt, V. Y. (2002). Protecting the ingroup stereotype: Ingroup identification and the management of deviant ingroup members. *British Journal of Social Psychology, 41*, 365-385.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Eagly, A.H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*, 543-558.

Ellemers, N. (1993). The influence of socio-structural variables on identity enhancement strategies. *European Review of Social Psychology, 4*, 27-57.

Ellemers, N., & Van Rijswijk, W. (1997). Identity needs versus social opportunities: The use of group-level and individual-level identity management strategies. *Social Psychology Quarterly, 60*, 52-65.

Feather, N. T. (1994). Attitudes towards high achievers and reactions to their fall: Theory and research concerning tall poppies. *Advances in Experimental Social Psychology, 26,* 1-73.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117-140.

Gardner, W. L., Gabriel, S., & Hochschild, L. (2002). When you and I are "we," you are not threatening: The role of self-expansion in social comparison. *Journal of Personality and Social Psychology, 82,* 239-251.

Groesz, L. M., & Murnen, S. K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders, 31,* 1-16.

Haslam, S.A. (2001). *Psychology in organizations: The social identity approach*. London: Sage.

Jetten, J., Branscombe, N. R., Schmitt, M. T., & Spears, R. (2001). Rebels with a cause: Group identification as a response to perceived discrimination from the mainstream. *Personality and Social Psychology Bulletin, 27*, 1204-1213.

Jetten, J., Schmitt, M.T., Branscombe, N.R., & McKimmie, B.M. (in press). Maintaining group identification under conditions of devaluation: Intergroup differentiation and intragroup respect as social creativity responses. *Journal of Experimental Social Psychology*.

Kemmelmeier, M., & Oyserman, D. (2001). The ups and downs of thinking about a successful other: self construals and the consequences of social comparisons. *European Journal of Social Psychology, 31,* 311-320.

Lockwood, P., Jordan, C. H., & Kunda, Z. (2002). Motivation by positive or negative role models: Regulatory focus determines who will best inspire us. *Journal of Personality and Social Psychology, 83*, 854-864.

Marques, J. M., Yzerbyt, V. Y., & Leyens, J. P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology, 18,* 1-16.

Martinot, Delphine; Redersdorff, Sandrine; Guimond, Serge; Dif, Stephane. (2002). Ingroup versus outgroup comparisons and self-esteem: The role of group status and ingroup identification. *Personality and Social Psychology Bulletin, 28*, 1586-1600.

McFarlin, D. B., & Blascovich, J. (1984). On the Remote Associates Task (RAT) as an alternative to illusory performance feedback. *Basic and Applied Social Psychology, 5,* 223–229.

Musseiler, T. (2003). "Everything is relative": Comparison processes in social judgment. *European Journal of Social Psychology, 33,* 719-733.

Mussweiler, T., Gabriel, S., & Bodenhausen, G. V. (2000). Shifting social identities as a

strategy for deflecting threatening social comparisons. *Journal of Personality and Social Psychology, 79,* 398–409.

Onorato, R., & Turner, J.C. (2004). Fluidity in the self-concept: the shift from personal to social identity. *European Journal of Social Psychology, 34*, 257-278.

Pleban R, & Tesser, A. (1981). The effects of relevance and quality of another's performance on interpersonal closeness. *Social Psychology Quarterly, 44,* 278-285.

Schmitt, M.T., & Branscombe, N.R. (2001). The good, the bad, and the manly: Effects of threats to one's prototypicality on evaluations of ingroup members. *Journal of Experimental Social Psychology, 37,* 510-517.

Schmitt, M. T., Silvia, P. J., & Branscombe, N. R. (2000). The intersection of self-evaluation maintenance and social identity theories: Intragroup judgment in interpersonal and intergroup contexts. *Personality and Social Psychology Bulletin, 26,* 1598-1606.

Simon, B., & Hamilton, D.L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology. 66,* 699-711.

Spears, R., Doosje, B., & Ellemers, N. (1997). Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification. *Personality and Social Psychology Bulletin, 23,* 538-553.

Spears, R., & Manstead, A.S.R. (1989). The social context of stereotyping and differentiation. *European Journal of Social Psychology, 19*, 101-121.

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302). New York: Academic Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (2nd ed., pp. 7-24). Monterey,

CA: Brooks-Cole.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. *Advances in Experimental Social Psychology, 21,* 181-227.

Turner, J .C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford: Blackwell.

Turner, J.C., & Onorato, R. (1999). Social identity, personality, and the self-concept: A self-categorizing perspective. In T.R. Tyler, R.M. Kramer, et al., (Eds.) *The psychology of the social self. Applied social research* (pp. 11-46). Mahwah, NJ: Lawrence Erlbaum.

Weber, J. G., & Brewer, M. B. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology, 66,* 268-275.

Wood, J. V., & Wilson, A. E. (2003). How important is social comparison? In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 344-366). New York: Guilford.

Ybema, J. F., & Buunk, B. P. (1993). Aiming at the top? Upward social comparison of abilities after failure. *European Journal of Social Psychology, 23,* 627-645.

Footnote

[1] Because our theoretical predictions only apply to participants who felt they have performed poorly, we excluded participants who answered more than half of the trigrams correctly ($n = 10$). Importantly, these excluded participants were not differentially distributed across the four conditions of the experimental design, $\chi^2 (3) = 2.62, p = .454$. Both with and without these participants, the number of correct trigrams did not significantly covary with either of the two manipulations or their interaction.

Table 1. Means and standard deviations for group identification, Experiment 1

	Comparative Context	
Target Performance	Interpersonal	Intergroup
Poor-Performing	5.70 (.85)	5.49 (.90)
Good-Performing	5.42 (.85)	6.07 (.52)

Note. Standard deviations are shown in parentheses.

Figure Captions

Figure 1. Prototypicality of self-descriptions as a function of comparative context and target performance, Experiment 1. Means are adjusted for the number of positive, negative, and neutral self-descriptions. Standard deviations, from left to right: 2.20, 2.29, 2.21, 2.29.

Figure 2. Intergroup differentiation as a function of comparative context and target performance, Experiment 2. Standard deviations, from left to right: 4.06, 3.21, 3.41, 3.47.