Organizational Behavior and Human Decision Processes 122 (2013) 80–91

Contents lists available at SciVerse ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Just how comparative are comparative judgments?

Joseph R. Radzevick [a,*], Don A. Moore [b]

[a] Department of Management, Gettysburg College, 300 North Washington St., Gettysburg, PA 17325, United States
[b] Haas School of Business, University of California at Berkeley, United States

ARTICLE INFO

Article history:
Received 22 June 2012
Accepted 1 May 2013
Available online 1 June 2013
Accepted by Xiao-Ping Chen

Keywords:
Social comparison
Reaction time
Differential weighting

ABSTRACT

Three studies investigate the psychology of comparative judgment, examining the circumstances under which judgments tend to concentrate disproportionately on one of the two elements that underlie the comparison (i.e., focused comparisons). We examine these tendencies at the judgment formation and information retrieval stages by examining judgment content as well as the speed and efficiency with which people make comparisons. The results replicate prior findings of differential weighting, indicating that focusing occurs in the formation stage of comparative judgments. However, focusing is absent in the reaction time data, suggesting that both elements of the comparison are equally accessible when individuals retrieve comparison-relevant information. These findings clarify the process by which people arrive at comparative judgments and demonstrate precisely when focused comparisons occur.

Introduction

Some of the most important decisions in life depend on comparative judgments. For instance, the decision of whether to start a new business ought to depend on how much better your product will be than the products of the competition. The decision of whether to marry someone ought to depend on how compatible you are with that person, relative to other potential life partners. We ask whether people make these sorts of comparative judgments optimally, or whether the features of one target in a comparison (such as your own business product or current romantic partner) receive undue weight when making the comparison.

Rational theories of comparative judgment include what Tversky (1977) called the symmetry axiom: $\delta(a,b) = \delta(b,a)$, where δ represents the distance between two stimuli in feature space. In other words, the distance between a and b must be equal to the distance between b and a. The curious fact about the way people judge similarities and differences is that we routinely violate the symmetry axiom. As Tversky points out, "We say, 'the portrait resembles the person' rather than 'the person resembles the portrait'" (p. 328). The way we construct our sentences parallels the way we think about comparisons: We focus on a target for any comparative judgment, and resist the implication of the symmetry axiom that we should be willing to switch it with the referent to which we are comparing it and arrive at the same assessment of the size of the difference between the two.

Judgments of similarity are likewise prone to focus on a particular target. Tversky and Gati (1978, p. 85) point out that comparative tasks can assume two distinct forms (which we will subsequently refer to as balanced and focused, respectively):

1. "Assess the degree to which a and b are similar to each other.
2. Assess the degree to which a is similar to b."

Tversky and Gati observe that when making balanced comparisons, the features of a and b should be made salient and weighted equally and thus the similarity of a to b is equal to the similarity of b to a. Conversely, in focused comparisons the "subject" of the comparison (more commonly referred to as the target in the comparative judgment literature) will be more salient and weighted more heavily than the referent. This asymmetry leads to substantive differences between the similarity of a to b and the similarity of b to a. We will employ Tversky and Gati's terminology by referring to unbiased comparisons as "balanced" and comparisons that overweight the target as "focused."

Additional research provides evidence that individuals are subject to focused comparisons in which they seem to overweight the target of the comparison relative to the referent to which the target is being compared (e.g., focalism; see Kruger & Burrus, 2004; Windschitl, Kruger, & Simms, 2003). They do so when comparing inanimate objects (Giladi & Klar, 2002), people (Klar & Giladi, 1997, 1999), and even when the self is the referent against which other people are being compared (Moore & Kim, 2003). Such focused comparisons affect numerous judgment contexts. For example, sports fans are more certain their own team will succeed when the team is strong, regardless of the strength of the competition (Radzevick & Moore, 2008). Negotiators predict that final deadlines

* Corresponding author.
E-mail address: jradzevi@gettysburg.edu (J.R. Radzevick).

0749-5978/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.obhdp.2013.05.001

will prove more harmful to themselves than their opponents, despite the fact both sides are subject to the same constraint (Moore, 2004, 2005).

Single target focusing

Research on comparative judgment has been marked by an even bolder claim than Tversky made. Some prior evidence has suggested that "ostensibly comparative social judgments are often myopic solo judgments of a single actor" (Moore & Kim, 2003, p. 1121). Klar and Giladi (1997) called this "singular target focusing" and described it this way: "rather than judging how friendly the singular target is relative to the other group members, people simply judge the friendliness of the singular target. As a result, what is supposed to be a comparative judgment turns into a pseudo-comparative or even non-comparative judgment in which the singular target is evaluated with no clear counterpart or reference group" (p. 900).

Evidence for single target focusing suggests that comparative judgments may not actually arise from a truly comparative process and instead may depend primarily on single individual judgments. For instance, when Kruger (1999) asked people how their skills compared with those of others, he found that their answers had more to do with their own personal skills than those of others. Accordingly, people report that they are better than others at riding a bicycle but worse than others at juggling. Kruger (1999) described his results this way: "The answer to the question 'How do I compare with my peers?' is based considerably more on 'I' than 'my peers'" (p. 230).

Likewise, when people are asked how happy they are relative to others, they tell you about their own personal happiness (Klar & Giladi, 1999). When assessing relative health risks, individuals often base their estimates on their own levels of risk but not the risk levels of others (Aucote & Gold, 2005; Covey & Davies, 2004; Gold, 2007). When an instructor decides to allow students to consult their books during an exam, their expectations for getting a good grade go up, even when it is common knowledge that the exam will be graded on a forced curve (Windschitl et al., 2003). Individuals base their expectations about future life events more upon their own likelihood of experiencing the events than upon the likelihood of others doing so (Chambers, Windschitl, & Suls, 2003; Kruger & Burrus, 2004).

Other consequential decisions appear similarly biased by focused comparisons. Individuals are willing to bet more money that they will beat an opponent on an easy trivia quiz than on a difficult quiz, though both competitors face the same level of quiz difficulty (Moore & Small, 2007). Excess entrepreneurial entry can be attributed in part to "reference group neglect," whereby individuals fail to account for relevant reference groups even though success depends on relative outcomes (Camerer & Lovallo, 1999; Moore, Oesch, & Zietsma, 2007).

Comparative processes

Chambers and Windschitl (2004) describe a useful model of comparative judgment in which focused comparison mechanisms operate (see Fig. 1). Once decision makers have interpreted the comparative question in their minds (stage A), decision makers move to stage B: they retrieve from memory the information relevant to the individual elements of the subsequent comparative judgment (e.g., target and referent, self and other, etc.). Decision makers then make absolute assessments of the individual elements (stage C). Next, the decision makers form the desired comparative judgment between the elements (stage D). And finally, they report the result of that comparative judgment (stage E).

Chambers and Windschitl speculate that single target focusing effects can manifest across different stages of judgment. For example, consider a student predicting the likelihood that he or she will outscore a particular classmate on an upcoming exam. A focused comparison (in this case egocentrism) may arise in the acquisition stage as the student brings to mind more information about his or her own knowledge and ability than the classmate's due to the fact that self-relevant knowledge is more mentally accessible. A second distinct possibility is that focusing may occur in the judgment formation stage whereby information about the self and other are equally accessible, but the student weighs information about his or her own knowledge and skill more heavily when calculating the final comparative judgment.

The present studies

In this paper, we employ a new approach to trace the psychological processes at work when people make comparative judgments. The results imply that focusing occurs at the judgment-formation stage and not at the information-acquisition stage.

Building on the theoretical foundation laid by Chambers and Windschitl (2004), we examine the extent to which focused comparative processes occur at the information retrieval and comparison formation stages. In each of the three studies we present, participants evaluate several pairs of individual elements and comparisons between them. These include individual as well as group level elements, respectively corresponding to more generalized comparative judgments as well as more specific judgments subject to single target focusing. We vary the order of these judgments and examine judgments across a variety of domains. We utilize a multi-method approach to test at the appropriate stages. For information retrieval, we observe the reaction times associated with the judgments. We also examine the actual weighting of the elements in the comparisons that result from the comparative judgment that follows retrieval, and expect to find differential weighting in these judgments.

If focused comparisons are driven by focused information retrieval, we should expect the following. First, because such differential weighting implies that the comparison has more to do with the target than the referent of the comparison, we should find that comparative judgments are facilitated more by coming after individual judgments of the target than they are by coming after the referent. Second, we should also find that comparative judgments facilitate evaluations of the target more than they do evaluations of the referent. Third, we should find that the target is overweighted in comparative judgments relative to the referent. Our first study tests these hypotheses.

We also thought it was important to examine the effects of several potential moderators of the process of comparative judgment. First, because we are interested in many different types of comparative judgments, we varied the topic of the comparison. We included both self-relevant and impersonal judgments since prior research on egocentrism gives us some reason to expect that self-relevant judgments may exacerbate focused comparisons, particularly with the self as target (Chambers & Suls, 2007; Dunning & Hayes, 1996; Epley & Dunning, 2000; Krizan & Suls, 2008; Mussweiler & Bodenhausen, 2002). Finally, we varied the inherent ambiguity of the response scale: some use subjective verbally-anchored scales while others use objective numerically-anchored scales. Prior research has found that subjectively labeled scales necessitate the interpretation of scale labels to formulate an answer (Biernat, 2003; Schwarz, 1999). This process of disambiguation opens the door to differences between people in their interpretation of scale labels, with some important consequences (Biernat & Manis, 1994; Heine, Lehman, Peng, & Greenholtz, 2002; Moore, 2007). Individuals may respond faster on subjective scales since they can adapt those responses to their own idiosyncratic standards (Dawes, 1977; Windschitl & Wells, 1996).

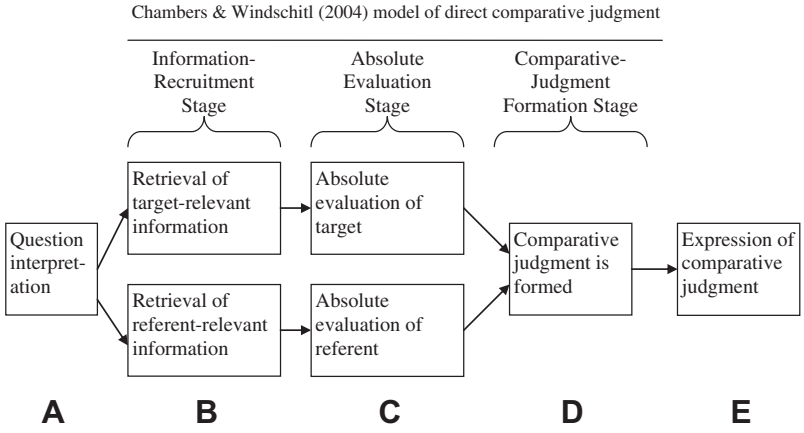


Fig. 1. Chambers and Windschitl's (2004) three-stage model of processes mediating direct comparative judgment, with the sensory stimulus input and the expressive output stages added.

Study 1

Method

Participants

Eighty-two individuals completed the study. Twenty-four were students at a university in Pittsburgh, Pennsylvania who participated for course credit while the remaining participants were recruited from a pool of local residents (both students and other community members) who participate in research for pay (34 female, 48 male; Mean age = 24.0 years, *SD* = 5.9).

Design

For each of 50 topics, participants provided two individual assessments for "Element A" and "Element B" and a comparison between the two, "Element A compared to Element B." By the conventions of this comparative judgment format, this made Element A the "target" of the comparison and Element B the "referent."

First, participants saw a description of the topic at the top of the computer screen (e.g., "Number of Academy Award nominations this person has received"). For each of the two individual assessments, a single element (e.g., "Meryl Streep" or "Elizabeth Taylor") appeared on either the left or right side of the screen and a response scale appeared at the bottom of the screen. This response scale was the same for each of the two individual judgments. Participants also made a comparative judgment (e.g., "Meryl Streep compared to Elizabeth Taylor") on a response scale that spanned the range of possible differences between the individual judgments. The order for the three judgments (Element A, Element B, Comparison) varied randomly for each topic. This meant that a given participant encountered different orders across topics.

The 50 topics came from two broad categories: factual and social. The 19 factual topics covered history, geography, science, and popular culture (e.g., "The number of scenes in this play"). For each topic, participants reported two individual assessments (e.g., *Othello* and *Macbeth*) and one comparison (e.g., *Othello* compared to *Macbeth*). We included these factual topics to create a contrasting group of "neutral" items that may not be as susceptible to weighting effects since they do not involve social comparison.

The remaining 31 social topics involved comparing people on characteristics and behaviors (e.g., their ages). These included nine items relevant to health behaviors (e.g., "The number of meals per week this person eats red meat") used by Klein and Kunda (1993), sixteen items representing more general social topics (e.g., "The number of pairs of blue jeans this person owns"), and six items predicting performance across tasks of various difficulty levels (e.g., "The number of US presidents this person can name"). The two individual elements for each social topic were drawn randomly from the following set: "You," "Another participant in this study," "Your best friend," "An average Pittsburgh resident," "Your mother," and "An average college student." Note that three of these represent collectives (study participants, Pittsburgh residents, and college students) and three do not (self, mother, and best friend). We needed to vary the set of individuals being judged because the alternative – just having participants evaluate self and other – made the task too predictable.

Response scales

Participants recorded their answers by clicking somewhere on a 101 point response scale labeled at five equally spaced points. We varied whether these labels were objective (numerical values such as 0, 5, 10, 15, and 20) or subjective (verbal descriptions such as Very few, Few, Some, Many, and Very many). Objective scales were calibrated so as to be wide enough to include 90% of the responses given on a pre-test of the same topics. The corresponding scales on the comparative judgments included the full possible range implied by the two individual judgments, extending from the negative maximum values of the individual scales to the positive maximum values of the individual scales with zero as the midpoints (e.g., −20, −10, 0, 10, 20). We based the comparative judgment scales upon their corresponding individual judgment scales in a similar manner (e.g., Far fewer, Some fewer, About the same, Some more, Many more).

Procedure

The experimenter provided instructions out loud and addressed any questions. Participants then were instructed to complete each item as quickly as possible without sacrificing accuracy. They completed the task on computers that timed all responses and randomized the order of topics, the response scale for a given topic (either objective or subjective), the two elements for each social topic, the designation of elements as A or B, and the order in which the three judgments (the two individual elements and the comparison) appeared within a given topic.

Results

Descriptive statistics for the reaction times and judgments across various categories appear in Appendix A. In this and the following studies, we used a natural log transformation of reaction times for all statistical analyses in order to reduce the influence

of skew and outliers. However, we report means in seconds in the interest of interpretability. We removed 18 observations that fell more than five standard deviations from the mean and dropped data from two participants because they failed to follow instructions. This left 3982 topic observations (each with three judgments) from 80 participants. The effects of our independent variables are shown in Fig. 2. We conducted the following analyses as a linear mixed model, controlling for the individual participants responding across the 50 topics and using the following three independent variables: question order (comparative judgment's position in the sequence: first vs. middle vs. last), element category (factual vs. non-self social vs. self-related social), and response scale (subjective vs. objective).

Response scale type

People were faster responding to subjectively labeled scales ($M = 4.2$ s, $SD = 2.8$) than objectively labeled scales ($M = 5.1$ s, $SD = 3.2$), $F(1,91) = 16.6$, $p < .001$; $d = .29$. This by itself is interesting, because it suggests that the vague verbal scale labels may represent a less effortful response mode. The fact that we observe this difference for all judgments and across all orders suggests that it affects a stage which participants go through for all questions, including those where the actual content of the judgment has been primed on a previous question. Because the response scale is most relevant to stage E (the expression stage), we infer that the response scale (subjective vs. objective) speeds the translation of judgments to responses on the scale. However, because the effect of subjective vs. objective response scales is not the focus of this paper, and because we did not observe any interactions between our manipulation of response scale type and other variables of interest, we do not explore it further in our analyses of the results.

Comparison facilitation in information retrieval

If comparative judgments depend on both individual elements, we should observe a main effect of element order so that comparative judgments are made more quickly when they come after just one of the individual judgments and faster still when they come after both.

Indeed, this main effect of order emerges strongly, $F(2,107) = 227.4$, $p < .001$. Pairwise comparisons show that participants made comparative judgments faster when they came in the middle ($M = 4.7$ s, $SD = 2.8$) rather than first ($M = 5.8$ s, $SD = 3.2$), $p < .001$; $d = .40$, and fastest when they came last ($M = 3.4$ s, $SD = 2.7$) rather than in the middle, $p < .001$; $d = .61$. These results clearly show that comparative judgments were facilitated by coming after individual judgments.

The order × element category interaction also attains significance, $F(4,3812) = 4.51$, $p < .001$. As Fig. 2 shows, this interaction results from the fact that order made more of a difference for factual and non-self social topics than it did for social topics involving the self. When comparative judgments came first, participants made self-relevant social comparisons more quickly than any other type of comparison. To test this directly, we conducted a new mixed model for only cases in which the comparison occurred first. The effect for topic category is significant, $F(2,105) = 16.8$, $p < .001$. Pairwise comparisons show that participants made comparative judgments more quickly when they involved the self ($M = 5.1$ s, $SD = 3.0$) than they did for either factual ($M = 6.2$ s, $SD = 3.1$), $p < .001$; $d = .39$, or social comparisons that did not involve the self ($M = 6.1$ s, $SD = 3.6$), $p < .001$; $d = .28$.

In short, it seems clear from this initial evidence that comparative judgments depend on, and are facilitated by, both the individual judgments.[1] These results do not support the most radical form of a focused comparison account, namely that comparative judgments are, at heart, really just a single individual judgment.

If differential accessibility leads to focused comparisons, we should expect to find two additional results in the reaction times: (1) comparative judgments in the middle should be faster when they follow the target than when they follow the referent, and (2) when they follow the comparative judgment, individual evaluations of the target should be faster than evaluations of the referent. But neither of these effects is present in the data.[2] With respect to the first, we find that participants made comparative judgments just as fast when they came after Element B ($M = 4.7$ s, $SD = 2.8$) as when they came after Element A ($M = 4.8$ s, $SD = 2.7$), and a test of the difference in a mixed model yields $F(1,1307) = .09$, $p = .77$; $d = 03$. With respect to the second, when Element A followed the comparison, it was no faster ($M = 3.5$ s, $SD = 2.5$) than when Element B followed the comparison ($M = 3.6$ s, $SD = 2.4$), $F(1,162) = 1.28$, $p = .26$; $d = .09$.[3]

These results provide no evidence that differential accessibility at the information-recruitment stage (stage B) of the comparative judgment causes focused comparisons.

Differential weighting in judgment formation

Next, we wanted to test whether our results reveal differential weighting of the target over the referent in the formation of the comparative judgment (stage D). We conducted a series of regression analyses on the comparative judgments across each domain using the corresponding individual judgments as independent

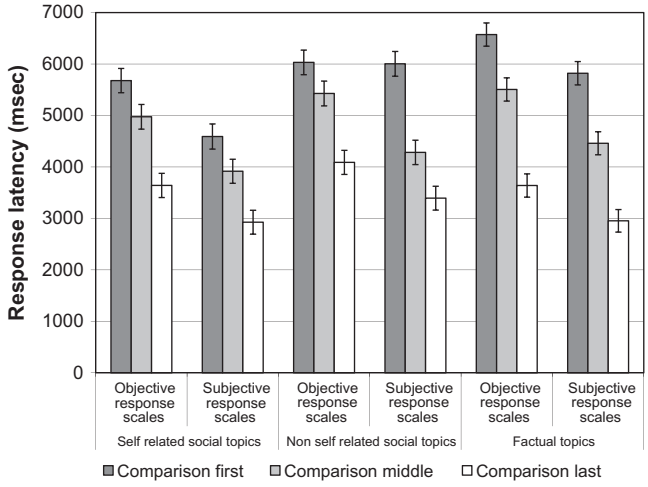


Fig. 2. Response latencies of comparative judgments as a function of comparison position (first vs. middle vs. last), response scale (subjective vs. objective), and topic category (self related social vs. non-self related social vs. factual). Error bars show standard errors (Study 1).

[1] To examine whether the speed of comparisons was affected by the collective vs. individual nature of the referent category, we tested for differences between our referent categories. When "best friend" was the referent, it was indeed the case that comparative judgments were made faster than all other referents (at least $p < .05$) except for "An average college student" ($p = .10$). No other significant differences between referents emerged. It is important to note that the referent category variable did NOT interact with the comparison position variable ($F(8,1145) = .309$, $p = .963$), consistent with our primary findings. Thus our pattern of results (here and also in subsequent studies) holds regardless of whether the comparison other represented an individual or collective entity.

[2] This pattern holds across the various question orders, question types, and response scales.

[3] Alternatively, this finding may reflect a task routinization effect whereby participants became more accustomed to giving the two individual judgments after the comparison. This could lead to similar reaction times even if Element A was the only one specifically activated by the comparative judgment. While supplementary analyses indicate that respondents in general make judgments faster in later rounds, this occurs at an equivalent rate for Element A, Element B, and the comparison of the two. Furthermore, these analyses reveal no significant effects for such learning over time for either Element A or Element B specifically when it follows the comparison.

Table 1
Results of regressions for comparative judgments on Element A and Element B absolute judgments (Study 1).

Independent variable	Self related social comparisons		Non-self related social comparisons	Factual comparisons
	Self is Element A	Self is Element B		
Element A absolute judgment	0.75***	0.66**	0.69***	0.66***
Element B absolute judgment	−0.50***	−0.58**	−0.49***	−0.55***
Adjusted R^2	0.60	0.53	0.55	0.43

Notes: Table show standardized beta weights. All coefficients are significant at $p < .001$. Asterisks indicate a significant Wald test for differences between the absolute weightings of Element A and Element B coefficients in a given model.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

variables (see Table 1). In general, both elements significantly affected the comparative judgment (all coefficient p's < .001). We find, however, that the weight placed on the two elements is not equal: the comparative judgment weights the target (Element A) more heavily than the referent (Element B). This holds across all topics (all p's < .01). The differential weighting pattern is consistent with a focalism account. This means that for self-relevant comparisons, the self is weighted more heavily only when it appears as Element A and non-self elements are weighted more heavily when they appear as Element A.

Discussion

The results of this first study present a bit of a puzzle. On the one hand, the reaction time evidence we present suggests that individuals engage in relatively balanced information retrieval in the initial stage. This is supported by the fact that comparisons were facilitated by prior assessment of the individual elements being compared in an equivalent manner and vice versa.

On the other hand, this balanced approach did not carry over to the judgment formation stage. The actual content of participants' comparisons shows differential weighting of target over referent, similar to that found in previous research. This differential weighting strongly suggests that our respondents' comparative judgments did depend more strongly on the absolute assessment of the target than of the referent.

The apparent contradiction helps make a clear distinction between activation vs. application. Together these observed patterns suggest that individuals consider the two elements relatively equally in the information-recruitment stage of judgment but then weight the target more heavily when it comes time to actually form the comparative judgment. In other words, the findings suggest that participants initially approached the judgments as balanced comparisons but then translated them into focused comparisons when recording their responses.

Given this initial evidence, we were curious whether comparisons were necessarily focused, or whether focused comparisons were the result of the way we posed the questions. Therefore, in our second study, we sought to change the way in which we asked people to made comparative judgments. Prior research has shown that such structural changes can alter the pattern of differential weighting and the features of the comparison at hand (e.g., Giladi & Klar, 2002; Moore & Kim, 2003). However, prior manipulations only sought to shift the focus between targets, not to induce more balanced weighting. By presenting respondents with a format more conducive to balanced comparisons, we sought to achieve two aims. First, we could contrast the standard elicitation format, which, as we have shown, produces a discrepancy between the information retrieval and comparison formation stages, with an alternative format that had the potential to exhibit greater consistency between the two stages. Second, we could disentangle the egocentrism and focalism mechanisms associated with single target focusing.

To meet these goals, we chose the most neutral elicitation method of which we could conceive, opting for a pictorial representation without any words other than the names of the two elements. In this format, each element receives equal prominence since neither receives a clear designation as target or referent. Thus respondents can frame comparisons as either balanced or focused as they see fit. In selecting this neutral question format, we thus sought to avoid the well-documented and powerful effects of question wording on their answers (Loftus, 1975; Schwarz, 1999).

As an additional impetus for our second study, we note that our first study employed types of questions that stand in noticeable contrast to the topics conventionally used to examine single target focusing. Many of our factual (e.g., Academy Award nominations) and social (e.g., the number of pairs of blue jeans owned) topics possess the kind of correctness not normally found in other research. In Study 2, we constructed a topic list of comparative questions more representative of the single target focusing literature to ground our work within that prior research.

Study 2

In Study 2, we employed topics used in prior research that has documented differential weighting in comparative judgment. We varied (between subjects) whether participants were asked to make their comparisons using the traditional focused question format or using a more neutral balanced question format. We expected to replicate findings of differential weighting in the comparison formation stage using the traditional focused questions but not using the balanced question formats. We also predicted that the way in which individuals activate information in the retrieval stage will be consistent (i.e., balanced) across response formats.

Method

The experimental method we used for Study 2 corresponded to that used for the first study, except where noted below.

Participants

One hundred twelve individuals (53 female, 59 male; Mean age = 26.0 years, SD = 9.8) from the same paid research pool participated in the study for $6 and the chance to earn additional money during the session.

Topics

We constructed a new set of 28 topics, including several drawn from previous studies that found evidence for differential weighting in comparative judgments. We included eight abilities used by Kruger (1999): four easy (operating a computer mouse, driving a car, riding a bicycle, and saving money) and four hard (programming a computer, juggling, playing chess, and telling a really good joke). We included four quiz topics used by Windschitl, Rose,

Stalkfleet, and Smith (2008): two easy (fast food chains and pop culture) and two hard (world rivers and Baroque music). We included eight events from Kruger and Burrus (2004): four common (living past 70, traveling to Europe, getting in an auto accident, and gaining five pounds in the next 10 years), and four rare (living past 100, traveling to the moon, getting a traffic ticket for driving too slow, and gaining five pounds in the next week). We also added eight neutral likelihood judgments (e.g., "The likelihood that this company will file for bankruptcy in the next 10 years" and "The likelihood that the population of this country will reach 1.5 billion in the next 20 years") in place of the factual topics from the first study.

Response formats

Respondents saw either a focused comparison format or a balanced comparison format. The focused format was replicated from Study 1. In the balanced format, individual elements (A or B) appeared in the center of the screen in either a blue or red colored font. The scale labels for each single element appeared in the same color as the element. The comparative judgment screen featured the "blue" element (B) in the lower left corner and the "red" element (A) in the upper right corner of the screen, with a diagonal black line between the two. We positioned elements on a diagonal so that neither would assume primacy in terms of the left–right and top–bottom orientations and we included the line between elements to visually reinforce when the comparative judgment occurred. The scales for the comparison increased in favor of a given element (e.g., higher numerical values or more extreme word labels) toward the end of the scale where that element was located, so that the scales consisted of a midpoint of zero increasing to the maximum value of the individual scale at both the left and right endpoints. To further reinforce this distinction, the color of the labels on each half of the scale corresponded to the color of the element. Appendix B shows three example screen representations for the three different judgments.

Participants were randomly assigned to one of four conditions that varied in the response scale format (objective vs. subjective) and the comparison format (focused target-referent presentation vs. balanced presentation). We made an exception for the ability judgments from Kruger (1999) since they did not translate well to an interpretation on an objective response scale. All participants answered using the subjective response scale for those topics.

We made one more exception to the procedures from the first study. When participants were asked to compare performances on each of the four trivia quizzes, the elements were always (1) the self and (2) another participant in the study (though we still varied which element was A or B). We did this in order to more closely replicate the Windschitl et al. (2008) procedure.

Additional tasks

Participants completed the four trivia quizzes taken from Windschitl et al. Two quizzes represented relatively easy domains (i.e., pop culture and fast food restaurants) and two represented difficult domains (i.e., world rivers and Baroque music). Each quiz consisted of eight multiple choice items.

Procedure

Like Windschitl et al., our participants completed these quizzes at the start of the experiment. Their instructions stated that they should answer the questions to the best of their ability. After finishing the quizzes, they read a description of the computer judgment task and then completed the judgments for the series of 28 topics.

Results and discussion

Descriptive statistics for the reaction times and judgments across various categories appear in Appendix C. For our statistical analyses, we removed 16 observations whose reaction times fell more than five standard deviations from the mean. This left 3120 topic observations (each with three judgments) from 112 participants.

Comparison facilitation in information retrieval

We conducted this analysis using a linear mixed model controlling for individual participants across the 28 topics and the independent variables of question order, topic category, and comparison format.[4]

These results again show that comparative judgments were facilitated by coming after individual judgments, $F(2,229) = 208.7$, $p < .001$. Pairwise comparisons reveal that participants made comparative judgments faster when they came in the middle ($M = 4.9$ s, $SD = 3.3$) than when they came first ($M = 6.1$ s, $SD = 3.7$), $p < .001$; $d = .40$, and fastest when they came last ($M = 3.7$ s, $SD = 3.0$) instead of in the middle, $p < .001$; $d = .53$.

The order × category interaction attains significance, $F(4,593) = 4.92$, $p = .001$. Again we find that when comparisons appeared first in the series, participants made self-relevant comparisons more quickly ($M = 5.4$ s, $SD = 3.5$) than social comparisons that did not involve the self ($M = 6.5$ s, $SD = 3.7$), $p < .05$; $d = .37$. Participants also made these self-relevant comparisons faster than they made factual comparisons ($M = 6.8$ s, $SD = 3.8$), $p < .001$; $d = .46$.

Again, we tested whether there was evidence in participants' reaction times for the notion that the activation of information at the retrieval stage consisted more heavily of one element than the other. Consistent with Study 1, we find little evidence for such focusing in the retrieval of information. First, when the comparative judgment came in the middle, we find that participants did not make those comparative judgments significantly faster when they came after Element A ($M = 5.1$ s, $SD = 3.4$) than when they came after Element B ($M = 4.8$ s, $SD = 3.1$), $F(1,1022) = 1.79$, $p = .18$; $d = .10$. Second, when Element A followed the comparison, it was not significantly faster ($M = 3.9$ s, $SD = 2.5$) than when Element B followed the comparison ($M = 3.9$ s, $SD = 2.6$), $F(1,991) = .62$, $p = .43$; $d = .03$. This is true for both the balanced and the focused question formats, suggesting that similar information retrieval processes operate regardless of whether the comparative judgment format is focused or balanced.

Differential weighting in judgment formation

This study enables us to test the effect of question format on the weighting of the individual elements in the comparative judgments. To do this, we again conducted regression analyses on the comparative judgments across each domain using the corresponding absolute judgments as independent variables. Results for the focused comparison condition appear in Table 2. While both individual elements significantly predict the comparison of those elements (all p's < .001), the focused question format also produced evidence of differential weighting that is consistent with Study 1 and with previous research (e.g., Chambers et al., 2003; Klar & Giladi, 1997; Kruger, 1999; Windschitl et al., 2008). Across element types (self-related social, non-self social, and factual), the target of the comparison (Element A) was weighted more heavily than the referent (Element B). This suggests that differential weighting is a pervasive feature of this kind of comparison elicitation.

[4] Alternative model specifications that included our manipulation of response scale (subjective vs. objective) produced no significant main effect or interactions.

Table 2
Results of regressions for comparative judgments on Element A and Element B absolute judgments in the focused comparison conditions (Study 2).

Independent variable	Self related social comparisons		Non-self related social comparisons	Factual comparisons
	Self is Element A	Self is Element B		
Element A absolute judgment	0.86***	0.76***	0.85***	0.71***
Element B absolute judgment	−0.45***	−0.63***	−0.61***	−0.46***
Adjusted R^2	0.55	0.39	0.51	0.45

Notes: Table show standardized beta weights. All coefficients are significant at $p < .001$. Asterisks indicate a significant Wald test for differences between the absolute weightings of Element A and Element B coefficients in a given model.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 3
Results of regressions for comparative judgments on Element A and Element B absolute judgments in the balanced comparison conditions (Study 2).

Independent variable	Self related social comparisons		Non-self related social comparisons	Factual comparisons
	Self is Element A	Self is Element B		
Element A absolute judgment	0.83*	0.63	0.69	0.65
Element B absolute judgment	−0.64*	−0.73	−0.63	−0.71
Adjusted R^2	0.59	0.52	0.48	0.58

Notes: Table show standardized beta weights. All coefficients are significant at $p < .001$. Asterisks indicate a significant Wald test for differences between the absolute weightings of Element A and Element B coefficients in a given model.
* $p < .05$.

Conversely, results from the balanced comparison condition display much weaker evidence of differential weighting (see Table 3). Both elements significantly predict the comparison of those elements (all p's < .001). In all but one case (social comparisons involving the self as Element A), the weights of these individual elements are not significantly different from each other. Perhaps it should not be surprising that our balanced question format did not produce differential weighting, since the question does not even specify target and referent designations. Nevertheless, our results suggest that it is at least possible for alternative question formats to attenuate differential weighting effects.[5] This difference in weighting between response formats stands in contrast to the relative uniformity of the reaction time data across formats. While the former suggests that focused comparisons can manifest in the comparative judgment formation stage, the latter suggests that a more balanced process operates during the information retrieval stage.

Study 3

So far, we have contended that the evidence contrary to focused comparison accounts in our reaction time data is a product of the balanced information retrieval that occurs during the initial stage of the comparative judgment process. This facilitation account holds that prior judgments facilitate subsequent judgments insofar as they are relevant. However, we cannot entirely rule out the alternative explanation that our reaction time data reflects learning effects within a given judgment series or across the several topics that participants encountered in our previous studies. We refer to this alternative explanation as the learning account.

For example, we observe faster reaction times for comparative judgments appearing after their individual components. This may result from the facilitating effect of the prior individual components. Alternatively, it is possible that individuals become more familiar with a given topic as they make additional judgments within that topic regardless of the specific elements involved (for example, participants did not need to re-read the topic description and they did not need to recalibrate the response scale). Similarly, we find that judgments for the target element and the referent element are made at an equivalent speed when following the comparison between the two. This again may reflect an equivalent activation of each element in the retrieval stage. However, it also is possible that the multi-topic structure of the task allows participants to anticipate that they will later make individual judgments for the elements of the comparison even if they do not need to activate a certain element for the actual comparative judgment.

We designed Study 3 to rule out these alternative accounts. Participants here completed only one three-judgment series instead of many. Moreover, some participants completed a judgment series in which the comparison consisted of different elements than the two individual judgments.

Method

Participants

Two hundred and fifty-seven individuals completed the study. We recruited participants from Amazon.com's Mechanical Turk (MTurk) website. Our advertisement explained that participants would earn $1 for taking part in a short study in which they would complete a series of judgments.

Design and procedures

Participants saw one of four topics used in Study 2. These included two likelihood judgments (living past the age of seventy and living past the age of one hundred) and two ability ratings (juggling and riding a bicycle). As in our previous studies, participants provided assessments for two individual elements and a comparison between two elements using the focused format (i.e., "Element A compared to Element B"). The individual elements were "You" and "Another participant in this study." However, in Study 3 we varied whether the comparison consisted of the same two individual elements or two additional elements ("Your best friend" and "An average college student"). Participants recorded their answers on equivalent subjective response scales used for the same topics in Study 2. All comparative questions in this study also used the traditional focused format.

[5] We considered the possibility that the balanced method may have failed to produce significant weighting differences because it produced greater error variance. However, error in the balanced method generally was no greater and in many cases was less than that found in the focused condition.

From the MTurk website, participants followed a link to a web-based survey platform. They first read instructions describing the judgment task. They were instructed to complete each judgment as quickly as possible without sacrificing accuracy. Then participants moved onto the series of judgments. The program again timed all responses and randomized the topic, the order of the three judgments (for the two individual judgments and the comparative judgment), the target-referent distinction between elements in the comparison, and the congruency of the elements comprising the comparison (same as the individual judgments or the two different elements).

Results and discussion

We removed data from one participant because the response times fell outside five standard deviations of the mean, leaving 256 observations in the final data set. Descriptive statistics for the reaction times and judgments appear in Appendix D.

Order effects on the speed of comparative judgments

We conducted this analysis as a linear model with the independent variables of order and congruency. These results support the facilitation account rather than the alternative explanation of topic learning. The main effect for order was not significant, $F(2,250) = 0.62$, $p = .54$. However, there was a significant order $\times$ congruency interaction, $F(2,250) = 4.13$, $p < .05$. The sequential position of the comparative judgment mattered little for the response time when it did not include the same elements as the individual judgments ($M = 11.9$ s, $SD = 7.1$ when first and $M = 11.6$ s, $SD = 6.4$ when last), $d = .07$. Conversely, when the comparison elements were the same as the individual judgments, those comparisons were made faster when they came last ($M = 8.1$ s, $SD = 5.0$) rather than first ($M = 10.1$ s, $SD = 6.5$), $d = .46$.

Focused comparisons in responses and reaction times

We performed additional analyses on the sets of responses for which the individual and comparative judgments were congruent (i.e., those consistent with the kinds of judgments made in Study 1 and Study 2). Consistent with our previous findings in Study 1 and Study 2, evidence for focused comparisons is absent in the reaction time data associated with the information retrieval stage of the judgments. Participants did *not* make comparative judgments any faster when they came after the target Element A ($M = 11.3$ s, $SD = 5.3$) than when they came after the referent Element B ($M = 9.4$ s, $SD = 3.8$), $F(1,51) = 1.27$, $p = .27$; $d = .31$. Additionally, when Element A followed the comparison, it was not significantly faster ($M = 7.8$ s, $SD = 3.4$) than when Element B followed the comparison ($M = 10.4$ s, $SD = 9.3$), $F(1,45) = .26$, $p = .61$; $d = .16$.

Patterns of differential weighting in Study 3 were consistent with those found for focused comparisons in the previous studies. Both the target Element A ($B = 0.59$, $p < .001$) and the referent Element B ($B = -0.18$, $p < .05$) significantly predict the comparison between the two (adjusted $r^2 = .27$). However, the target is weighted significantly greater than the referent ($F(1,162) = 23.13$, $p < .001$).

Taken as a whole, the results of Study 3 provide no support for the alternative explanation that the reaction time patterns of the previous studies should be attributed to the learning account. Instead, the results are consistent with our facilitation account that comparisons involve a relatively balanced informational activation and retrieval process, with any focused comparison processes occurring primarily during the formation stage of the comparative judgments.

General discussion

The evidence presented here provides important empirical clarification to previous accounts of focused comparisons. Whereas previous accounts assumed that focused processes operate in both the early (retrieval) and late (formation) stages of comparative judgment, our results suggest that it is the formation stage alone that is likely to exhibit these patterns. During retrieval, information about each element appears equally accessible. This is perhaps most clearly demonstrated in Study 2. The study directly contrasts a standard focused comparison format distinguishing target and referent with a neutral, balanced format that includes no such differentiation between elements in the comparison. As expected, comparisons made using this latter format exhibit much weaker patterns of differential weighting during judgment formation than do those using the focused comparison format. However, these divergences evaporate when we turn attention to the earlier judgment stage of information recruitment, where focused comparisons display striking similarities to their balanced counterparts.

The fact that more balanced comparisons are possible shows that information retrieval need not be focused. We illustrated this with Study 2's balanced presentation format, much as Tversky and Gati (1978) proposed with their non-directional similarity question. Our reaction time results suggest that focus does not characterize the way people retrieve information for comparative judgments, but that it is instead a result of the way they think about the judgment, and that this mind-set can be directed by question wording.

One way to describe our core finding is as a null result: we find no significant differences in response times as a function of our manipulations of target vs. referent. One should always be cautious interpreting null results, because they can arise from so many different causes. Null results can result from small sample sizes, weak manipulations, or noisy dependent measures. A dissociation design, however, increases the interpretability of a null result. A dissociation design measures a dissociation between two dependent variables, one that is affected by the independent manipulation and one that is not. If one variable is strongly affected by the manipulation it helps rule out weak manipulation and small sample size as explanations for the null result on the other variable.

Our design includes exactly such a dissociation. Our manipulation of focusing produces strong differential weighting effects. All our studies replicate differential weighting in comparative judgments. These effects are strong, reliable, and consistent. And they co-exist with a complete lack of evidence for any sort of differential accessibility in response times, providing a clear dissociation between the two dependent measures. The obvious implication is that differential accessibility, as measured by reaction times, cannot account for the strong differential weighting effects in our data.

It can still be the case, however, that the null result comes from a poorly-measured variable while the significant effect comes from a well-measured variable. If this was the cause of our dissociation, then we should not expect to be able to find many significant differences between our reaction time measures. This, however, is obviously not the case. We find considerable differences in reaction times based on the order in which judgments occur. When a specific judgment occurs early in the sequence it is invariably slower than when it occurs late in the sequence, strongly suggesting facilitation by the other relevant judgments.

Our results also shed light on the nature of information storage and retrieval. Not all information is equally accessible in the human mind. The speedier response times for self-relevant information confirm the fact that this information is more accessible than is information either about other people or about other factual (non-social) topics. This was especially true for subjective response scales. This ought to be surprising to anyone who imagines that human memory stores objective facts that are subject

to later interpretations. On the contrary, when we asked our participants to recall objective facts the longer reaction times imply that recall was more laborious than when we asked them only for vague subjective assessments. They were able to tell us more quickly that they scored "well" on the quiz about fast food chains than that they scored "7 out of 8." The impressionistic simplifications most available in memory have to be enriched, supplemented, and reconstructed to provide specific details (Loftus & Loftus, 1980). It will always be easier to recall, for instance, whether we liked or disliked a restaurant than to recall what we ate.

Whither focused retrieval?

While our studies imply that individuals adopt a balanced process during information retrieval, we do not mean to suggest that balanced comparisons are more common or more important than focused comparisons. Although we utilized a wide variety of targets and topics across the three studies, the list is by no means exhaustive. There very well could be additional combinations of topics and targets that produce more focused information retrieval than exhibited in judgments here. Indeed, studies that examine high-stakes choices do suggest that individuals may engage in a comparative process that is focused throughout and predictably succumb to the biases that arise from it (Camerer & Lovallo, 1999). Consistent with this, Moore and Cain (2007) also find that when deciding whether to compete with others, people are more likely to enter the fray when the task is easy than when it is difficult.

Similarly, we do not wish to conflate single target focusing (typically involving one vs. many comparisons) with a more general class of target-referent comparisons spanning a wider variety of forms. Our results show consistent patterns across both types of judgments, but this need not always be the case. In this area there also could be unexplored combinations of topics and targets that exhibit more differentiated results than we found here.

It also is important to distinguish between retrieval of information from memory and the acquisition of new information. While this paper deals exclusively with the activation of information already possessed by the decision maker, focused comparisons may play a vital role in the ways in which individuals search for and acquire knowledge from external sources about the elements in comparative judgments. For example, Radzevick and Moore (2008, Study 3) asked participants to estimate the likelihood that a target football team (*team a*) had beaten its opponent (*team b*) in a recent game. Participants first had the opportunity to view pieces of information about both teams. They chose to view more information about *team a* than about *team b* and looked at information about *a* longer than they did information about *b*. So in terms of information acquisition, individuals acted consistently with focused comparison accounts. Differentiating these types of processes (internal vs. external retrieval) are promising avenues for future research.

Conclusion

There can be little doubt that social comparison matters a great deal. Social comparisons drive which careers people choose, where they choose to live, and how satisfied they are with what they have (Luttmer, 2005; Tesser, 1988). And so it is a big deal if these judgments are systematically biased by focused comparison processes. This research attempts to advance our understanding of the psychological forces at work behind comparative judgments. Where, in the process of arriving at a comparative judgment, do comparisons become focused on the target?

Our results suggest that focusing does not appear to be an inherent feature of information retrieval, but instead has its strongest effect when people are directly forming comparative judgments. We find evidence across three studies of a relatively balanced activation of knowledge in the initial information retrieval stage. One use of this understanding might be to attempt to reduce negative consequences of focused comparisons. The evidence suggests that it makes sense to concentrate on ways in which we can improve the elicitation of comparisons during judgment-formation rather than on earlier information-recruitment phases in the process. Thus comparative judgments can appear more focused or balanced, depending on the particular question people are answering.

Acknowledgments

We are grateful for the facilities and resources provided by the Center for Behavioral Research at Carnegie Mellon University. We also would like to thank participants at the 2009 Society for Judgment and Decision Making Annual Conference and the 2008 Academy of Management Annual Meeting for helpful comments on earlier versions of this paper. Finally, we appreciate the constructive suggestions for strengthening the paper provided by our editor and anonymous reviewers.

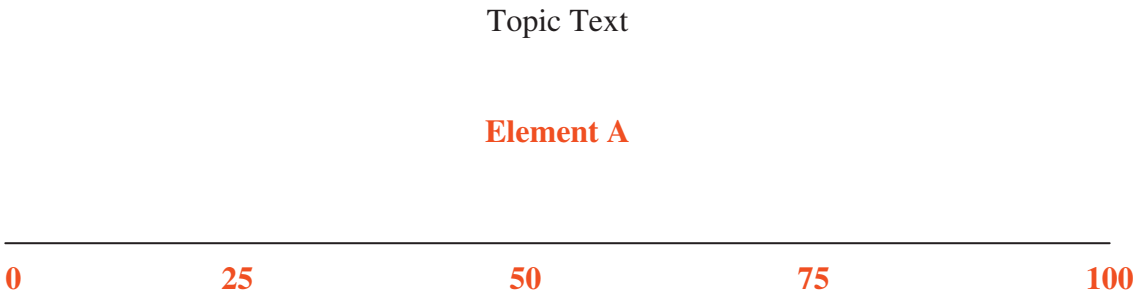
Appendix A

Means and standard deviations (Study 1).

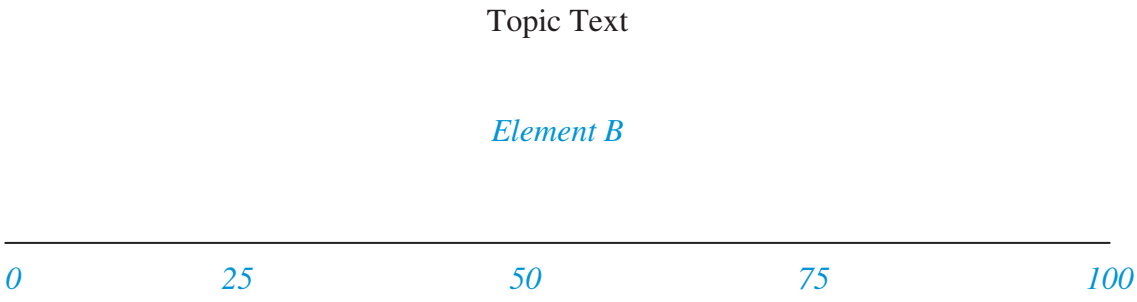
	Element A				Element B				Element C			
	RT		Judgment		RT		Judgment		RT		Judgment	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Factual	3.53	2.93	49.0	23.5	3.51	2.73	48.9	23.0	4.78	3.02	51.3	18.4
Non-self social	3.35	2.84	44.0	27.0	3.37	2.59	44.6	26.4	4.83	3.22	50.1	23.2
Self social (You as target)	3.11	2.58	37.5	30.3	3.10	2.42	43.7	26.9	4.18	2.80	46.6	24.1
Self social (You as referent)	3.36	2.52	43.8	28.1	3.16	2.62	40.5	30.3	4.28	2.95	53.7	22.7
Objective scale	3.72	3.15	41.2	27.1	3.58	2.64	40.8	26.6	5.07	3.23	52.4	18.4
Subjective scale	3.05	2.35	48.4	26.0	3.13	2.60	50.0	24.7	4.20	2.80	48.7	24.3

Appendix B

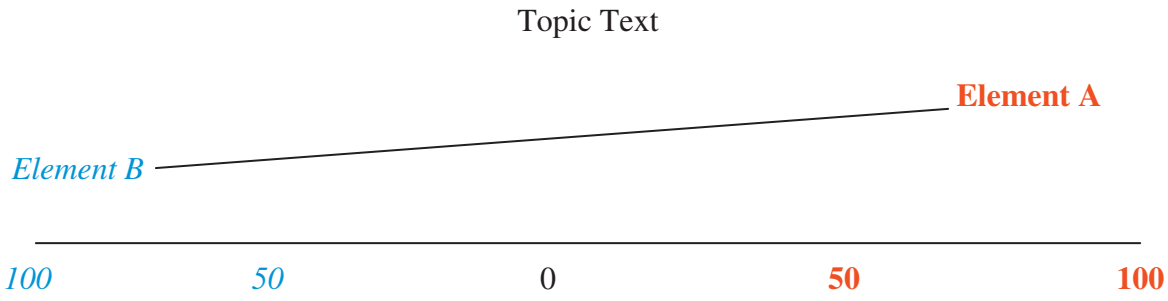
Screen templates for the balanced condition in Study 2.
(Bold font indicates red font and italic font indicates blue font in the actual program).
Element A individual judgment.



Element B individual judgment.



Comparative judgment between Element A and Element B.



Appendix C

Means and standard deviations (Study 2).

	Element A				Element B				Element C			
	RT		Judgment		RT		Judgment		RT		Judgment	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Focused Comp.												
Factual	4.10	4.46	50.8	30.7	4.04	5.87	53.5	29.6	5.55	3.92	50.7	25.7
Non-self social	3.67	2.72	49.4	30.1	3.59	3.07	48.1	30.4	5.50	3.61	49.8	25.1
Self target	3.08	2.29	49.1	32.1	3.71	2.74	49.9	27.1	4.92	3.96	50.1	25.2
Self referent	3.55	2.22	48.7	26.2	3.23	2.53	50.1	32.2	4.61	3.20	49.8	22.4
Objective scale	3.69	3.23	47.0	30.0	3.65	2.82	47.6	30.7	5.07	3.43	50.5	24.7
Subjective scale	3.60	3.13	52.1	29.6	3.67	4.82	53.3	29.0	5.34	3.97	49.8	24.8
Balanced Comp.												
Factual	3.84	3.25	51.9	31.6	3.47	2.46	54.7	29.8	5.41	4.56	50.4	28.1
Non-self social	3.81	3.29	49.0	30.4	3.56	2.86	49.2	30.9	5.14	4.80	49.0	26.1
Self target	2.98	2.24	48.4	33.5	3.48	2.61	48.6	28.8	4.45	4.06	45.9	24.8
Self referent	3.96	3.24	49.7	29.0	3.82	8.63	51.6	33.1	4.65	4.48	51.5	24.8
Objective scale	3.61	2.89	47.3	31.3	3.61	5.67	48.7	30.6	4.73	3.78	48.5	25.4
Subjective scale	3.73	3.28	52.3	30.9	3.53	3.03	53.5	30.6	5.21	5.16	49.9	26.9

Appendix D

Means and standard deviations (Study 3).

	Element A				Element B				Element C			
	RT		Judgment		RT		Judgment		RT		Judgment	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Congruent	8.22	5.20	56.9	28.1	8.40	5.31	55.9	27.1	9.43	5.49	55.5	20.4
Non-congruent	9.41	4.65	53.1	27.6	8.75	7.99	54.8	25.4	11.00	6.26	52.4	20.9

References

Aucote, H. M., & Gold, R. S. (2005). Non-equivalence of direct and indirect measures of unrealistic optimism. *Psychology, Health, & Medicine, 10*(2), 194–201.

Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist, 58*(12), 1019–1027.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*(1), 5–20.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review, 89*(1), 306–318.

Chambers, J. R., & Suls, J. (2007). The role of egocentrism and focalism in the emotion intensity bias. *Journal of Experimental Social Psychology, 43*(4), 618–625.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of nonmotivational factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130*(5).

Chambers, J. R., Windschitl, P. D., & Suls, J. (2003). Egocentrism, event frequency, and comparative optimism: When what happens frequently is "more likely to happen to me". *Personality and Social Psychology Bulletin, 29*(11), 1343–1356.

Covey, J. A., & Davies, A. D. M. (2004). Are people unrealistically optimistic? It depends how you ask them. *British Journal of Health Psychology, 9*(1), 39–49.

Dawes, R. M. (1977). Suppose we measured height with rating scales instead of rulers. *Applied Psychological Measurement, 1*(2), 267–273.

Dunning, D., & Hayes, A. F. (1996). Evidence of egocentric comparison in social judgment. *Journal of Personality and Social Psychology, 71*, 213–229.

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79*(6), 861–875.

Giladi, E. E., & Klar, Y. (2002). When standards are wide of the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. *Journal of Experimental Psychology: General, 131*(4), 538–551.

Gold, R. S. (2007). The link between judgments of comparative risk and own risk: Further evidence. *Psychology, Health, & Medicine, 12*(2), 238–247.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology, 82*(6), 903–918.

Klar, Y., & Giladi, E. E. (1997). No one in my group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology, 73*(5), 885–901.

Klar, Y., & Giladi, E. E. (1999). Are most people happier than their peers, or are they just happy? *Personality and Social Psychology Bulletin, 25*(5), 585–594.

Klein, W. M. P., & Kunda, Z. (1993). Maintaining self-serving social comparisons: Biased reconstruction of one's past behaviors. *Personality and Social Psychology Bulletin, 19*(6), 732–739.

Krizan, Z., & Suls, J. (2008). Losing sight of oneself in the above-average effect: When egocentrism, focalism, and group diffuseness collide. *Journal of Experimental Social Psychology, 44*(4), 929–942.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77*(2), 221–232.

Kruger, J., & Burrus, J. (2004). Egocentrism and focalism in unrealistic optimism (and pessimism). *Journal of Experimental Social Psychology, 40*(3), 332–340.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*, 560–572.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35*(5), 409.

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics, 120*(3), 963–1002.

Moore, D. A. (2004). Myopic prediction, self-destructive secrecy, and the unexpected benefits of revealing final deadlines in negotiation. *Organizational Behavior and Human Decision Processes, 94*(2), 125–139.

Moore, D. A. (2005). Myopic biases in strategic social prediction: Why deadlines put everyone under more pressure than everyone else. *Personality and Social Psychology Bulletin, 31*(5), 668–679.

Moore, D. A. (2007). When good = better than average. *Judgment and Decision Making, 2*(5), 277–291.
Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior & Human Decision Processes, 103*, 197–213.
Moore, D. A., & Kim, T. G. (2003). Myopic social prediction and the solo comparison effect. *Journal of Personality and Social Psychology, 85*(6), 1121–1135.
Moore, D. A., Oesch, J. M., & Zietsma, C. (2007). What competition? Myopic self-focus in market entry decisions. *Organization Science, 18*(3), 440–454.
Moore, D. A., & Small, D. A. (2007). Error and bias in comparative social judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology, 92*(6), 972–989.
Mussweiler, T., & Bodenhausen, G. V. (2002). I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology, 82*(1), 19–32.
Radzevick, J. R., & Moore, D. A. (2008). Myopic biases in competitions. *Organizational Behavior & Human Decision Processes, 107*(2), 206–218.
Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54*(2), 93–105.
Tesser, A. (1988). Toward a self evaluation maintenance model of social behavior. In L. Berkowitz (Ed.). *Advances in experimental social psychology* (Vol. 21, pp. 181–227). New York: Guilford.
Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327–352.
Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categoriztion* (pp. 79–98). Hillsdale, NJ: Erlbaum.
Windschitl, P. D., Kruger, J., & Simms, E. (2003). The influence of egocentrism and focalism on people's optimism in competitions: When what affects us equally affects me more. *Journal of Personality and Social Psychology, 85*(3), 389–408.
Windschitl, P. D., Rose, J. P., Stalkfleet, M. T., & Smith, A. R. (2008). Are people excessive or judicious in their egocentrism? A modeling approach to understanding bias and accuracy in people's optimism within competitive contexts. *Journal of Personality and Social Psychology, 95*(2), 253–273.
Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied, 2*(4), 343–364.