Journal of Personality and Social Psychology
2003, Vol. 85, No. 3, 467–481

0022-3514/03/$12.00 DOI: 10.1037/0022-3514.85.3.467

What Friends Are For! The Use of Routine Standards in Social Comparison

Thomas Mussweiler and Katja Rüter
Universität Würzburg

Social comparisons are frequently used for self-evaluation. As a consequence, judges have to be efficient in each comparison step. Standard selection, however, is traditionally seen as an elaborate process in which judges deliberately select similar standards. We propose that often judges do not engage in such deliberate selection processes. Instead, they use routine standards—standards that have frequently been used for self-evaluation. Consistent with this assumption, Studies 1–3 demonstrate that student participants activate knowledge about their best friend—a likely routine standard—during self-evaluation. In Study 1, lexical decisions for the best friend's name are facilitated after self-evaluation. In Study 2, participants judge their best friend faster after evaluating themselves on the same dimension. In Study 3, this is even the case if the best friend is dissimilar and consequently undiagnostic. Finally, in Study 4, self-evaluations are primarily influenced by comparison information about an experimentally created routine standard.

Life is much easier for those who know themselves. Without appropriate self-knowledge, the decisions we need to make during our daily routines—from the most trivial to the most existential—become difficult to master. Ordering the right pizza, selecting a matching partner, and choosing the right career path all require that we know who we are, what we like, and what we can do. Given this existential importance, it is little surprising that people have a keen interest in obtaining self-knowledge (Festinger, 1954). In fact, this interest may even be pursued if obtaining self-knowledge comes at ego-deflating costs: Oftentimes people seek self-knowledge even if it has the potential to shed unfavorable light on them (for a review, see Trope, 1986). The interest in learning who we are is also apparent in a striking egocentricity of our thoughts: Self-reflective thoughts make up a large portion of our mental activity. In one study (Csikszentmihalyi & Figurski, 1982), for example, about 8% of all thoughts participants had in the course of the day were found to pertain to the self. Thus, people appear to spend a considerable chunk of their time reflecting on themselves, trying to figure out who they are.

How do people obtain this apparently precious self-knowledge? How do they learn who they are? Social comparisons (Festinger, 1954) are one primary source of self-knowledge. Although temporal comparisons are also a viable source of self-knowledge (Wilson & Ross, 2000), our standing on those dimensions that are most indicative of our social, romantic, and economic welfare are often determined in comparison with others. To say that one is smart, funny, or good-looking is to say that one is smarter, funnier, and better looking than others (Huttenlocher & Higgins, 1971). In fact, social comparisons play so prominent a role in self-evaluation that they are even sought in situations in which objective standards are readily available (Klein, 1997). Thus, to understand who we are, we often focus on what other people are.

Given that self-evaluative thoughts—and thus social comparison activities—occupy such a sizeable portion of our mental life, it is essential for the social judge to be efficient in making social comparisons. Achieving a thorough understanding of who we are has little benefit if it is too much of a burden for our limited cognitive resources. Finding out that one has outstanding skills as a computer programmer, for example, has little adaptive value if doing so requires so much cognitive capacity that none is left for programming. Similar to most social arenas, it is thus essential that we also act as cognitive misers (Taylor, 1981) in our social comparison activities. How is this required efficiency achieved? For a process that is engaged as often as social comparison, proceduralization and automatization (Bargh, 1997; Smith, 1994) are likely to be important capacity-saving devices. As is true for any other psychological process that is carried out repeatedly, comparison processes are likely to become proceduralized so that they can be carried out in relatively automatic ways that require little cognitive capacity (Bargh, 1997; Smith, 1994). In fact, given that social comparisons are typically involved in self-evaluation, they are carried out so frequently that proceduralization is particularly likely.

This suggests that comparisons may be relatively automatic processes that require little cognitive capacity. Consistent with this assumption, it has been demonstrated that comparisons are indeed often spontaneous and effortless (Dunning & Hayes, 1996; Gilbert, Giesler, & Morris, 1995). To evaluate another person, for example, judges appear to spontaneously compare this person with the self (Dunning & Hayes, 1996). Apparently, social comparisons are

The present research was supported by a grant from the German Science Foundation. We thank Kai Epstude and the other members of the Emmy Noether Research Group on Comparison Processes and the Würzburg Social Cognition Group for stimulating discussions of this work. We are particularly indebted to Lysann Damisch, Tanja Hundhammer, Yvonne Moller, Martina Walter, Romy Weiland, and Sebastian Werner for their help in preparing the stimulus materials and collecting the data.

Correspondence concerning this article should be addressed to either Thomas Mussweiler or Katja Rüter, Psychologie II, Universität Würzburg, Röntgenring 10, Würzburg 97070, Germany. E-mail: mussweiler@psychologie.uni-wuerzburg.de or rueter@psychologie.uni-wuerzburg.de

often so natural and effortless that they are even engaged if the comparison standard is clearly irrelevant for self-evaluation (Gilbert et al., 1995). In one study, for example, participants compared their performance in detecting schizophrenia in photographed target persons with an inferior standard even if they knew that this standard had been deliberately misled to make wrong judgments. This comparison with the clearly nondiagnostic standard was primarily apparent if participants' cognitive resources were depleted by rehearsing an eight-digit number (Gilbert et al., 1995). These findings suggest that comparisons with others are indeed relatively automatic processes that require little cognitive capacity. In fact, social comparisons may often be so effortless that they are even engaged when normatively inappropriate, so that social judges then have to allocate extra resources to correct for their consequences (for a more detailed discussion, see Gilbert et al., 1995).

The Selection of Social Comparison Standards

Social comparison may thus often take the form of a fairly automatic and highly efficient process that requires little cognitive capacity and can easily be carried out under suboptimal conditions. We have argued that in light of the ubiquity of social comparison activities, their automatic and capacity-saving qualities are necessary to ensure our psychological functioning. If we do indeed constantly engage in social comparisons, then we cannot afford to allocate too many resources to them. This efficiency criterion equally applies to all stages of the comparison process. To use social comparisons efficiently, judges must not only be able to quickly compare themselves with a specific standard, they must also be efficient in selecting such a standard in the first place. Which standards do people select for comparison? What are the psychological mechanisms that underlie the selection process? And, how do these mechanisms relate to the important criterion of the efficiency of social comparison?

The answer that social comparison theory and research offers to these important questions seems simple at first glance. In particular, social comparison theory prescribes that to obtain valuable and diagnostic information about themselves, judges have to compare with others who are similar on the critical dimension itself (Festinger, 1954) or on attributes that are related to this dimension (Goethals & Darley, 1977; for a discussion of the limitations of both notions, see Miller & Prentice, 1996). This is assumed to be the case because comparisons with dissimilar others may not have clear implications for self-evaluation. If, in an attempt to evaluate your athletic abilities, you compare with your 80-year-old grand aunt, for example, your potentially superior performance can be easily attributed to the fact that both of you differ on a dimension that is closely related to athletic abilities (i.e., age). As a consequence, it is unclear whether the comparison outcome is due to your personal qualities or to the difference on an important related attribute so that it is of limited diagnostic value. These normative perspectives on standard selection are supported by considerable empirical evidence. In fact, both types of similarity—similarity on a critical dimension and similarity on related attributes—appear to influence with whom people compare (e.g., Gruder, 1971; Suls, Gastorf, & Lawhon, 1978; Wheeler, 1966; Zanna, Goethals, & Hill, 1975).

The diagnostic advantages of selecting similar others as comparison standards, however, are likely to come at a distinct cost. As is true for human judgment in other domains, obtaining optimal information requires a lot of effort and cognitive resources (Simon, 1956; Tversky & Kahneman, 1974). Finding a standard who is similar on the critical dimension or on related attributes is an elaborate process in which different dimensions, different potential standards, and different criteria have to be considered. Oftentimes, there seem to be too many choices and too little time. Imagine, for example, that in your attempt to evaluate your athletic abilities you had opted against using your grand aunt as a comparison standard. With whom should you compare? With your spouse, who is similar to you on many important attributes related to athletic abilities (e.g., age, nutrition, living conditions) but also differs on the important gender attribute? Or with your tennis partner who is of the same gender but considerably younger? Or with your 10-year-old niece who is even younger but seems to be a perfect match for your tennis abilities? As is apparent from this example, selecting a similar standard for comparison is an arduous task, and consequently one that we can rarely afford to engage in. Rather, as is true in other judgment domains, we are likely to "satisfice" (Simon, 1956) rather than optimize by using a standard that can be selected in an efficient manner.

Routine Standards

How is this objective achieved? What type of comparison standard can we select efficiently? One possibility is that people follow routines in selecting social comparison standards. Applying routines is one of the major ways by which judges simplify complex decisions and consequently save cognitive resources (e.g., Aarts & Dijksterhuis, 2000; Betsch, Haberstroh, Glöckner, Haar, & Fiedler, 2001; Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994; for an overview, see Verplanken & Aarts, 1999). Rather than deliberately deciding whether to walk, ride your bike, or drive your car to work by weighing all relevant information (e.g., weather conditions, weight to carry, distance) every single morning, for example, you may simply decide to do what you always do. Following this routine is likely to free cognitive resources for other pressing issues (e.g., deciding what to teach in class). This heuristic advantage of routine application is apparent in the fact that people resort more readily to their decision routines under suboptimal conditions (Betsch, Fiedler, & Brinkmann, 1998). Thus, to reduce the complexity of our daily lives, we often apply well-practiced routines.

In much the same way, we may also resort to routine comparison standards to simplify the complex task of standard selection. Instead of engaging in the arduous and often impossible task of finding the most diagnostic standard for comparison, we may simply compare with those standards we routinely use for comparison. Rather than agonizing about whether to select your grand aunt, your partner, your tennis partner, or your 10-year-old niece as a comparison standard for evaluating your athletic abilities, you may simply compare with your partner, because this is the person with whom you most often compare. This suggests that social comparisons may be shaped by processes of procedural priming (Smith, 1994). The more often a particular comparison is carried out, the more accessible and efficient the process of relating one's

attributes to this standard becomes, and the more likely one is to further engage in this specific comparison.

Such a procedural priming perspective on social comparison suggests a number of potential advantages of routine standard use in social comparison. To the extent that judges simply apply the best practiced procedure for self-evaluation and routinely compare with the same people, they may not have to engage in an elaborate standard selection process. Using routine standards may even allow us to skip the process of standard selection altogether. If we routinely compare with the same people, then there is not much selecting we need to do. Furthermore, comparison-relevant information may be readily accessible because it has already been used in previous comparisons so that it does not have to be searched or constructed. Finally, the comparison process of relating this standard knowledge to self-knowledge may be particularly efficient because it was extensively practiced in previous comparisons (Smith, 1989, 1994). In sum, the use of routine standards has the potential to simplify social comparison in at least three important ways: It may make standard selection easier or even unnecessary. It may ensure that the referent information about the standard that is needed to carry out the comparison is readily accessible. Finally, it may rely on a well-practiced and highly efficient comparison process. In the present research, we examine the use of routine standards in self-evaluative judgments.

The Present Research

In principle, there are two research strategies that can be used to examine the hypothesized use of routine standards in social comparison. The two strategies differ with respect to the nature of the potential routine standard. A first possibility is to examine the use of those actual comparison standards that are likely to be routine standards for a particular participant population. One may, for example, examine how participants use their partners, siblings, colleagues, and friends for comparison purposes. A second possibility is to experimentally create routine standards by engaging participants in repeated comparisons with a previously unknown fictitious person. The procedural priming perspective we have outlined holds that any standard that is repeatedly used for comparison may become a routine standard. If this is the case, then routine standards can be experimentally created by inducing judges to engage in a series of comparisons with the same standard. Clearly, both strategies have unique advantages. Examining the use of actual routine standards, on the one hand, is closer to the daily comparison activities people actually engage in. Examining the use of experimentally created routine standards, on the other hand, allows to control for potential differences between standards that may otherwise influence comparison activities. In this research, we applied both strategies.

In particular, we will initially focus on actually existing comparison standards and examine the use of the potential routine standard that is likely to be most important for our participant population of college undergraduates. Although a host of potential routine standards such as siblings, partners, and colleagues exist, their relative importance is likely to vary over the course of life. For college students, their best friends are likely to play a particularly prominent role. Because the initial college years constitute an important life transition (Compas, Wagner, & Slavin, 1986), they are a time in which peers with similar experiences become particularly important for self-evaluation (Hays & Oxley, 1986; Hirsch, 1980). Students' best friends in college are thus likely to be often used for social comparison and may consequently become important routine standards. Despite our initial focus on the use of best friends as comparison standards, the question we are examining is quite general in scope. We are not concerned about the specific role friends play in social comparison. Rather, we are interested in the role routine standards in general assume, and merely use best friends as one class of potential routine standards that is likely to be particularly important for our participant population.

To test whether people use routine standards when evaluating themselves, we first examine whether college students spontaneously activate information about their best friends during self-evaluation. If best friends are indeed used as routine comparison standards, as our reasoning suggests, then information about their best friends should be more accessible subsequent to a self-evaluative judgment. In the first study, we set out to demonstrate this increased accessibility with the help of a lexical decision task. Participants are instructed to evaluate themselves on a number of personality dimensions. If they do indeed think about and compare with their best friends in the course of these self-evaluative judgments, then the best friend's name should be made accessible so that it is more easily recognized in a subsequent lexical decision task. In the second study, we extend this reasoning and attempt to demonstrate that it is not only the potential routine standard's name that is rendered accessible, but knowledge about his or her standing on the critical dimension of self-evaluation. If in evaluating themselves our participants do indeed think about their friend's standing on the judgmental dimension, then they should subsequently be faster in evaluating their best friend on the critical dimension. Study 3 will then test for potential limits of the use of routine standards in self-evaluation by contrasting the use of routine standards with the use of those standards that have traditionally been assumed to be most influential in social comparison, namely standards who are similar to the self (Festinger, 1954; Goethals & Darley, 1977). Will participants still use routine standards for comparison if they are clearly dissimilar on the critical dimension and if a similar standard is available? Our third study examines this question concerning the ubiquity of routine standard use.

Our final study will then apply the second principle research strategy and examine the use of experimentally created routine standards in social comparison. If our reasoning is correct, then judges should tend to select those standards for comparison with whom they frequently compared themselves prior to self-evaluation. Study 4 uses the judgmental consequences of comparison (i.e., assimilation) as a diagnostic tool to test for this assumed dependency.

Study 1

In our first experiment we set out to demonstrate that in evaluating themselves, our participants do indeed think about a potential routine standard. In particular, we examined whether college student participants think about their best friends during self-evaluation. If this is indeed the case, then participants' best friend should be easily accessible subsequent to the comparison. Because thinking about a particular concept during a task increases the

accessibility of this concept (Higgins, 1996), thinking about one's best friend during self-evaluation should put this person at the top of participants' minds. To demonstrate this expected increase in the accessibility of participants' best friend, we used the classic method to examine knowledge accessibility effects, namely a lexical decision task (e.g., Neely, 1977). In this task, participants are confronted with a series of letter strings for which they have to indicate whether they do or do not constitute actually existing words. Previous research (for an overview, see Neely, 1991) has demonstrated that the more accessible a given concept is, the faster judges are in making related lexical decisions. Thus, participants should be faster in responding to the name of their best friend if its accessibility has been increased by its use in a preceding self-evaluation task. This dependency allows response latencies for lexical decisions to be used as diagnostic probes for the activation of standards during a preceding judgment. If judges are faster in responding to the name of their best friend subsequent to evaluating themselves on a given dimension, then this suggests that they have activated this standard during self-evaluation.

In Study 1, we had student participants either evaluate themselves or another person with respect to a number of characteristics (e.g., cheerfulness, sentimentality). Evaluating themselves should induce them to think about and compare with their best friend as a likely routine standard. Evaluating another person, however, should induce them to compare this target with the self, because the self is typically used as a standard in judgments about other people (Dunning & Hayes, 1996). As a consequence, the best friend as participants' likely routine standard for self-evaluation should only be rendered accessible in the first case. If this is indeed the case, then participants should be relatively faster in identifying the name of their best friend in a lexical decision task that follows self-evaluation. This, however, should not be the case subsequent to evaluating another person.

Method

Participants. Thirty-six students at the University of Würzburg, Würzburg, Germany, participated in Study 1. We embedded this study in a series of unrelated experiments during a 1-hr session. Participants were randomly assigned to one of four experimental conditions and were offered DM 12 (approximately $6 at the time) as compensation.

Materials and procedure. Upon arrival in the laboratory, participants were greeted by the experimenter and seated in front of personal computers with 17-in. monitors. In the instructions, we explained that during the next hour participants would take part in several unrelated experiments, some of them conducted at the computer and some with paper and pencil. In what was introduced as the first experiment, we collected names of participants' same-sex friends and family members ostensibly to examine participants' perceptions of their same-sex social network. A total of up to 18 names was collected, among them the two critical ones that served as target stimuli for the lexical decision task: participants' current best friend and a good school friend with whom they had lost contact (exfriend). We decided to use the exfriend as a control because such a former friend shares many of the qualities of one's current best friend. In particular, one's current and former friends seem good matches in terms of similarity to the self, amount of available information, and liking. To keep up the cover story, we subsequently asked participants to describe and rate themselves and the members of their social network on several dimensions (e.g., *höflich* [polite], *durchsetzungsfähig* [self-assertive], *attraktiv* [attractive]). Following this first part of the study, participants worked on a number of other, unrelated experiments, which lasted for a total of about 20 min.

The final experiment of the session was the critical one in which we assessed the accessibility of participants' routine standards after self-evaluation. To do so, we first informed participants about the ostensible purpose of this final study: examining how thinking about a person influences people's ability to concentrate. Toward this end, participants should first think about and evaluate a single person on several dimensions. We would then test their ability to concentrate with a reaction-time task.

Approximately half of the participants evaluated themselves on 10 personal attributes (e.g., *Leidenschaft* [passion], *Fröhlichkeit* [cheerfulness], *Empfindsamkeit* [sentimentality]). Our reasoning suggests that these self-evaluations induce participants to activate and compare with their routine standard. The other half evaluated a well-known same-sex celebrity (tennis players Steffi Graf or Boris Becker) on the same 10 attributes. These participants should not activate their routine standard during the evaluation. All evaluations were given along 9-point rating scales ranging from 1 (*very little*) to 9 (*very much*).

Subsequent to the evaluation task, we assessed the accessibility of the routine standard with a lexical decision task. In the cover story we introduced this task as a measure of participants' ability to concentrate. To minimize the time gap between the evaluation and the lexical decision task, the procedure of the latter was already explained to participants together with the evaluation task. This allowed participants to start working on the lexical decision task immediately after completing the evaluation task. Congruent with the typical procedure in the lexical decision paradigm (Neely, 1991), participants' task was to decide as fast as possible whether a letter string does or does not constitute a name. We instructed participants to put their right and left index fingers on the response keys before the first trial, and to keep that position throughout the task. The stimuli included the name of the current best friend (routine standard) and the name of the good school friend to whom participants had lost contact (exfriend). If participants used their best friend as a routine standard in the preceding evaluation task, then they should be able to identify her or his name faster than the name of the exfriend.

All lexical decision trials followed the same sequence. First a fixation point ("X") appeared in the middle of the screen for one second. It was immediately followed by the letter string. Participants responded to the letter string by pressing either the *D* key with the left index finger or the *Ö* key (the *semicolon* key on an American keyboard) with the right index finger. We counterbalanced the assignment of the keys to the respective responses. The assignment was indicated at the bottom left and right corners of the screen during the whole task. The letter string remained on the screen until participants had given their answers.

Altogether participants had to react to 15 letter strings. Five of those were same-sex names, five opposite-sex names, and five no names. The same-sex names were exclusively names of participants' personal social network. The critical trials appeared in the fourth and the seventh positions. Whereas for half of the participants the name of the best friend was presented first and the name of the exfriend second, this order was reversed for the other half. We randomized the presentation of the remaining stimuli with the limitation that no other same-sex name appeared before the two critical trials. After completing the lexical decision task, participants were thanked, fully debriefed, paid, and dismissed.

In sum, Study 1 is based on a 2 (evaluation: self vs. other) $\times$ 2 (lexical decision: best friend vs. exfriend) mixed-factorial design. Evaluation was varied between participants. Half of the participants evaluated themselves, whereas the other half evaluated a celebrity (tennis player). Subsequently, all participants had to identify the name of their best friend and the name of their exfriend in the lexical decision task. Hence, this variable was varied within participants.

Results

Our reasoning suggests that participants think of their best friends when evaluating themselves. Doing so will render their

best friend's name more accessible so that it is recognized faster in the lexical decision task. If participants do indeed use their routine standards as predicted, then they should be able to identify the name of their best friend faster than the name of their exfriend after evaluating the self. This, however, should not be the case after evaluating another person.

Following the suggestions of Fazio (1990), we performed logarithmic transformations (ln) of the response latencies to reduce the skewedness of the response distribution. Our analyses are based on these transformed latencies. For ease of interpretation, however, we report the untransformed means (in milliseconds).

As is apparent from Figure 1, participants did indeed react faster to the name of their best friend ($M = 576$ ms) than to the name of their exfriend ($M = 642$ ms) if they had previously evaluated themselves, $t(14) = 2.86$, $p < .01$ (two-tailed). This, however, was not the case if participants had evaluated another person (the tennis player). In this case, they did not identify the name of their best friend ($M = 680$ ms) any faster than the name of their exfriend ($M = 670$ ms), $t(20) = -0.24$, *ns*. This pattern of means produced a significant interaction effect in a 2 (evaluation: self vs. other) $\times$ 2 (lexical decision: best friend vs. exfriend) mixed-model analysis of variance (ANOVA) using the logarithmic transformations of response latencies in the lexical decision task as the dependent measure, $F(1, 34) = 4.13$, $p < .05$, $d = 0.70$. None of the main effects proved to be significant in this analysis (Fs < 2.86, ps $> .10$).

Discussion

These findings provide initial support for the assumption that self-evaluative judgments involve the activation and use of routine standards. As we have expected, participants were faster in responding to the name of their best friend than to the name of a former friend subsequent to self-evaluation. This, however, was not the case after participants had evaluated another person instead of themselves. Because faster lexical decisions typically indicate higher levels of accessibility, these findings demonstrate that evaluating oneself does indeed put one's best friend on the top of one's mind. This, in turn, is consistent with our assumption that participants think about and compare with their best friend during self-evaluation. Of importance, the present data indicate that the activation of the routine standard was specific to self-evaluation. Lexical decisions for the best friend's name were only facilitated by preceding evaluations pertaining to the self, not by evaluations pertaining to another target person. Our best friends do thus not appear to be used invariably in social judgment, but specifically for evaluations of the self.

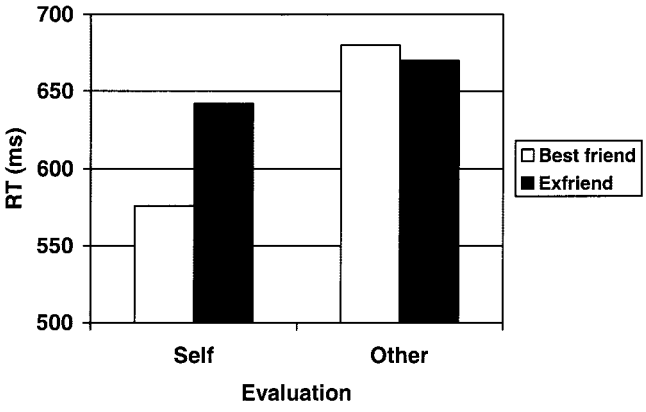

Figure 1. Response latencies for lexical decisions (best friend vs. exfriend) subsequent to evaluation of self versus other (Study 1). RT = reaction time.

Our second study attempts to extend these findings by looking more specifically at the activation of comparison-relevant information during self-evaluation. If one's best friend is indeed used for comparison, then judges have to activate information about his or her standing on the judgmental dimension during self-evaluation. If, for example, participants think of and compare with their best friend in the course of evaluating how cheerful they are, then they should activate information regarding their best friend's cheerfulness. Study 2 was designed to demonstrate this activation of comparison-relevant information about the routine standard.

Study 2

To do so, we asked our participants to judge their best friend or a former friend on the exact same dimension on which they had previously evaluated themselves. For example, participants first evaluated their own punctuality and then judged the punctuality of either their best friend or a former friend. Of importance, the time participants need to evaluate the other person depends on the accessibility of the knowledge they need to make this judgment (e.g., Strack & Mussweiler, 1997). Thus, if judges have already activated this knowledge in the preceding self-evaluation, then they can readily use this knowledge when judging the other person. As a consequence, participants are faster in making this judgment. In this respect, the time judges need to judge the other person can be used as a diagnostic probe indicating the extent to which they have activated judgment-relevant knowledge about this person during self-evaluation. Just as faster judgments about the self subsequent to judging another person indicate that self-knowledge regarding the critical characteristic was activated (Dunning & Hayes, 1996), faster judgments about another person subsequent to judging the self indicate that knowledge about this person has been activated during self-evaluation. If, for example, participants do indeed activate knowledge about their best friend's punctuality while evaluating their own punctuality, then they should subsequently be faster in judging how punctual their best friend is.

To demonstrate this dependency, we had our participants first think about and evaluate either themselves or another person, and then judge either their best friend or a former friend with respect to a number of different characteristics. If our reasoning is correct, then judgments about the best friend but not about a former friend should be facilitated by preceding evaluations of the self but not of the other person.

Method

Participants. Thirty-three students at the University of Würzburg participated in Study 2. As in Study 1, the experiment was part of a 1-hr session for which participants were offered DM 12 (approximately $6 at the time).

Materials and procedure. Upon arrival in the laboratory, participants were greeted by the experimenter and seated in front of personal computers with 17-in. monitors. In this experiment we used the same cover story as

in Study 1. Participants were made to believe that we were interested in their perception of their same-sex social network. In the first part of the experiment, we again asked them to name several of their friends and family members. As in Study 1, participants provided the names of their current best friend and of another friend with whom they had lost contact (exfriend 1). In addition, we asked for the name of a second former friend (exfriend 2).

In the second part of the experiment participants then worked on a series of evaluation and judgment tasks. We informed them that judging a person can take different forms depending on whether the judgment is made rather thoroughly or more spontaneously. Following this distinction, they were instructed to alternately work on two tasks that differ with respect to the elaborateness of processing they require. In the evaluation task, participants were instructed to first allow themselves 1 min to think deeply about the characteristics of one person (e.g., How polite is your friend Paul?), and then describe their thoughts and conclusions in writing. Hence, this task induced participants to engage in elaborate evaluative processes about one person. The subsequent judgment task then required participants to quickly judge two persons with respect to the same characteristic used in the evaluation task. We emphasized that participants should report these latter judgments as fast as possible. To allow them to do so, we asked them to position their fingers on the number keys in the top row of the keyboard with which answers were to be reported. The questions (e.g., How polite is your friend Peter?) were presented on the computer screen together with the 6-point scale ranging from 1 (*very little*) to 6 (*very much*) along which the judgments were made. Each of the individual trials of the judgment task followed the same sequence. Participants first judged the target person of the evaluation task with respect to the critical dimension (e.g., How polite is your friend Paul?). This was followed by a delay of 1,000 ms. Subsequently, they judged a second person along the same dimension (e.g., How polite is your friend Peter?). After reporting this second judgment, participants proceeded with the next evaluation task, which pertained to another dimension.

Altogether, this second part of the experiment consisted of 10 trials. In each trial, participants first evaluated and described a person concerning a personal characteristic, then positioned their fingers on the number keys and finally judged the previously described person and a second person with respect to the same characteristic. We manipulated the independent variable evaluation of self versus other (exfriend 2) in the evaluation task and the independent variable judgment of best friend (routine standard) versus exfriend 1 in the judgment task. Study 2 is thus based on a 2 (evaluation: self vs. other) $\times$ 2 (judgment: best friend vs. exfriend 1) experimental design. Both factors were manipulated within participants.

The four conditions were realized in four different trials that were presented at the fifth, sixth, eighth, and ninth positions. The first four trials took approximately 12 min to complete, therefore a substantial delay resulted between the first part of the experiment (naming the friends) and the critical evaluation and judgment tasks. The critical trials pertained to the characteristics *Pünktlichkeit* [punctuality], *Sportlichkeit* [sportiveness], *Ehrgeiz* [ambition], and *Ordentlichkeit* [tidiness]. Whereas the order of the dimensions was the same for all participants, the conditions were counterbalanced. For example, a quarter of the participants thought about their own punctuality and subsequently rated their best friend on that dimension. Another quarter first thought about how punctual their exfriend 2 is and subsequently rated their exfriend 1. In the noncritical trials, participants thought about and rated other members of their same-sex social network.

After completing the second part of the experiment, participants continued with other, unrelated studies. Upon completion of all tasks, participants were thanked for their participation, fully debriefed, paid, and dismissed.

Results

Evaluations. To ensure that the two critical standards—participants' best friend and their exfriend 1—do not differ with respect to the variable that is traditionally seen as the most important determinant of standard selection, namely similarity to the self (Festinger, 1954; Goethals & Darley, 1977), we analyzed participants' ratings. Specifically, we calculated difference scores for participants' self-evaluation and their evaluation of their best friend versus their exfriend 1. The dimensions on which those differences were calculated varied between subjects because of the counterbalancing described above. If participants saw themselves as more similar to one of the critical standards, then this should be apparent in evaluations that are closer to one another and consequently in a lower difference score. The mean difference between evaluations of the self and the best friend ($M = 1.45$), however, was similar to that between evaluations of the self and the exfriend 1 ($M = 1.55$), $t(32) = 0.33$, *ns*. Thus, there is no reason to assume that the best friend was seen as more similar to the self than the exfriend 1.

Response latencies. Our central dependent measures are the response latencies for evaluations of participants' best friend and their exfriend 1, respectively. A preliminary inspection of the untransformed response latencies revealed that 5 out of the total of 132 responses ($< 4\%$) were clear outliers. Each of these responses was at least 2,000 ms slower than the slowest of the remaining responses and all took longer than 8,000 ms. We excluded these outliers from our analyses, and we were left with a total of 127 response latencies ($M = 2{,}959$ ms, $SD = 1{,}412$ ms) or 29 participants for the within-participants analysis. We conducted logarithmic transformations (ln) of the remaining response latencies prior to analysis but nevertheless present the untransformed means (in milliseconds) for ease of interpretation.

Similar to our predictions in Study 1, we expected participants to be faster in judging their best friend than their exfriend 1 subsequent to a self-evaluation. The judgment of the best friend should be facilitated if participants have used their best friend as a routine standard. As is apparent from Figure 2, the data are consistent with this hypothesis. Participants were faster ($M = 2{,}556$ ms) in judging their best friend than their exfriend 1 ($M = 2{,}922$ ms) after evaluating themselves on the same dimension, $t(28) = 2.17$, $p < .04$ (two-tailed). This, however, was not the case after judging exfriend 2. In fact, judging exfriend 2 did not facilitate the judgment of the best friend relative to the exfriend 1. Rather, the reverse was the case. After judging exfriend 2, partic-

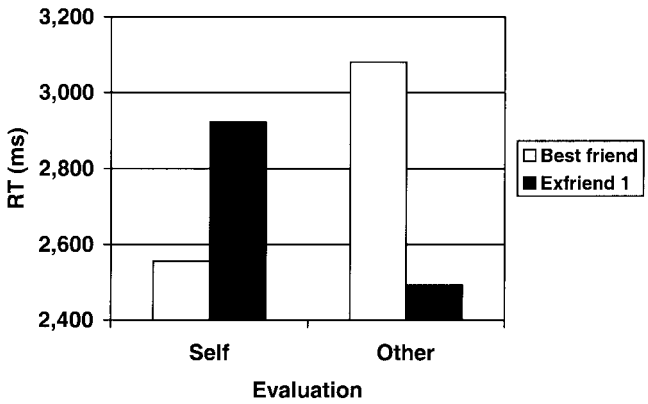

Figure 2. Response latencies for judgments of best friend versus exfriend 1 subsequent to evaluation of self versus other (Study 2). RT = reaction time.

ipants rated exfriend 1 ($M = 2{,}493$ ms) faster than their best friend ($M = 3{,}081$ ms), $t(28) = -2.96$, $p < .01$ (two-tailed). This may be the case because exfriend 1 and exfriend 2 are both former school friends who are likely to share many features. As a consequence, in evaluating the one, participants may think about the other and use him or her as a comparison standard. This pattern of means was borne out in a significant interaction effect in a 2 (evaluation: self vs. other) × 2 (judgment: best friend vs. exfriend 1) within-participants ANOVA using the logarithmic transformations of response latencies in the judgment task as the dependent measure, $F(1, 28) = 13.86$, $p < .001$, $d = 1.41$. None of the main effects approached significance in this analysis (Fs < 1).

To statistically control for potential differences in perceived similarity to the critical standards, we conducted an analysis of covariance using the difference between the two difference scores for evaluations of self–best friend and evaluations of self–exfriend 1 as a covariate. In this analysis, the critical interaction remained unchanged, $F(1, 27) = 14.02$, $p < .001$, $d = 1.44$.

Discussion

Consistent with our hypotheses, the results of Study 2 demonstrate that in the course of evaluating themselves, participants activated knowledge about their best friend's characteristics in the critical domain. Because doing so rendered this knowledge readily accessible, judges did not have to activate and generate this knowledge anew while judging their best friend. As a consequence, these judgments were facilitated and participants were able to provide them more rapidly. Of importance, this facilitation effect is only apparent in judgments about the best friend, indicating that it is indeed knowledge that pertains specifically to this routine standard that is rendered accessible. If more general semantic knowledge were rendered accessible, then this knowledge would be applicable to judgments about any person so that judgments about the best friend and judgments about the former friend should be facilitated in similar ways. Clearly, this is not the case. Furthermore, the facilitation of judgments about participants' best friends only occurred if participants had previously evaluated themselves, not when they had judged another person. This suggests that the use of the routine standard is specific to judgments about the self.

Together with Study 1, these findings are consistent with the notion of routine standard use in self-evaluation. In line with our theoretical assumptions, the activation of knowledge about one's best friend appears to play an important role in self-evaluation. When judging themselves, judges appear to think about their friend's characteristics in the critical domain and use this information for comparison. Study 3 was designed to examine the robustness of this mechanism.

Study 3

The use of routine standards may at times come at the cost of diagnosticity. At least in some situations, our routine standards may not be the most diagnostic because they may be quite different from us either on the critical dimension itself (Festinger, 1954) or on attributes that are related to this dimension (Goethals & Darley, 1977). Do people still use knowledge about their routine standards in the self-evaluation process if these standards are clearly dissimilar and thus potentially undiagnostic? Or will people forego the use of routine standards when normative concerns of diagnosticity are violated? Our next experiment examines this important question.

To do so, we put participants in similar conditions as in Study 2. Again, they first evaluated either themselves or another person before judging either the routine standard or another potential standard. This time, however, the diagnostic value of the comparison with the routine standard was deliberately undermined by choosing the dimension for self-evaluation on which the best friend was perceived as most dissimilar from the self. Because, according to social comparison theory (Festinger, 1954), similarity breeds diagnosticity, the routine standard should thus be undiagnostic for self-evaluation. Would information about the routine standard's characteristics in the critical domain still be activated during self-evaluation so that subsequent judgments about this standard are facilitated? To provide a particularly strong test for the use of routine standards, we used a different potential standard as a control this time. In particular, subsequent to evaluating themselves, some of our participants judged the one of their acquaintances who was most similar, and thus most diagnostic, to them on the critical dimension. Would judgments about the dissimilar best friend or about the similar acquaintance be primarily facilitated by the preceding self-evaluation?

Method

Participants. We recruited 58 students at the University of Würzburg as participants. As in the previous studies, the experiment was part of a 1-hr session for which participants were offered DM 12 (approximately $6 at the time).

Materials and procedure. As in the preceding studies, participants first named several same-sex members of their social network. This time we wanted to examine whether participants use their routine standard even if he or she is quite different from them on the comparison dimension. Therefore, we used a different cover story. Participants were made to believe that we were interested in the importance of similarity in personal attributes for the maintenance of friendships. Using this cover story allowed us to ask participants on which dimension they perceive their best friend as least similar to themselves. Participants could choose 1 out of 12 personal attributes (e.g., *intelligent* [intelligent], *aktiv* [active], *offen* [open], *selbstbewusst* [self-confident]). Subsequently, they were instructed to name another same-sex person of their social network who is very similar to them on the previously chosen dimension. In this first block of the experiment, we thus assessed the name of the best friend, a dimension on which the best friend is very dissimilar to the participant, and the name of another same-sex person (acquaintance) who is very similar to the participant on that specific dimension. After finishing the first block, participants worked on several other, unrelated experiments. The time gap between the first and the second block of the experiment was at least 15 min.

Participants' task in the second block was similar to the task in the second block of Study 2. This time we explained that we wanted to assess participants' perceptions of people who are more or less close to them. Participants were first induced to either evaluate themselves or another person. In particular, they were instructed to take a minute and think about a person with respect to a given personal characteristic and to subsequently write down their thoughts. Next, participants positioned their fingers on the number keys of the computer keyboard and finally judged several persons with respect to the same characteristic used in the evaluation task. Again, we emphasized that participants should take their time in the evaluation task, but should react as fast as possible in the judgment task. In contrast to Study 2, this whole block consisted of only one trial. Only the dimension

on which participants perceived their best friend as very dissimilar was used. On that critical dimension, half of the participants first thought about and described themselves. The other half thought about and described the well-known same-sex celebrity (tennis players Steffi Graf or Boris Becker) already used in Study 1. Subsequently, all participants rated eight persons with respect to the same dimension along a 6-point scale ranging from 1 (*very little*) to 6 (*very much*). The first person was always the one participants had evaluated in the evaluation task (the self or the tennis player). The judgment of the second person was the critical one, for which the response latencies constitute our dependent variable. Whereas about half of our participants rated their best friend, the other half rated their acquaintance. The remaining target persons were other friends and celebrities (e.g., Madonna, Joschka Fischer).

To ensure that participants' best friends and acquaintances fulfill the criteria necessary to test our hypotheses, we asked a series of manipulation check questions after the judgment task. First, to ensure that participants are indeed in closer contact with their best friend, so that this person is more likely to be routinely used for comparison than the acquaintance, participants answered the following questions: "How often do you have contact with your friend [name]?" and "How close is your friend [name] to you?" To ensure that the best friend is at the same time more dissimilar on the critical dimension than the acquaintance, participants indicated their perceived similarity to both friends on the critical dimension. If, for example, *open* was the critical dimension chosen in the first block, participants were asked, "How open is your friend [name] compared to you?" and "Think about attributes that make a person open. How similar to you is your friend [name] on those attributes?" Furthermore, to assess whether differences in the perceived general similarity between the self and the respective standards exist, participants were asked the following: "In general, how similar to you is your friend [name]?" All answers were given along 9-point rating scales and were asked for both, the best friend and the acquaintance. After completion of the second block of the experiment, participants were thanked, fully debriefed, paid, and dismissed.

In sum, Study 3 is based on a 2 (evaluation: self vs. other) $\times$ 2 (judgment: best friend vs. acquaintance) between-participants design. All participants first thought about and evaluated either themselves or the tennis player and subsequently judged either their best friend or their acquaintance. Unlike in Study 2, we used a between-participants design this time because doing so allowed us to use only the one characteristic in which the best friend was perceived as least similar in all conditions.

Results

In Study 3 participants evaluated themselves along a dimension on which the best friend as a routine standard is fairly undiagnostic. Participants deliberately selected an attribute on which their best friend is very dissimilar to themselves. At the same time, they named an acquaintance who, in contrast to their best friend, is very similar to them on that specific attribute. If similarity is the critical factor in standard selection, then participants should use their similar acquaintance rather than their dissimilar best friend as a comparison standard in self-evaluation. If, however, the best friend is used because he or she is routinely used in social comparison, similarity should only play an inferior role. In this case, participants should compare themselves with the routine standard even if he or she is undiagnostic. The preference of one over the other standard should become apparent in the response latencies in the subsequent judgment. Whoever is used as a comparison standard in the self-evaluation, the similar acquaintance or the dissimilar best friend, should subsequently be judged faster.

Preliminary analysis. Most participants (28%) chose the attribute *gründlich* [thorough] as the one on which they are least similar to their best friend. Of the remaining 11 attributes, 2 (*intelligent* [intelligent] and *herzlich* [hearty]) were chosen by none. To check whether the best friend, the acquaintance, and the selected dimension fulfill the criteria necessary to test our hypothesis, we asked participants a series of manipulation check questions. The results, summarized in Table 1, demonstrate that our manipulation was successful. Altogether the answers are as expected. Participants indeed indicated that they meet their best friend more often and perceive him or her to be closer to them than their acquaintance. Thus, the best friend is indeed more likely to be routinely used for comparison than the acquaintance. At the same time, these data indicate that the best friend is a less diagnostic standard than the acquaintance. Participants rated their best friend as less similar to themselves than the acquaintance, both on the critical dimension itself and on related attributes. In addition, this difference in the specific similarity on the selected dimension is also apparent in participants' responses in the judgment task. In particular, the mean difference between participants' judgments on the critical dimension of themselves and their best friend was more pronounced ($M = 2.41$) than the mean difference between judgments of themselves and their acquaintance ($M = 0.86$), $t(57) = 9.39$, $p < .001$. In contrast to the specific similarity on the critical dimension, there is no indication that overall perceived similarity between the self and the best friend versus the acquaintance differed (Table 1).

Response latencies. We used the response latencies for the judgment of the best friend versus the acquaintance as our main dependent variable. Short latencies indicate activation during the previously conducted evaluation of the self versus the other person and thus the use of this person as a standard in the comparison process.

We classified three response latencies ($< 5.5\%$) as outliers. Each of these responses was at least 3,000 ms slower than the slowest of the remaining responses ($M = 3{,}955$, $SD = 1{,}549$) and all took longer than 10,000 ms.

Table 1
Differences in the Perception of the Best Friend and the Acquaintance (Study 3)

Item	Best friend	Acquaintance	$t(57)$
How often do you have contact with your friend [name]?[a]	7.31	6.03	3.81*
How close is your friend [name] to you?[a]	7.38	6.22	6.23*
How [critical dimension] is your friend [name] compared to you?[b]	2.45	0.71	9.17*
Think about attributes, which make a person [critical dimension]. How similar to you is your friend [name] on those attributes?[a]	4.64	5.87	−4.00*
In general, how similar to you is your friend [name]?[a]	5.64	5.88	−0.80

Note. $N = 58$. The critical dimension was chosen out of 12 personal characteristics. Participants selected the characteristic on which they perceive their best friend as least similar to themselves.
[a] Response scale ranging from 1 (*very little*) to 9 (*very much*). [b] Absolute difference to midpoint (*the same*) of the scale with the endpoints 1 (*much less*) and 9 (*much more*).
* $p < .001$.

Unlike the previous study, Study 3 only involved one critical trial. To control for the variability of response latencies between participants, we calculated individual baselines and subtracted them from the critical response latencies. It should be noted that it was not necessary to use a baseline in Studies 1 and 2, because in these studies we manipulated the critical factor (best friend vs. other friend) within participants so that the baseline is already "built in." For the baseline in Study 3, we calculated the mean response latency for judgments about the three celebrities used as filler items in the judgment task. Prior to subtraction, all latencies were log transformed (ln). For ease of interpretation we again report the untransformed means (in milliseconds).

We hypothesized that our participants may use their best friend as a routine standard even if he or she is dissimilar to them and therefore less diagnostic. If this were indeed the case, then participants should be faster in judging their best friend than their acquaintance after evaluating themselves on the same dimension. Evaluating another person, however, should not facilitate judgments of the best friend. The means reported in Figure 3 are consistent with this hypothesis. Participants were fastest if they judged their best friend subsequently to self-evaluation ($M = 213$ ms). This latency is significantly shorter than the time participants needed to judge their acquaintance ($M = 1{,}066$ ms) in the same condition, $t(26) = 2.02$, $p < .05$ (two-tailed). In contrast, there was no difference in response latencies if participants judged the best friend ($M = 1{,}031$ ms) or the acquaintance ($M = 787$ ms) after evaluating the tennis player, $t(25) = -0.28$, *ns*. This pattern was borne out in a significant interaction effect in a 2 (evaluation: self vs. other) $\times$ 2 (judgment: best friend vs. acquaintance) ANOVA using the log-transformed response latencies for the judgment task as the dependent variable, $F(1, 51) = 5.02$, $p < .03$, $d = 0.63$. In this analysis, none of the main effects proved to be significant (Fs < 2.16, ps $> .14$). Furthermore, the critical interaction effect still held if differences in the general perceived similarity of best friend and self versus acquaintance and self were statistically controlled for. Using this difference as a covariate did not reduce the significance of the interaction effect, $F(1, 50) = 5.56$, $p < .03$, $d = 0.68$.

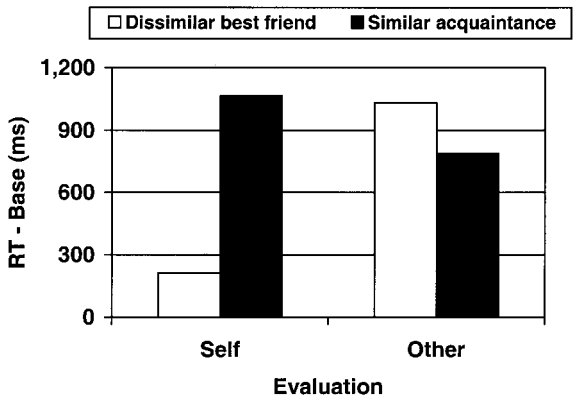

Figure 3. Response latencies (difference to baseline) for judgments of best friend versus acquaintance subsequent to evaluation of self versus other (Study 3). RT = reaction time.

Discussion

These findings show a striking robustness of the use of routine standards in the self-evaluation process. Even if participants' best friends were only minimally diagnostic for comparisons on the critical dimension, did participants activate information about them. In fact, the present results indicate that after self-evaluation information pertaining to the undiagnostic best friend was more accessible than information pertaining to the maximally diagnostic acquaintance who was deliberately selected to be similar to the self. This is apparent in the fact that self-evaluation facilitated only judgments about the best friend, not about the similar acquaintance. In light of the fact that knowledge accessibility is primarily determined by recent activation (Higgins, Rholes, & Jones, 1977), this suggests that in evaluating themselves, judges thought more about their undiagnostic best friend than about their diagnostic acquaintance. In this respect, judges' tendency to rely on their routine standards for self-evaluation appears to be so robust that it even prevails in situations in which comparing with the routine standard is unlikely to provide diagnostic information for self-evaluation.

At the same time, the current results suggest that participants do not select their best friends as comparison standards because they see them as more similar to themselves. Perceived similarity does not appear to be the driving force behind the use of routine standards. No differences in the overall similarity between the self and the respective standards were detected, and the critical interaction effect still held if differences in overall similarity were statistically controlled for. This finding is consistent with the assumption that routine standards are not selected for comparison because of specific attributes they hold, but because of their procedural qualities. Our final study was designed to provide more direct evidence for this assumption.

Study 4

To do so, we applied the second principle research strategy and experimentally created routine standards by inducing participants to repeatedly compare with one particular standard before engaging in self-evaluation. Using such experimentally created routine standards to examine standard selection in social comparison has two distinct advantages. First, it allows to control for differences in the specific attributes of different potential comparison standards. In Studies 2 and 3 we have statistically controlled for one particularly important attribute, namely similarity to the self. In principle, however, an endless number of such differences (e.g., liking) exists so that it is impossible to statistically control for all of them. Creating standards for which participants have no prior knowledge circumvents this potential problem. Second, the use of experimentally created standards allows us to directly manipulate the procedural qualities that we assume to underlie the use of routine standards, namely their repeated use in prior comparisons. As a consequence, Study 4 allows us to take a more direct look at the procedural underpinnings of routine standard formation and use in social comparison.

To experimentally create routine standards, participants repeatedly compared with the same standard. They received information about 15 different preferences (e.g., favorite novel, movie, cuisine) of a fictitious person named Daniela and were instructed to form

an impression of her. This information was presented in a comparative format so that participants were induced to list their own preferences right after reading about those of Daniela. Thus, in forming an impression of Daniela, participants repeatedly compared themselves to her. Participants also received information about a second fictitious person named Martina. This information, however, was presented in an isolated format so that in forming an impression of Martina, participants processed her preferences without relating them to their own. In forming impressions about both potential standards, judges are thus likely to compare themselves more elaborately with Daniela than with Martina so that comparisons with Daniela are better practiced. In this respect, Daniela constitutes an experimentally created routine standard. If participants do indeed preferentially use routine standards for comparison, then they should primarily use Daniela as a comparison standard for subsequent self-evaluation.

To see whether this is indeed the case, we applied an additional indirect measure of standard selection in social comparison. In particular, we used the self-evaluative consequences of comparison as a diagnostic tool to examine standard selection. Social comparisons influence subsequent self-evaluations in predictable ways (for an overview, see Mussweiler, 2003). Our previous work has established that in this specific social comparison paradigm, judges typically assimilate self-evaluations toward the standard that was used for comparison (Mussweiler, 2001a, 2001b; Mussweiler & Strack, 2000). This is the case, because in comparing themselves with a pertinent standard, judges selectively activate self-knowledge indicating that their standing on the judgmental dimension is similar to that of the standard. Because this knowledge is then used as a basis for self-evaluations, these are assimilated toward the standard (Mussweiler, 2003). This assimilative link between self-evaluations and the critical standard of the comparison process allows us to use self-evaluative comparison consequences as an unobtrusive indicator of standard selection. If self-evaluations are assimilated to the routine-standard Daniela, as we assume, then this indicates that judges used Daniela as a comparison standard during the self-evaluation process.

Method

Participants. We recruited 28 students at a cafeteria of the University of Würzburg as participants and offered them an ice cream cone or a bar of chocolate as a compensation. On agreement to participate, they were led to a separate room where they completed the experiment in groups of up to 10.

Materials and procedure. Unlike Studies 1–3, our final study is a paper-and-pencil experiment. On arrival in the lab, participants were greeted by the experimenter, led to a separate table, handed the questionnaire, told to read instructions carefully, and to work through the materials in the given order.

The first page of the questionnaire included general instructions and explained the ostensible purpose of the study. In particular, we pointed out that we were interested in the influence of different kinds of information on person perception. During the experiment, participants would be asked to form an impression of two fictitious persons of their own gender. They would read a list of preferences of these two people and a short paragraph about their experiences at the university. To allow us to control for potential effects participants' specific characteristics may have on their evaluation of the target persons, participants would also provide information about their own preferences and experiences.

On the second page the routine standard was created by engaging participants in an impression formation task that was structured in a comparative format. Participants were instructed to form an impression of a fictitious person whose gender matched theirs. For clarity, we only describe the female version throughout the *Method* section. Participants received a list of 15 preferences of a fictitious person named Daniela (e.g., her favorite novel, movie, cuisine) and were instructed to list their own preference in the respective domain right underneath Daniela's. Beneath each of Daniela's preferences there was a blank line for participants to fill in their own preference. For example, participants read, "*Daniela's favorite cuisine is Italian. My favorite cuisine is* _____." Thinking about one's own preference right after reading about Daniela's is likely to involve a comparison with Daniela, so that this impression formation task is likely to involve 15 comparisons with the same standard. Such repeated practice in social comparison is likely to establish Daniela as a routine standard. After completion of all 15 preferences, participants briefly described their first impression of Daniela in an open-ended format.

Next, participants were instructed to form an impression of a second fictitious person named Martina. This second impression formation task was structured in an isolated format so that it was less likely to involve comparisons with the self. In particular, in forming an impression of Martina participants were presented with a list of her preferences without indicating their own preferences (e.g., *Martina's favorite cuisine is Thai*.). The listed preferences of Martina and Daniela were counterbalanced so that no differences in the inferred characteristics of both standards exist. Once again, the impression formation task was concluded by an open-ended question about the first impression of the described person.

After participants had formed impressions of the routine-standard Daniela and the nonroutine-standard Martina, they proceeded with the self-evaluation task. Instructions for the self-evaluation task pointed out that participants would now receive additional information about Daniela and Martina as well as three further persons (Klara, Susanne, Ursula), which may be helpful in forming an impression of these people. This information would pertain to their experiences during their first semesters as students at the University of Würzburg and would describe how well they managed to adjust to life at the university. We informed participants that we would ask them a series of questions concerning their own adjustment to life at the university and that they would have the opportunity to read about the experiences of Daniela, Martina, and the others before answering these questions.

The experiences of the five standard students, which were presented on the next page, consisted of five brief paragraphs in which each of them provided some general background information (start of the studies, graduation from college) and commented on their adjustment to college life. All five paragraphs were printed on a single page with the description of Daniela and Martina in the second and the fourth positions. Order was counterbalanced between conditions. The two descriptions of the critical standards Daniela and Martina were modeled after those used by Mussweiler (2001b). One of the paragraphs described a person who had adjusted very well to college because she had moved a lot during childhood and liked to approach new life settings as a kind of challenge, and was thus portrayed as a high-comparison standard for adjustment to college. The other paragraph portrayed a person who had a hard time adjusting because she had never moved before and was skeptical to face new life settings, and thus constituted a low standard of comparison. For all participants, either Daniela or Martina were described as adjusting very well or very poorly to college life. The remaining three descriptions were more neutral in tone and were identical for all participants. The descriptions of Daniela and Martina, however, were varied between participants. For half of them, the routine-standard Daniela was described as the high standard and the nonroutine-standard Martina constituted the low standard. For the other half, this assignment was reversed. Thus, the two critical standards were described in opposite ways for each participant. If Daniela adjusted easily to college, then Martina adjusted poorly and vice versa.

To assess the self-evaluative consequences of exposure to the five potential standards, we subsequently asked a series of questions concerning participants' own adjustment to life at the university. These questions were as follows: On average, how often do you visit home during a semester? How many friends do you have in Würzburg? How many weeks did you need to adapt to life in Würzburg? How often do you go out per month? How often do you call home per month? How well did you adapt to life at the university? (1 = *very badly,* 9 = *very well*); In general, how well do you adapt to new life settings? *(*1 = *very badly,* 9 = *very well*).

At the end of the experiment, participants judged Daniela's and Martina's adjustment to college by answering the same seven questions again (e.g., What do you estimate, how often does Martina go out per month?). In addition, participants indicated how much they liked Daniela and Martina, how similar these two are to themselves concerning their adjustment to life at the university, and how similar they are to themselves in general. These ratings were given on 9-point scales ranging from 1 (*not at all*) to 9 (*very much*). Finally, participants were thanked for participation, debriefed, and offered their compensation.

In sum, Study 4 is based on a simple two-cell design: Participants were either exposed to comparison information portraying the routine-standard Daniela as a high standard and the nonroutine-standard Martina as a low standard, or they were exposed to comparison information portraying Daniela as a low standard and Martina as a high standard.

Results

The critical standard descriptions used in this study were adapted from those used by Mussweiler (2001b). In this previous research, participants typically assimilated toward the standard (see also, Mussweiler, 2001a, 2003; Mussweiler & Strack, 2000) and evaluated their own adjustment to college more positively after a comparison with a high rather than a low standard. In Study 4 we use this established assimilation effect as an unobtrusive indicator of standard selection. Participants were exposed to both, the description of a high and a low standard. Their self-evaluations should be primarily influenced by the standard with which they primarily compared themselves. Therefore, if they subsequently describe themselves as adjusting well to college, then this indicates that they had used the high rather than the low standard as a standard of comparison. We expected that participants primarily compare and thus assimilate their self-evaluation with the routine standard established in the first part of the experiment. If this is indeed the case, then participants should evaluate their own adjustment to college more positively if the routine-standard Daniela was described as a high rather than a low standard. If, however, Daniela is not preferentially selected and does not assume a particularly prominent role in the self-evaluation process, then similar self-evaluations should be given in both conditions, because across all five standards the same information was provided.

Preliminary analyses. To check whether participants did indeed perceive the high standard as better adjusted to life at the university than the low standard, we combined the ratings of the standards in two single scores. We *z*-transformed and averaged the answers to the seven critical questions and, where necessary, recoded the answers so that higher values indicate better adjustment. As expected, participants judged adjustment to college to be better for the high standard ($M = 0.41$) than for the low standard ($M = -0.41$), $t(27) = 4.51$, $p < .001$ (two-tailed).

To ensure that the expected differences in standard selection are not produced by differences in perceived liking of and similarity to Daniela and Martina, we examined participants' ratings of these attributes. No differences between Daniela and Martina emerged (ts < 1.7, ps $> .10$).

Self-evaluations. Our central measure is participants' self-evaluations of their adjustment to college. To combine the seven ratings into a single score (Cronbach's $\alpha = .54$), we *z*-transformed and, where necessary, recoded the individual ratings so that for the resulting mean, higher values indicate better adjustment. Consistent with our hypothesis, participants' self-evaluations were assimilated toward the routine-standard Daniela so that more positive self-evaluations were given if Daniela was described as a high ($M = 0.22$) rather than a low standard ($M = -0.32$),[1] $t(26) = 2.94$, $p < .01$ (two-tailed), $d = 1.15$. Furthermore, neither controlling for differences in perceived liking of, $F(1, 25) = 9.73$, $p < .01$, $d = 1.25$, nor for differences in similarity to Daniela and Martina, $F(1, 25) = 8.97$, $p < .01$, $d = 1.20$, changed the significance of this effect.

Discussion

These findings indicate that self-evaluations were assimilated to the standard that was previously turned into a routine standard by repeated comparisons. Self-evaluations are primarily influenced by the standard that was predominantly used during the self-evaluation process. In this respect, this self-evaluation effect constitutes additional converging evidence, suggesting that participants mainly compared themselves with the routine standard. In light of the fact that the routine standard was experimentally created, this finding extends those of Studies 1 through 3 in important ways. First, it demonstrates the generality of routine standard selection in social comparison. Not only one's best friend may serve as a routine standard, but any standard that has been extensively used for comparison. Given that with our friends, partners, family members, and colleagues a multitude of potential routine standards exists, the use of routine standards for comparison is likely to be fairly general in nature. Second, because for experimentally created standards no differences in specific attributes such as liking and similarity to the self exist, these potentially important influences on standard selection cannot be responsible for the demonstrated preference to compare with the routine standard. The only systematic difference between the two standards is that participants had repeatedly compared themselves with the routine standard but not with the nonroutine standard. Although identical information was provided about the two, only the information about the routine standard was repeatedly compared with the self. From our perspective, these repeated comparisons served as a procedural priming that rendered comparisons with the routine standard more likely to be carried out. Finally, because in Study 4 we directly manipulated those processes that lead to the development of a routine standard, these findings provide more direct evidence that repeated comparisons with a given standard do indeed increase the chances that this standard will be selected for subsequent comparisons.

[1] The mean *z* values do not add up to zero because of unequal cell sizes and missing values on some of the individual judgments.

General Discussion

In the present research, we have examined the use of routine standards in social comparison. Consistent with our assumptions, our findings suggest that in evaluating themselves, people do indeed activate information about others with whom they are likely to compare on a routine basis. For one, we have demonstrated that in evaluating themselves, our student participants activated information about their best friends, which—on a priori grounds (Hays & Oxley, 1986; Hirsch, 1980)—appear to be a likely routine standard for this population. Furthermore, we have shown that an experimentally created routine standard is more likely to be used for comparison so that it primarily determines self-evaluative comparison consequences.

In the present findings, the use of routine standards in self-evaluation is apparent on three distinct measures that are all indirect rather than explicit measures of standard selection. Study 1 demonstrates that subsequent to a series of self-evaluative judgments, participants were faster in recognizing the name of their best friend but not the name of an exfriend in a lexical decision task. Studies 2 and 3 further demonstrate that participants were faster in judging their best friend, but not another friend, on the dimension on which they had previously evaluated themselves. This suggests that participants had activated information about the best friend's standing on the judgmental dimension during self-evaluation. Furthermore, Study 4 demonstrates that the self-evaluative consequences of comparison are primarily influenced by the nature of the routine standard. Taken together, these studies provide substantial converging support for the notion of routine standard use in social comparison.

Extending these basic findings, Study 3 provides a first test of potential limits of the use of routine standards by deliberately choosing a dimension on which the routine standard is dissimilar to the self. Even under such conditions of low diagnostic value do judges appear to activate knowledge about their best friend during self-evaluation.

At first sight, one potential ambiguity of these data may appear to remain. In particular, one could argue that the enhanced accessibility of knowledge about participants' best friend demonstrated in Studies 1–3 may not necessarily result from a comparison of self and best friend. Rather, best friends may be so closely associated with the self, that any activation of the self spreads to the representation of one's best friend. Such a spreading activation mechanism, however, is unlikely to be responsible for our data. For one, in light of the ubiquity of comparison processes in self-evaluation, it seems unlikely that information about a potential standard that is easily accessible would not be used for self-evaluative purposes. In fact, coaccessibility has been suggested as a primary determinant of social comparison activities (Kruglanski & Mayseless, 1990). Furthermore, the fact that in Study 4 self-evaluations were clearly influenced by the accessible routine standard suggests that judges did indeed use these standards for comparison. If knowledge about the routine standard had only been accessible but had not been related to the self, then it should not influence self-evaluations. Thus, it is difficult for a spreading activation mechanism to account for the data we obtained across the four studies. In addition, some of our most recent evidence (Rüter & Mussweiler, 2003) demonstrates that the accessibility of knowledge about the routine standard is not necessarily increased when people activate self-knowledge on the critical dimension, as is implied by a spreading activation mechanism. Rather, routine standard knowledge only becomes accessible in self-evaluative situations that are comparative in nature. In one of these studies, for example, participants who described their eating habits in concrete absolute terms (e.g. "What do you eat? Where do you eat?"), which did not require them to engage in comparisons, did not show increased accessibility of knowledge about their best friends. Participants who evaluated themselves in terms that are relative in nature (e.g., "How healthy are your eating habits?") and require comparisons with others, however, subsequently had knowledge about their best friend more accessible. These findings suggest that the mere activation of self-knowledge in self-evaluation is not sufficient to produce the accessibility effects we have examined. In light of these findings, the increased accessibility of routine standard knowledge we have demonstrated indicates that the routine standard has indeed been compared with the self.

Taken together, the present findings thus suggest that our participants had a strong inclination to use routine standards for self-evaluation. In fact, they even activated information about their best friends under circumstances for which using this information violates normative considerations to obtain diagnostic information through the comparison (Festinger, 1954; Goethals & Darley, 1977). Furthermore, our data suggest that the use of routine standards in social comparison constitutes a rather general and stable effect. Across the four studies, evidence for routine standard use was obtained with different types of routine standards, for self-evaluations on a multitude of different dimensions, and was apparent in measures that are quite different in nature.

Implications for Social Comparison Theory and Research

The present research has important implications for social comparison theory and research. Most importantly, our results indicate that the selection of social comparison standards may not primarily depend on normative concerns to select the most diagnostic standard, as classic treatments of standard selection suggest (Festinger, 1954; Goethals & Darley, 1977). Instead of using standards who promise to offer the most diagnostic self-evaluative information because they are highly similar to the self, judges may simply select standards they routinely use for comparison. In fact, the results of Study 3 suggest that judges may even use their routine standards for comparison if the self-evaluative dimension was deliberately selected to be one for which the routine standard is most different from the self. That is, even under conditions that minimize the diagnostic value of the best friend as a comparison standard do judges activate information about this routine standard during self-evaluation. These findings suggest that under specific conditions, efficiency considerations may weigh more heavily in the selection process than normative considerations to select the most diagnostic standard. At times, the heuristic value of using a potential standard may thus become the primary determinant of its selection. Standards may be primarily selected because they are easy to use, not because they promise to provide the most relevant information for self-evaluation.

In fact, this potential primacy of routine standard use is not limited to self-evaluative comparisons. It is also apparent in evaluations of others. When judging another person, people typically compare this person with themselves (Dunning & Hayes, 1996;

Mussweiler & Bodenhausen, 2002). In this respect, the self serves as a routine standard in social judgment. This suggests that routine standard use may well be a bidirectional mechanism: Judges use others as routine standards when evaluating themselves and use themselves as a routine standard when evaluating others.

Such a widespread use of routine standards is well consistent with an emerging perspective on social comparison processes that emphasizes its spontaneous and automatic qualities (e.g., Dunning & Hayes, 1996; Gilbert et al., 1995; Mussweiler & Bodenhausen, 2002; for a discussion, see Mussweiler, 2003). Comparisons do play a core role in human judgment in general and in self-evaluation in specific, so that social judges have to be efficient comparers. Because we constantly compare, we cannot afford to allocate too many resources to this process. As a consequence, not only the comparison itself (Gilbert et al., 1995) but also the selection of comparison standards has to occur in a relatively automatic fashion. Normative concerns to obtain valuable information through social comparison may only come into play once the comparison has been carried out. If sufficient mental resources are available, judges may then correct for unwanted or inappropriate comparison consequences (Gilbert et al., 1995).

Mechanisms of Routine Standard Formation and Use

The formation and use of routine standards in social comparison is also consistent with a more general perspective on social judgment that emphasizes its procedural nature (Smith, 1989, 1994). From this perspective, social judgments are conceptualized as a case of procedural learning in which practice facilitates those inferences that are involved in the judgment process so that it continuously becomes more efficient. Because people typically act as cognitive misers (Taylor, 1981), they are particularly inclined to apply those cognitive processes that are highly efficient, so that well-practiced processes are preferentially engaged. In much the same way, judges may be particularly inclined to apply those comparison processes that are well practiced and consequently highly efficient.

For social judgment, the consequences practice has for the efficiency as well as for the content of judgment have been closely examined (for an overview, see Smith 1994). This research established that practice has two types of efficiency advantages: a general and a specific one (Smith, 1989). For example, repeatedly judging whether a behavioral term (e.g., *help*) is indicative of a target trait (e.g., *friendly*) facilitates subsequent judgments so that judges become faster in relating new terms (e.g., *ask*) to the target trait (e.g., *friendly*). This general practice effect thus increases the efficiency of relating any stimulus to the target trait. In addition, there is a specific practice effect so that judges become especially fast in relating a previously encountered term to the target trait. This specific practice effect thus particularly increases the efficiency of relating the same stimulus to the target (Smith, 1989). Practice, however, does not only determine how fast and efficient a given judgment can be made, it also influences the very content of the judgment. Judges as cognitive misers (Taylor, 1981) preferentially apply those processes that are highly efficient and require little of their precious processing capacity. If alternative processes can be used for a given judgment, judges thus resort to the most efficient one, so that well-practiced processes are preferentially engaged. In social judgment, judges with practice in relating behaviors to the target trait friendliness, for example, are more likely to interpret an ambiguous behavior in terms of friendliness than in terms of intelligence. For judges with practice in relating behaviors to intelligence, however, the reverse is the case (Smith, 1989). A procedural perspective on social judgment (e.g., Smith, 1989, 1994) thus suggests that relating a particular stimulus to a target particularly increases the efficiency of subsequently relating the same stimulus to this target. As a consequence, this specific well-practiced procedure is preferentially used for social judgment.

In light of the fact that social comparisons are in essence social judgments, similar types of practice effects may also lead to the formation and use of routine standards in social comparison. Such a procedural perspective suggests that repeatedly relating a particular other to oneself may not only facilitate social comparison processes in general but also the specific comparison of the self and the repeatedly used standard. That is, with repetition and practice, a specific comparison is likely to become increasingly efficient so that it is more likely to be engaged. Judges who resort to the most efficient process for self-evaluation are thus likely to repeatedly compare with the same standards. In this respect, the more often we compare with the same standards, the more efficient this comparison will become and the more likely we are to repeatedly use them for subsequent comparisons. From this perspective, routine standards may be preferentially used for comparison, because judges are more efficient in comparing themselves with a routine standard than with a nonroutine standard. Although this potential efficiency advantage is clearly consistent with the reported data, providing direct support for this implication is beyond the scope of the present research. In light of the extensive literature demonstrating efficiency advantages of well-practiced processes in social judgment (Smith, 1989, 1994), however, routine standard use is likely to contribute substantially to the facilitation of self-evaluative judgments.

Limits of Routine Standard Use

Nevertheless, some limits to the use of routine standards in social comparison do appear to exist. Clearly, there exist circumstances under which obtaining diagnostic information is so valuable that judges are likely to be highly strategic and guided by normative concerns in their choice of comparison standards. If the outcome of the comparison is of high personal importance, judges are thus likely to be more strongly led by the diagnosticity of the information that is gained by the comparison. Thus, when judges are highly motivated to obtain diagnostic information so that the selection of comparison standards is more deliberate, they may forgo the use of routine standards in the self-evaluation process.

This may also be the case in situations in which self-protective rather than normative concerns are of primary importance. In particular, protecting one's self-regard may sometimes be the strongest concern during social comparison (Taylor, Wayment, & Carrillo, 1996; Wood, 1989; Wood & Taylor, 1991), especially if the positive self-image most people are motivated to entertain is threatened. Under such conditions, people may also forgo the use of the routine standard and, instead, select a standard who sheds the most favorable light on them. One self-protective selection strategy that may be used in such a way is the selection of a downward standard, that is, a person who is worse off than oneself

(Wills, 1981). Because one's own weaknesses may pale relative to those who do even worse, such downward comparisons may offer self-protective benefits. In this respect, motivational concerns for self-protection may guide standard selection and work against the use of a routine standard.

There are, however, also conditions that are likely to further promote the use of routine standards. As we have suggested before, decision routines are particularly likely to be used under suboptimal conditions that preclude more deliberate or strategic processing (e.g., Betsch et al., 1998). Thus, if judges have limited cognitive capacities available for evaluating themselves or if they have to do so under time pressure, then they are particularly likely to use routine standards for comparison. Notably, because self-evaluative thoughts make up such a large portion of our total mental activity (Csikszentmihalyi & Figurski, 1982), we will rarely be able to engage in deliberate and thus capacity-consuming modes of standard selection. Instead, we are likely to resort to our routine standards in most situations.

It is also important to note that the efficiency considerations that speak for the use of routine standards and the diagnosticity considerations that speak for the use of similar standards do not necessarily have to be at odds with one another. Because we typically surround ourselves with people who are generally similar to us (Berscheid & Reis, 1998; Byrne, 1997), and because in the domain of social comparison similarity breeds diagnosticity (Festinger, 1954; Goethals & Darley, 1977), on average, routine standards are likely to be reasonably diagnostic. Thus, in many situations the use of routine standards may well offer the best of both worlds—they simplify the selection and comparison process and are at the same time sufficiently diagnostic. In this respect, routine standards may often be the optimal choice for people trying to "satisfice" (Simon, 1956) in their self-evaluative endeavors.

Conclusion

Comparisons with others play a core role in self-evaluation processes. To understand who we are, we have to compare with others. In fact, such social comparisons play such an important role in our lives and are so often engaged that they have to operate in a relative automatic fashion. If social comparisons were mostly deliberate processes that involve a conscious and strategic choice of comparison partners, we would be overwhelmed by the sheer number of comparison opportunities. Deliberate social comparisons are too costly to be routinely engaged. As a consequence, the social judge is in dire need of comparison strategies that simplify comparisons, so that they can be efficiently engaged. Using routine standards appears to be one such strategy, which allows us to constantly compare with others without paralyzing us by overusing scarce cognitive resources.

References

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal directed behavior. *Journal of Personality and Social Psychology, 78,* 53–63.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Hillsdale, NJ: Erlbaum.

Berscheid, E., & Reis, H. T. (1998). Attraction and close relationships. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 193–281). New York: McGraw-Hill.

Betsch, T., Fiedler, K., & Brinkmann, J. (1998). Behavioral routines in decision making: The effects of novelty in task presentation and time pressure on routine maintenance and deviation. *European Journal of Social Psychology, 28,* 861–878.

Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., & Fiedler, K. (2001). The effects of routine strength on adaptation and information search in recurrent decision making. *Organizational Behavior and Human Decision Processes, 84,* 23–53.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships, 14,* 417–431.

Compas, B. E., Wagner, B. M., & Slavin, L. A. (1986). A prospective study of life events, social support, and psychological symptomatology during the transition from high school to college. *American Journal of Community Psychology, 14,* 241–257.

Csikszentmihalyi, M., & Figurski, T. J. (1982). Self-awareness and aversive experience in everyday life. *Journal of Personality, 50,* 15–28.

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

Fazio, R. H. (1990). A practical guide to the use of response latencies in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74–97). Newbury Park, CA: Sage.

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

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology, 69,* 227–236.

Goethals, G. R., & Darley, J. M. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 259–278). Washington, DC: Hemisphere Publication Services.

Gruder, C. L. (1971). Determinants of social comparison choices. *Journal of Experimental Social Psychology, 7,* 473–489.

Hays, R. B., & Oxley, D. (1986). Social network development and functioning during a life transition. *Journal of Personality and Social Psychology, 50,* 305–323.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13,* 141–154.

Hirsch, B. J. (1980). Natural support systems and coping with major life changes. *American Journal of Community Psychology, 8,* 159–172.

Huttenlocher, J., & Higgins, E. T. (1971). Adjectives, comparatives, and syllogisms. *Psychological Review, 78,* 487–504.

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

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

Miller, D. T., & Prentice, D. A. (1996). The construction of social norms and standards. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 799–829). New York: Guilford Press.

Mussweiler, T. (2001a). Focus of comparison as a determinant of assimilation versus contrast in social comparison. *Personality and Social Psychology Bulletin, 27,* 38–47.

Mussweiler, T. (2001b). "Seek and ye shall find": Antecedents of assim-

ilation and contrast in social comparison. *European Journal of Social Psychology, 31,* 499–509.
Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110,* 472–489.
Mussweiler, T., & Bodenhausen, G. (2002). I know you are but what am I? Self-evaluative consequences of judging ingroup and outgroup members. *Journal of Personality and Social Psychology, 82,* 19–32.
Mussweiler, T., & Strack, F. (2000). The "relative self": Informational and judgmental consequences of comparative self-evaluation. *Journal of Personality and Social Psychology, 79,* 23–38.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 3,* 226–254.
Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–337). Hillsdale, NJ: Erlbaum.
Rüter, K., & Mussweiler, T. (2003). *Bonds of friendship—Comparative self-evaluations evoke the use of routine standards.* Unpublished manuscript, Universität Würzburg, Würzburg, Germany.
Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129–138.
Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology, 25,* 500–523.
Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 99–152). Hillsdale, NJ: Erlbaum.
Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73,* 437–446.
Suls, J., Gastorf, J. W., & Lawhon, J. (1978). Social comparison choices for evaluating a sex- and age-related ability. *Personality and Social Psychology Bulletin, 4,* 102–105.
Taylor, S. E. (1981). The interface of cognitive and social psychology. In J. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 182–211). Hillsdale, NJ: Erlbaum.
Taylor, S. E., Wayment, H. A., & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (pp. 3–27). New York: Guilford Press.
Trope, Y. (1986). Self-assessment and self-enhancement in achievement motivation. In R. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 1, pp. 350–378). New York: Guilford Press.
Tversky, A., & Kahneman, D. (1974, September). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1130.
Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behavior: Is habit an empty construct or an interesting case of goal-directed automaticity. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 10, pp. 101–134). Chichester, England: Wiley.
Verplanken, B., Aarts, H., van Knippenberg, A, & van Knippenberg, C. (1994). Attitude versus general habit: Antecedents of travel mode choice. *Journal of Applied Social Psychology, 24,* 285–300.
Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology, 2*(Suppl. 1), 27–31.
Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90,* 245–271.
Wilson, A. E., & Ross, M. (2000). The frequency of temporal self and social comparisons in people's personal appraisals. *Journal of Personality and Social Psychology, 78,* 928–942.
Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin, 106,* 231–248.
Wood, J. V., & Taylor, K. L. (1991). Serving self-relevant goals through social comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 23–49). Hillsdale, NJ: Erlbaum.
Zanna, M. P., Goethals, G. R., & Hill, J. F. (1975). Evaluating a sex-related ability: Social comparison with similar others and standard setters. *Journal of Experimental Social Psychology, 11,* 86–93.

Received October 31, 2002
Revision received April 10, 2003
Accepted April 17, 2003 ■

Correction to Greenwald et al. (2003)

The article "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm," by Anthony G. Greenwald, Brian A. Nosek, and Mahzarin R. Banaji (*Journal of Personality and Social Psychology,* 2003, Vol. 85, No. 2, pp. 197–216), contained several errors.

On page 203, the data lines in Figure 2 are incorrectly labeled. As in Figure 1, the line with filled squares as data points should be labeled MEAN, the line with filled diamonds as data points should be labeled MEDIAN, and the line with unfilled squares as data points should be labeled RECIPROCAL.