

## When Comparisons Arise

**Daniel T. Gilbert**

Department of Psychology University of Texas at Austin

**R. Brian Giesler**

Baylor College of Medicine

**Kathryn A. Morris**

Department of Psychology University of Texas at Austin

### ABSTRACT

People acquire information about their abilities by comparison, and research suggests that people restrict such comparisons to those whom they consider sources of diagnostic information. We suggest that diagnosticity is often considered only *after* comparisons are made and that people do not fail to make nondiagnostic comparisons so much as they mentally undo them. In 2 studies, participants made nondiagnostic comparisons even when they knew they should not, and quickly unmade them when they were able. These results suggest that social comparisons may be relatively spontaneous, effortless, and unintentional reactions to the performances of others and that they may occur even when people consider such reactions logically inappropriate.

---

This research was supported by a research grant (R01-MH49613) and a Research Scientist Development Award (KO2-MH00939) from the National Institute of Mental Health (NIMH). The support of NIMH is gratefully acknowledged.

We thank Brett Pelham and Abraham Tesser for their comments on an earlier version of this article, and Susan Brupbacher, Rebecca Pratt, Leah Shimatsu, and Michelle Skeen for their help with the execution of these experiments.

Correspondence may be addressed to Daniel T. Gilbert, Department of Psychology, University of Texas, Austin, Texas, 78712.

Electronic mail may be sent to [GILBERT@PSYVAX.PSY.UTEXAS.EDU](mailto:GILBERT@PSYVAX.PSY.UTEXAS.EDU)

Received: September 14, 1994

Revised: February 2, 1995

Accepted: February 2, 1995

---

If we say that a thing is great or small by its own standard of great or small, then there is nothing in all creation which is not great, nothing which is not small.

—Chuang Tzu (circa 400—300 BC), *Autumn Floods*

If we tried to describe the world in absolute terms we would probably find ourselves with little to say. Although some properties of objects can be considered absolutely (the numerosity of grapes, the mortality of people), many more can be considered only as relations. A haiku is short, a pizza is hot, and an African elephant is slightly lumpy, but only in the context of other poems, pies, and pachyderms. As it is with physical characteristics such as size and temperature, so it is with our own psychological characteristics such as our dispositions and capacities. Indeed, the things we most want to know about ourselves—Am I attractive? Am I smart? Am I likable?—can often be known only by comparing our thoughts, feelings, and behaviors with those of the people around us. We are intelligent and interesting precisely because others are so dim and dull.

Social psychologists have long been interested in these sorts of comparisons. [Festinger \(1954a\)](#) described the causes and consequences of social comparison and, in the four decades that followed, researchers refined and extended his theoretical effort, generated considerable empirical evidence for and against specific hypotheses, and ultimately came to agree on a number of fundamental truths. (For recent reviews see [Kruglanski & Mayselless, 1990](#) ; [Olson, Herman, & Zanna, 1986](#) ; [Suls & Wills, 1991](#) ; [Taylor, Buunk, & Aspinwall, 1990](#) ; [Wood, 1989](#) .) One of these fundamental truths is that people do not compare themselves with just anyone, but rather only with particular others. A professor may compare her salary with that of a colleague and feel angry, embarrassed, or overjoyed, but she is unlikely to compare her salary with a seven-year-old child's weekly allowance or to an oil baron's annual income. The first comparison is rich with diagnostic information, the subsequent comparisons are not, and because "people engage in social comparison for one reason, and that is to gain information" ( [Gibbons & Gerrard, 1991, p. 317](#) ), they seek the former and avoid the latter.

What is the evidence for this fundamental truth? How do we know that people avoid making nondiagnostic comparisons? The answer is simple. Social comparisons have both cognitive and affective consequences (e.g., [Brickman & Bullman, 1977](#) ; [Morse & Gergen, 1970](#) ; [Salovey & Rodin, 1984](#) ; [Taylor & Lobel, 1989](#) ; [Tesser, 1991](#) ; [Tesser, Millar, & Moore, 1988](#) ), and when people are provided with nondiagnostic comparison information they appear to experience neither of these. A professor who learns that a colleague of the same rank earns twice as much as she does may experience changes in her beliefs ("I must not be as good a teacher as he is") and her emotions ("I think I'll weep"), but a professor who learns that the president of Exxon earns several times her salary probably will not. This lack of cognitive and affective consequences is taken as evidence that the professor did not compare herself with the executive. When self-evaluations are unchanged by nondiagnostic comparison information, it is reasonably assumed that no comparison was made.

Although this assumption is indeed reasonable, we suspect that it is often wrong, and that what appears to be a failure to engage in social comparison may instead be a rapid and deliberate repudiation of its effects. We suggest that people cannot always avoid nondiagnostic comparisons, and that they often consider factors such as diagnosticity only after having made comparisons whose effects they must then work to reverse.

## Mental Control of Social Comparisons

[Festinger \(1954b\)](#) introduced his insights about social comparison this way: "We have made the following derivations about the conditions under which a social comparison process arises" (p. 217). Festinger did not say that he had determined when, how, or why people choose to compare themselves with others; rather, he specifically used language that made the process, and not the processor, the active agent—as though he was thinking of social comparison as a reaction to the environment rather than as an action played on it. Although later theorists have generally eschewed explicit claims about the degree to which the social comparison process is under the conscious control of the individual, the literature can sometimes give one the sense that comparisons are mental operations that people choose to perform or not to perform, rather than natural, effortless, or even inevitable reactions to the behavior of others. As [Wood \(1989, p. 232\)](#) correctly noted, "By emphasizing the individual's ability to select comparisons, the literature has largely reflected the view that the social environment is in the background," and thus the literature has given "little attention to situations in which the environment imposes comparisons on the individual." The not-so-tacit message is that social comparison is indeed a choice that people make rather than a reaction they have (cf. [Brickman & Bullman, 1977](#); [Taylor et al., 1990](#); [Wood, 1989](#)).

Such a view of social comparison is somewhat odd in light of what we know about the ease and spontaneity of other forms of mental comparison. [Pelham and Wachsmuth \(1994\)](#) noted that social comparisons are essentially contrast effects ("Tom's perfect score on the civil service exam made me feel particularly bad about mine"), and contrast effects in both the perceptual and cognitive domains often "occur at an early stage in cognitive processing, require minimal resources, and are therefore beyond the subject's control" ([Wedell, 1994, p. 1007](#)). On some occasions contrast effects do result from deliberate processing ([Martin, Seta, & Crelia, 1990](#); [Schwarz & Bless, 1992](#); [Wedell, 1990](#)), but there is little doubt that on other occasions they simply arrive unbidden. [Petty and Wegener \(1993\)](#) argued that such unwanted contrast effects can bedevil judgment and decision making (a job opportunity in Columbus may seem especially unappealing to an executive who had hoped to relocate to Paris) and that individuals must therefore make post hoc corrections for these unwanted influences ("The Columbus job is probably a great opportunity for me, and I shouldn't let pictures of the Eiffel Tower color my opinion").

Similar reasoning underlies a host of recent *correction models* which, taken together, suggest that people's thoughts and feelings may be uncontrollably influenced in a variety of undesirable ways and that people often exert control over their thoughts and feelings by "undoing" or correcting these undesirable effects after they have happened rather than by avoiding them in the first place (see [Gilbert, 1991](#); [Wilson & Brekke, 1994](#)). Because social comparisons are the products of more general inferential mechanisms ([Kruglanski & Mayseless, 1990](#)), it seems reasonable to suspect that, like many inferences, they may arise spontaneously and be undone subsequently and with effort. Although it may be logically inappropriate to compare one's backhand with a tennis pro's or one's standing in polite society with a prisoner's, it may be that such logical considerations affect our

thoughts and feelings only after the inappropriate comparisons have arisen. Unfortunately for the social comparison theorist, people who avoid comparisons and people who mentally undo them ultimately look alike; that is, neither appears to have been affected by the performances of others. The individual who decides that a friend's perfect score on a civil service exam is irrelevant to a reasonable assessment of his own ability will look exactly like the individual who never learned how his friend scored in the first place. So how can one tell whether a particular individual avoided making a comparison or made a comparison and then mentally undid its effects?

### **The Telltales of Correction**

Two research strategies have been particularly useful in distinguishing between people who have not changed their minds and people who have changed their minds and then quickly changed them back again. Imagine the simple case in which an office manager encounters a piece of information whose effects she would rather avoid—for example, a nasty rumor about a new employee. A correction model of belief (e.g., [Gilbert, Krull, & Malone, 1990](#)) suggests that hearing the rumor necessarily changes the manager's belief about her employee and that only after experiencing this change in belief may the manager remind herself that the rumor is unsubstantiated and thus make an effort to recover her original opinion. Although the manager may report the same opinion both before and after hearing the rumor, and may therefore appear to have been unaffected by it, there are at least two ways in which the researcher can demonstrate that belief change and correction did, in fact, take place in the interim.

First, the researcher may show that when the office manager's ability to correct her beliefs is impaired, the rumor-induced belief remains. For example, if the manager is mentally preoccupied with her quarterly reports at the moment she hears the rumor, then the likelihood that she will later hold a negative opinion of the employee should be substantially increased. This increase may be taken as evidence for a correction model of belief. Second, the researcher may show that even when the manager is ultimately able to correct her rumor-induced belief, she may still show signs of having briefly held it. For example, if the manager momentarily believed the rumor that the employee was a recovering ax murderer, she may have felt a transient rush of fear. Thoughts and beliefs can be changed quickly ("Mabel starts that silly ax-murderer rumor every time I hire a new typist. Probably no reason to believe it this time"), but emotions subside rather more slowly ("Gosh, my palms are all sticky"). Thus, the office manager's fear may linger measurably even after her belief has evaporated, and such lingering affect may also be taken as evidence for a correction model of belief.

In the following experiments we used these two general strategies to show that what appears to be a failure to make a nondiagnostic social comparison may actually be a success at correcting one. We attempted to show that people do indeed compare themselves with others even when they realize that such comparisons will be nondiagnostic, and that they then "uncompare" themselves if they can. In Experiment 1 we impaired our participants' ability to perform mental corrections by placing them under cognitive load, and we tried to show that the nondiagnostic comparisons they made

remained. In Experiment 2 we tried to show that even when participants are able to make mental corrections, they still experience the affective consequences of the nondiagnostic comparisons they made and then corrected.

## **Experiment 1**

### **Method Overview**

Female participants viewed an instructional videotape in which a female confederate was shown 18 pairs of photographs of target persons and was asked to determine which of the target persons in each pair was schizophrenic. The confederate either performed poorly (i.e., she responded correctly on 4 trials) or performed well (she responded correctly on 16 trials). When the confederate did well, participants were told that she had been given instructions about how to detect schizophrenia prior to her videotaped performance. When the confederate did poorly, participants were told that she had been deliberately misled about how to detect schizophrenia prior to her videotaped performance. After viewing the videotape, half the participants were made cognitively busy with a digit-rehearsal task, and the remaining participants were not. All participants then performed the schizophrenia detection task themselves and were given bogus feedback indicating that they had responded correctly on 10 of the 18 trials. Participants then rated their own and the average student's competence at the task.

### **Participants**

Participants were 68 female undergraduates at the University of Texas at Austin who took part in exchange for credit in their introductory psychology course.

### **Procedure The schizophrenia detection task.**

On arrival at the laboratory, participants were greeted by a male experimenter who explained that he was studying the effects of concentration on pattern learning and that the participant would be asked to perform a "schizophrenia detection task" under conditions of either full or impaired concentration. This task was similar to one developed by [Aronson and Carlsmith \(1962\)](#).

The experimenter explained that the schizophrenia detection task required the participant to view 18 pairs of photographs, 1 pair at a time. Each of the 36 photographs showed the face and upper torso of a male or female undergraduate at the University of Texas at Austin who had previously agreed to allow his or her photograph to be used in the experiment. The photographs in each pair were matched for gender and roughly matched for general appearance and attractiveness. The experimenter explained that one target person in each pair of photographs had a normal psychiatric history and that the other target person had been diagnosed as schizophrenic or had a family member who had been so diagnosed and thus was at risk for schizophrenia. The experimenter explained that on each trial he would hold one photograph in each hand and ask the participant to point to the photograph of the target person whom she felt was most likely to have (or be at risk

for) schizophrenia, at which time the experimenter would indicate whether or not the participant had chosen correctly. Ostensibly, participants would learn by trial and error to detect the subtle patterns of facial features that indicated the presence of schizophrenia.

### **Manipulation of cognitive busyness.**

The experimenter explained that to discover how concentration affected pattern learning, he would ask some participants to perform the schizophrenia detection task while rehearsing an 8-digit number. Half the participants were randomly assigned to the busy condition, were shown an 8-digit number just prior to performing the task, and were told to keep the number in mind throughout the task and to be prepared to report the number when the task was finished. The remaining participants were assigned to the nonbusy condition and were not given a number to rehearse.

### **Manipulation of confederate's performance.**

Before participants performed the task, they were shown a short "instructional" videotape ostensibly to illustrate the task procedures. The videotape showed the experimenter and a female confederate performing the schizophrenia detection task. Participants were told that the confederate in the videotape had been a participant in a previous experiment and that she was performing precisely the same task that the present participant would be asked to perform. Participants in the good-performance condition saw the videotaped confederate give correct responses on 16 of 18 trials, and participants in the poor-performance condition saw the videotaped confederate give correct responses on 4 of the 18 trials.

Participants in the good-performance condition were told that, prior to performing the task, the videotaped confederate had been taught to recognize the pattern of facial features that indicated schizophrenia, and thus they should expect her performance to be particularly good. Participants in the poor-performance condition were told that, prior to performing the task, the videotaped confederate had been purposely misled about the pattern of facial features that indicated schizophrenia, and thus they should expect her performance to be particularly poor.

It is important to note that participants in the busy condition saw a version of the instructional videotape in which the experimenter gave the confederate an 8-digit number to rehearse as she performed the schizophrenia detection task. This was done so that busy participants would not conclude that their own score differed from the confederate's score because they (but not the confederate) had performed the task under difficult conditions. As such, participants in both the busy and nonbusy conditions saw the confederate perform the schizophrenia detection task under precisely the same conditions that they would encounter themselves.

After seeing the videotaped confederate perform poorly or well, and after receiving or not receiving an 8-digit number to rehearse, all participants performed the schizophrenia

detection task. Participants were given feedback indicating that they had given the correct response on 10 of the 18 trials. Participants then completed the dependent measures.

### **Dependent Measures**

Participants were asked to recall their own and the videotaped confederate's score on the schizophrenia detection task. Participants were then asked to estimate their own and the average student's ability to detect people at risk for schizophrenia on a 7-point scale that was anchored at the endpoints with the phrases *detect people at risk very poorly* and *detect people at risk very well*. At the conclusion of the experimental session, participants were probed for suspicion and understanding of basic information, thoroughly debriefed, thanked, and dismissed.

### **Results and Discussion Excluded Data**

Two participants reported misunderstanding the instructions, 1 participant misremembered the confederate's score by more than two correct responses, and 4 participants did not recall or reported not understanding that the confederate had been given training or had been misled about how to detect schizophrenia. Data from these 7 participants were excluded from all analyses.

### **Self-Perceived Competence**

Did participants compare themselves with a confederate whose performance was known to have been shaped by prior training? The difference between a participants' rating of her own ability and her rating of the average student's ability was taken as an index of her self-perceived competence.<sup>1</sup> We submitted this index to a 2 (confederate's performance: good or poor)  $\times$  2 (busyness: busy or nonbusy) analysis of variance (ANOVA), which revealed a main effect of confederate's performance,  $F(1, 60) = 15.19, p < .0003$ . This effect was qualified by the predicted Confederate's Performance  $\times$  Busyness interaction,  $F(1, 60) = 4.32, p < .04$ . As [Table 1](#) shows, the self-perceived competence of nonbusy participants was not reliably affected by the confederate's performance,  $F(1, 60) = 1.81, p = .18$ , but the self-perceived competence of busy participants was affected,  $F(1, 60) = 16.47, p < .0002$ . It is important to keep in mind that busy participants learned that the confederate's performance was nondiagnostic *before* they became busy and before they performed the task themselves. As such, they were just as aware of the nondiagnosticity of the performance as were nonbusy participants, who demonstrated that awareness by using it later to correct their inferences.

Finally, an inspection of [Table 1](#) reveals that the statistical interaction was driven largely by the condition in which the confederate performed poorly. This suggests either that ratings in the nonbusy condition are already at the psychological floor (i.e., participants simply will not rate themselves as much less capable than the average person) or that busyness caused a slight and inexplicable elevation of ratings. In either case, the important point is that the ratings of busy participants were reliably affected by changes in the confederate's performance, whereas the ratings of nonbusy participants were not.

## Exactly What Does Busyness Do?

Although all participants seemed to realize that the confederate's prior training made her a source of nondiagnostic comparison information, busy participants were apparently unable to use this fact to undo the comparison they seemed unable to avoid making. This finding is consistent with our contention that the diagnosticity of a comparison may be taken into account only after a comparison is made.

The interpretation of busyness effects, however, is always tricky business. Past research suggests that busyness truncates sequential mental operations, or "stops people in their mental tracks," so to speak (e.g., [Gilbert & Hixon, 1991](#) ; [Gilbert, Pelham, & Krull, 1988](#) ; [Kantowitz, 1974](#) ; [Martin, Seta, & Crelia, 1990](#) ). As such, busy participants are generally thought of as having "gotten stuck" at a stage through which nonbusy subjects have successfully passed, and their reports are thought to be identical to the reports that nonbusy participants *would have made* had it been possible to interrogate them in the moments before they changed their minds. Although this is a viable interpretation, it is not the only viable interpretation. One might argue, for example, that busyness did not truncate the information processing sequence that nonbusy participants normally follow, but rather that it caused busy participants to follow an entirely different sequence. In other words, busyness may have caused participants to make comparisons that they would not have made had they not been busy. If busyness gives rise to an abnormal information processing sequence rather than simply interrupting a normal information processing sequence, then the reports of busy participants may say little about the mental life of nonbusy participants.

The busyness technique simply cannot overcome these ambiguities, and thus we felt it important to generate converging evidence for our hypothesis by using a very different technique. As we mentioned earlier, cognitions are often fleeting, but they may leave affective traces that take somewhat longer to decay. Because emotions have a greater half-life than do thoughts, they may be taken as relatively enduring evidence of the transient cognitions that generated them, just as fossil imprints may be taken as evidence of the transient life forms that left them behind. As [Tesser \(1991, p. 116\)](#) noted, emotion "serves as a marker. Its presence provides information about whether social comparison and reflection processes are engaged." In Experiment 2 we attempted to find the affective fossils left by very brief social comparisons. The basic plan was to expose nonbusy participants to diagnostic or nondiagnostic social comparison information and then measure their cognitive and affective reactions. We expected that cognitive measures would suggest that nonbusy participants had used the diagnostic information and ignored the nondiagnostic information but that affective measures would clearly show that in both cases they had, in fact, made social comparisons.

## Experiment 2

### Method Overview

Female participants were connected by electrode to a machine that could ostensibly detect the valence of their affect. Participants were shown 18 pairs of photographs and were asked to determine which of the male target persons in each pair was expressing insincere emotions. Participants were asked to report the valence of their affect at several points during the task. After performing the task, participants watched a female confederate perform the task. Participants were then told that they had responded correctly on 10 of 18 trials and that the confederate had responded correctly on either 16 of 18 trials or 4 of 18 trials. Participants knew that the confederate was either performing a different task than the participant had performed or the same task that the participant had performed. Finally, all participants rated their ability and the average woman's ability to detect the sincerity of men's emotional expressions.

## **Participants**

Participants were 53 female undergraduates at the University of Texas at Austin who participated in exchange for credit in their introductory psychology course.

## **Procedure**

Participants arrived at the laboratory, where they met another female college student (a confederate) who had ostensibly arrived a bit earlier. The experimenter told the participant and confederate that the experiment consisted of two unrelated tasks. The primary task involved reading "insincerity in men's emotional expressions," and the secondary task involved "validating some new lie-detection machinery."

## **Emotion-reading task.**

The participant and confederate were told that the purpose of the primary task was to determine how well women can distinguish between men's sincere and insincere emotional expressions. The experimenter explained that the participant and confederate would view a series of photographs of men who had ostensibly undergone a mood induction procedure in which they had been made to feel either happy or sad. The experimenter had then ostensibly instructed each man to smile for the camera as she photographed him. Thus, the men who were in good moods as a result of the mood induction procedure were ostensibly expressing sincere emotions when they smiled for the camera, whereas the men who were in bad moods as a result of the mood induction procedure were ostensibly expressing insincere emotions when they smiled for the camera. (In fact, the photographs used for the emotion-reading task were of male college students who had not undergone any mood induction procedure and who had simply been asked to smile. These students all gave permission for their photographs to be used in later experiments.) The participant and the confederate were told that they would be shown 18 pairs of photographs and that in each pair one man would be expressing sincere emotions and one man would be expressing insincere emotions. They were told that their primary task was to view each pair of photographs and determine which of the two men pictured was expressing insincere emotions (i.e., which one was smiling while actually feeling sad).

### **Manipulations of confederate's performance.**

The participant performed the emotion-reading task and then watched the confederate perform the task. After both the participant and the confederate had performed the task, the experimenter announced the participant's score and then the confederate's score. All participants were told that they had responded correctly on 10 of the 18 trials.

Half the participants were randomly assigned to the same-task condition. These participants and the confederate were told that the confederate would be shown the same photographs that the participant had just seen and would thus perform the same emotion-reading task that the participant had just performed. Half these participants saw the confederate perform well (i.e., they heard the experimenter tell the confederate that she had responded correctly on 16 of 18 trials), and half saw the confederate do poorly (i.e., they heard the experimenter tell the confederate that she had responded correctly on only 4 of 18 trials).

The remainder of the participants were assigned to the different-task condition. These participants and the confederate were told that the confederate would be shown the same photographs that the participant had seen but that in each photograph the left side or the right side of the man's face would be covered by a piece of paper. The experimenter explained that prior research had shown that all men have a "strong and a weak side to their faces" and that the confederate would be seeing only the strong side of each man's face. Half the participants in the different-task condition were told that women who see only the strong side of a man's face tend to make very few errors on the emotion-reading task (ostensibly because false cues from the weak side are not present to confuse them). These participants then saw the confederate perform well (i.e., they heard the experimenter tell the confederate that she had responded correctly on 16 of 18 trials). The remaining participants in the different-task condition were told that women who see only the strong side of a man's face tend to make many errors on the emotion-reading task (ostensibly because useful cues from the weak side are not present to help them). These participants then saw the confederate perform poorly (i.e., they heard the experimenter tell the confederate that she had responded correctly on only 4 of 18 trials). In short, participants in the different-task condition saw the confederate perform poorly or well, and in both cases the confederate's performance was easily explained by the relative difficulty or ease of the task that the confederate had been assigned to perform.

### **Equipment-validation task.**

The purpose of this experiment was to determine whether participants' thoughts and feelings were affected by the confederate's performance. We expected that participants in the different-task condition would recognize that social comparisons would be nondiagnostic under these circumstances and thus might be particularly reluctant to admit that their affective states had been influenced by the confederate's performance. As such, we employed a bogus pipeline procedure ( [Jones & Sigall, 1971](#) ) to increase the veracity of participants' self-reports (see [Roese & Jamieson, 1993](#) , for a review of this technique).

Prior to performing the detection task, the experimenter pointed to an aggregation of interconnected electronic machinery (e.g., a computer, signal light, transfer box, wires and cables, etc.). The experimenter told the participant and confederate that this was a new, sophisticated device that could reliably detect the valence of a person's current affective state. The participant and confederate were asked to participate in a "validation trial" that would help confirm the accuracy of the affect-detection device. The experimenter pointed to a signal light that was visibly connected to the affect-detection device and told the participant and confederate that at several times during the detection task the signal light would flash and that when it did, they were to report their current affective state. They were told that these written reports would later be compared with the affect-detection device's readings and that this would enable the experimenter to validate the accuracy of the affect-detection device. Before the participant and confederate performed the emotion-reading task, the experimenter cleaned their middle and index fingers with rubbing alcohol and fastened electrodes to those fingers with Velcro straps. These electrodes were visibly attached to the affect-detection device.

### **Summary of procedure.**

Participants were connected to a bogus affect-detection device and were signaled to report their current affective states at several points throughout the experiment. Participants performed the emotion-reading task and then watched the confederate perform either the same task or a different task. The experimenter then announced both the participant's score (10 of 18 correct) and the confederate's score (either 16 of 18 correct or 4 of 18 correct). Finally, participants were asked to complete a questionnaire that contained several dependent measures.

### **Dependent Measures**

The signal light flashed 20 times throughout the experiment, and thus the participant made 20 written reports of her current affective state. Each report was made on a 9-point Likert-type scale that was anchored at the endpoints with the words *positive* and *negative*. Two of these reports were considered critical. The first critical report was made immediately after the participant learned how she had performed on the primary task. We refer to this as the *self-feedback measure*. The second critical report was made immediately after the participant learned how the confederate had performed on the primary task. We refer to this as the *other-feedback measure*.

After the participant and confederate had completed the emotion-reading task and received their scores, the experimenter asked them to complete a questionnaire. The participant's questionnaire asked her to recall her own score and the confederate's score, to rate her own ability to detect sincerity in men's faces (on a 9-point scale anchored with the phrases *very bad* and *very good*), and to rate the average University of Texas student's ability to detect sincerity in men's faces (on the same scale). After completing the questionnaire, the participant and confederate were instructed to remove the electrodes from their fingers. The participant was probed for suspicion, debriefed, thanked, and dismissed.

## Results and Discussion Excluded Data

Of the 53 participants who took part in the experiment, 2 participants reported extreme suspicion, 1 participant was unable to understand the instructions, 1 participant was unable to concentrate on the task, and 1 participant's age made her inappropriate for the experiment. The data from these 5 participants were excluded from all analyses, leaving 48 participants in the data set. <sup>2</sup>

## Recall of Scores

All participants correctly recalled their own scores and the confederate's score.

## Self-Perceived Competence

Did participants compare themselves with a confederate whose performance was known to have been a function of the ease or difficulty of the task she had been assigned? The difference between a participant's rating of her own ability and her rating of the average student's ability was taken as an index of her self-perceived competence. We submitted this index to a 2 (confederate's performance: poor or good)  $\times$  2 (confederate's task: same or different) ANOVA. The analysis revealed a main effect of confederate's performance,  $F(1, 44) = 7.91, p < .01$ , but this effect was qualified by the predicted Confederate's Performance  $\times$  Confederate's Task interaction,  $F(1, 44) = 5.93, p < .02$ . As [Table 2](#) shows, participants' self-perceived competence was affected by the confederate's performance when the confederate performed the same task that the participant had performed,  $F(1, 44) = 15.09, p < .001$ , but not when she performed a different task than the participant had performed ( $F < 1$ ).

## Change in Affective State

Ratings of self-perceived competence suggest that participants did not engage in social comparison when the confederate's performance was determined by the ease or difficulty of her task. If participants in the different-task condition did not engage in social comparison, then their affective state should not have been influenced by the confederate's performance. The influence of the confederate's performance on the participant's affective state was operationalized as the difference between the self-feedback measure (the participant's report of her affective state after she received feedback about her performance) and the other-feedback measure (the participant's report of her affective state after she heard the confederate receive feedback about the confederate's performance). We submitted this index to a 2  $\times$  2 ANOVA (as above) that revealed only the predicted main effect of confederate's performance,  $F(1, 44) = 14.44, p < .001$ . As [Table 2](#) shows, participants experienced more positive changes in their affective states when the confederate did poorly than when she did well. Importantly, this was true in both the same-task condition,  $F(1, 44) = 6.38, p < .02$ , and the different-task condition,  $F(1, 44) = 8.06, p < .001$ . Clearly, participants in the different-task condition were affected by the performance of the confederate, despite their claims of immunity. It is worth noting that, by the end of the experiment, the affective consequences of social

comparison had entirely faded. Analysis of the participants' last affective ratings showed absolutely no effects of any manipulation (all  $F$ 's  $< 1$ ).

## General Discussion

If an insurance salesman from Wichita compares himself with Charles Manson, he will probably learn little about his own moral character. Although ordinary people seem to realize that such comparisons are not worth the bother, our research suggests that making them may not be much of a bother at all. In fact, social comparisons can sometimes be so natural and easy that people may make them even when they don't really want to, and when that happens, they may have little choice but to mentally undo the comparisons they made. Such efforts are not always successful, and when they are, the unwanted comparisons may continue to have emotional effects even after their cognitive effects have been reversed. In short, sometimes we choose to compare ourselves with others, but sometimes such comparisons are thrust upon us. When they are, we may not be able to escape their unwanted influence.

This view of the social comparison process raises three important questions. First, if social comparisons are relatively automatic responses to the performances of others, does this mean that people have no choice about which comparisons they will or won't make? Second, when people are not consciously choosing to compare themselves with one person rather than another, then what factors will determine the object of their comparisons? And third, if comparison is a relatively automatic response to the performance of another, then what does this say about other responses, such as reflection? We address each issue in turn.

### The Role of Choice in Social Comparison

Although our work suggests that the decision to compare or not to compare is not always ours to make, conscious choice does play an important role in a correction model of social comparison. Correction models are based on the assumption that people lack complete control over their processing of information—in other words, the mind often responds to information in ways that the mind's owner might wish it would not. But a lack of complete control is not a complete lack of control. Indeed, even a person who spontaneously compares himself with those whom he encounters can control his conclusions by (a) choosing which others to encounter, and (b) choosing which conclusions to revise. Perhaps a small-town mayor cannot avoid comparing her political stature with that of the President when she meets him, but she may be able to avoid meeting him, and if not, then she may be able to correct her ill-advised comparison after it has been made. Even when people are unable to control directly their processing of information, the two strategies of *exposure control* (determining the information to which one will be exposed in the future) and *unbelieving* (consciously repudiating the conclusions that one has reached in the past) allow them exceptional authority over the contents of their own minds ( [Gilbert, 1992](#) ).

These two strategies are key to understanding how a correction model of the social comparison process fits with more traditional conceptualizations. Previous work has been largely concerned with articulating the "rules" of social comparison—that is, describing the factors that determine when and with whom people consciously choose to compare—and these rules constitute the naive psychology of social comparison. Although our data suggest that such rules do not necessarily determine when people make mental comparisons, we suspect that they do indeed determine when people will engage in the more deliberate strategies of exposure control and unbelieving. For example, one of the most well-established social comparison rules is that people who are uncertain of their standing on an ability-linked dimension (and who wish to gain accurate, rather than merely flattering, information about themselves) tend to think of dissimilar others as nondiagnostic sources of comparison information and thus do not compare with them. Although our research suggests that this "similarity rule" is sometimes wrong inasmuch as people *do*, in fact, compare with dissimilar others, we suspect it is quite right inasmuch as accuracy-driven people probably do not seek encounters with dissimilar others (exposure control) and, when such encounters occur, may repudiate the thoughts and feelings that the encounter produced (unbelieving). In short, the well-established rules of social comparison may not predict when we will compare with those whom we encounter, but they probably do predict whom we will encounter and whether we will work to undo the effects of that encounter after it has taken place. Our conceptualization represents a departure from previous conceptualizations of the social comparison process, but it does not do so by rejecting previous work but rather by relocating it.

### **The Object of Comparison**

If social comparisons were thoroughly reflexive responses to the real or simulated presence of others, then people would mentally compare themselves with every person whom they encountered or imagined on a particular day. Given the sizable number of people with whom we have some form of social commerce, it seems logically absurd to suggest that social comparison is an inevitable consequence of observing another's performance. What, then, determines which performances will spark a comparison and which will not?

This problem in the study of social comparison is very much like the "context problem" in the study of psychophysical contrast. A librarian may consider *Totem and Taboo* especially light if he moved *The Rise and Fall of the Roman Empire* in order to reach it, but why was the first book judged in the context of the second book rather than in the context of books lifted hours before, groceries lifted last month, grandchildren lifted last year, or for that matter, every object ever lifted during the librarian's lifetime? Decades of psychophysical research suggest some answers to this question. Contextual stimuli are most likely to become the objects of comparison when (a) they have been recently encountered, (b) they are explicitly judged, and (c) their values are especially extreme ( [Parducci, 1992](#) ; [Parducci & Wedell, 1990](#) ). In other words, the librarian is more likely to experience contrast when he lifts the heavy book just moments before the lighter one, when he explicitly mutters, "That book is the heaviest damn thing I've lifted today," and when the book *is*, in fact, the heaviest damn thing he lifted today.

These three principles may provide clues about when social comparisons will and will not arise. In our studies, the confederate's performance was in close spatial and temporal proximity to the participant's, it was the most extreme performance the participant had ever observed, and its value was explicitly judged. As noted, these are the very factors that encourage contextual effects. Indeed, we would have been surprised if our participants had spontaneously compared themselves with a confederate whose moderate performance they had incidentally witnessed a year earlier. We suspect that these three factors—recency, extremity, and explicit judgment—created just the sort of climate in which comparisons are most likely to arise. Our research demonstrates that under some circumstances comparisons may arise without conscious effort, but future research must determine when they will and when they will not arise. These three principles seem to provide a reasonable point of departure.

### **Reflection and Correction**

As every proud parent knows, another person's performance can evoke reactions other than social comparison. When a close other's performance provides evidence of abilities that are irrelevant to an observer's own cherished identity, the self-evaluation maintenance model ( [Tesser & Campbell, 1983](#) ) suggests that observers engage in *reflection* rather than comparison and that they react to the other's performance as if it were their own ( [Tesser, 1984](#) ). Indeed, it would be a strange mother who was steaming rather than beaming when her daughter won first prize in the third grade math competition, and the odd coach who felt smugly self-satisfied when his star sprinter placed last in the 50-yard dash. A variety of experiments have shown that when those with whom we share special bonds perform in domains that are not relevant to our own sense of self-worth, we share in the thrill of their victories and the agony of their defeats ( [Tesser, 1986](#) ; [1991](#) ). How does our theorizing relate to this important work?

The self-evaluation maintenance model is a dual-process model. In other words, comparison and reflection are thought to be different processes that produce different cognitive and emotional responses, and the closeness of the other and the relevance of the other's behavior determine which of these processes is (or is most strongly) evoked. Of course, just because people show different responses does not mean that these responses are necessarily produced by different processes. Correction models provide an alternative to dual-process models by suggesting that different responses are sometimes the result of a single, multistage process, and that what appear to be the products of two separate processes may actually be the initial and updated outputs of one. This suggests that, at least in some cases, the jealousy of comparison and the pride of reflection could be the early and late results of a single psychological operation. Just as participants in our studies presumably corrected their spontaneous comparisons by reminding themselves that they should not have been affected by another's performance ("I feel pretty stupid. But then again, the other participant had prior training at schizophrenia detection, so I guess I shouldn't feel stupid after all"), it is conceivable that under other circumstances correction might involve reminding oneself that one *should* have been affected by another's performance, but in a different way than one was affected ("Gosh, Fred's Academy Award makes me feel jealous. But I'm a plumber, not a producer, so I guess I

should feel proud to be Fred's brother"). In this case, the sense of self-worth and feelings of pride that appear to be the products of a reflection process are actually the products of a corrected comparison.

We do not wish to suggest that *all* reflections are merely revisions of comparisons, but it may be useful to distinguish between two kinds of reflections—those produced when one regards another's irrelevant behavior as an extension of one's own and thus initially has positive reactions to success and negative reactions to failure, and those produced when one regards another's relevant behavior as belonging to the other, initially compares the other's performance with one's own, experiences positive reactions to the other's failure and negative reactions to the other's success, and finally, engages in the cognitive work necessary to reverse or undo these early reactions. In both cases one may ultimately take pride in the other's achievements and suffer the humiliation of the other's failures, but such conclusions may be achieved by distinctly different means. It is not unreasonable to suppose that the latter sort of reflection may produce weaker changes in and greater discontinuity between cognition and emotion. These issues are surely worthy of investigation.

## Coda

[Bertrand Russell \(1930\)](#) suggested that envy is one of the fundamental causes of human suffering, and he counseled his readers to avoid it:

The habit of thinking in terms of comparison is a fatal one ... which consists in seeing things never in themselves but only in their relations. ... You can get away from envy by enjoying the pleasures that come your way, by doing the work that you have to do, and by avoiding comparisons with those whom you imagine, perhaps quite falsely, to be more fortunate than yourself" (pp. 87—88).

Is it possible to break the fatal habit of social comparison? [Russell \(1930\)](#) thought so. After all, he argued, "Beggars do not envy millionaires, though of course they will envy other beggars who are more successful" (p. 90). In other words, because people clearly do not compare themselves with everyone they must therefore be capable of controlling their comparisons and hence of avoiding the envy that informs their daily misery. Alas, the fact that comparisons are less than ubiquitous does not mean they are under the voluntary control of the individual. We have argued that people can *indirectly* control their thoughts and feelings by avoiding the situations that give rise to unwise comparisons or by undoing those comparisons once they have been made. But our studies also suggest that when ordinary people are deprived of such indirect control—and when certain environmental factors that encourage comparison are in place—they may well experience the very comparisons they wish to avoid. A life without envy would itself be an envious life, but our research suggests that the injunction to live one may be somewhat impractical advice.

## References

- Aronson, E. & Carlsmith, J. M. (1962). Performance expectancy as a determinant of actual performance. (*Journal of Abnormal and Social Psychology*, 65, 178—182.)
- Brickman, P. & Bullman, R. J. (1977). Pleasure and pain in social comparison. (In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 149—186). Washington, DC: Hemisphere.)
- Festinger, L. (1954a). A theory of social comparison processes. (*Human Relations*, 7, 117—140.)
- Festinger, L. (1954b). Motivation leading to social behavior. (In M. R. Jones (Ed.), *Nebraska symposium on motivation* (Vol. 2, pp. 191—218). Lincoln: University of Nebraska Press.)
- Gibbons, F. X. & Gerrard, M. (1991). Downward comparison and coping with threat. (In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. (pp. 317—346). Hillsdale, NJ: Erlbaum.)
- Gilbert, D. T. (1991). How mental systems believe. (*American Psychologist*, 46, 107—119.)
- Gilbert, D. T. (1992). The assent of man: Mental representation and the control of belief. (In D. M. Wegner & J. Pennebaker (Eds.), *The handbook of mental control* (pp. 57—87). New York: Prentice-Hall.)
- Gilbert, D. T. & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. (*Journal of Personality and Social Psychology*, 60, 509—517.)
- Gilbert, D. T., Krull, D. S. & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. (*Journal of Personality and Social Psychology*, 59, 601—613.)
- Gilbert, D. T., Pelham, B. W. & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. (*Journal of Personality and Social Psychology*, 54, 733—740.)
- Jones, E. E. & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. (*Psychological Bulletin*, 76, 349—364.)
- Kantowitz, B. H. (1974). Double stimulation. (In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 83—132). Hillsdale, NJ: Erlbaum.)
- Krantz, D. L. & Campbell, D. T. (1961). Separating perceptual and linguistic effects of context shifts upon absolute judgments. (*Journal of Experimental Psychology*, 62, 35—42.)
- Kruglanski, A. W. & Mayseless, O. (1990). Classic and current social comparison research: Expanding the perspective. (*Psychological Bulletin*, 108, 195—208.)
- Martin, L. L., Seta, J. J. & Crelia, R. A. (1990). Assimilation and contrast as a function of people's willingness and ability to expend effort in forming an impression. (*Journal of Personality and Social Psychology*, 59, 27—37.)
- Morse, S. & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. (*Journal of Personality and Social Psychology*, 40, 624—634.)
- Olson, J. M., Herman, C. P. & Zanna, M. P. (Eds.) (1986). *Relative deprivation and social comparison: The Ontario symposium* (Vol. 4). (Hillsdale, NJ: Erlbaum)
- Parducci, A. (1992). Elaborations upon psychophysical contexts for judgment: Implications of cognitive models. (In H. G. Geissler, S. W. Link, & J. T. Townsend

- (Eds.), *Cognition, information processing, and psychophysics: Basic issues*. (pp. 207—224). Hillsdale, NJ: Erlbaum.)
- Parducci, A. & Wedell, D. (1990). The context of evaluative judgments: Psychophysics and beyond. (In H. H. Geissler (Ed.), *Psychophysical explorations of mental structures*. (pp. 94—103). Göttingen, Germany: Hogrefe & Huber.)
- Pelham, B. W. & Wachsmuth, J. O. (1994). *Feeling tall in the land of giants: A theory of social assimilation processes*. (Unpublished manuscript, University of California at Los Angeles)
- Petty, R. E. & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. (*Journal of Experimental Social Psychology*, 29, 137—165.)
- Roese, N. J. & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. (*Psychological Bulletin*, 114, 363—375.)
- Russell, B. (1930). *The conquest of happiness*. (New York: Liveright)
- Salovey, P. & Rodin, J. (1984). Some antecedents and consequences of social comparison jealousy. (*Journal of Personality and Social Psychology*, 47, 780—792.)
- Schwarz, N. & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. (In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217—245). Hillsdale, NJ: Erlbaum.)
- Suls, J. & Wills, T. A. (Eds.) (1991). *Social comparison: Contemporary theory and research*. (Hillsdale, NJ: Erlbaum)
- Taylor, S. E., Buunk, B. P. & Aspinwall, L. G. (1990). Social comparison, stress, and coping. (*Personality and Social Psychology Bulletin*, 16, 74—89.)
- Taylor, S. E. & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. (*Psychological Review*, 96, 569—575.)
- Tesser, A. (1984). Self-evaluation maintenance processes: Implications for relationships and development. (In J. C. Masters & K. Yarkin-Levin (Eds.), *Boundary areas in social and developmental psychology* (pp. 271—299). San Diego, CA: Academic Press.)
- Tesser, A. (1986). Some effects of self-evaluation maintenance on cognition and action. (In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (pp. 435—464). New York: Guilford Press.)
- Tesser, A. (1991). Emotion in social comparison and reflection processes. (In J. Sulz & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. (pp. 117—148). Hillsdale, NJ: Erlbaum.)
- Tesser, A. & Campbell, J. (1983). Self-definition and self-evaluation maintenance. (In J. M. Sulz & A. Greenwald (Eds.), *Social psychological perspectives on the self* (Vol. 2, pp. 1—31). Hillsdale, NJ: Erlbaum.)
- Tesser, A., Millar, M. & Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. (*Journal of Personality and Social Psychology*, 54, 49—61.)
- Wedell, D. H. (1990). Methods for determining the locus of context effects in judgment. (In J. P. Caverni, J. M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 285—302). New York: Elsevier Science.)
- Wedell, D. H. (1994). Contextual contrast in evaluative judgments: A test of pre- versus postintegration models of contrast. (*Journal of Personality and Social Psychology*, 66,

1007—1019.)

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

Wilson, T. D. & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. (*Psychological Bulletin*, 116, 117—142.)

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

## 1

We considered it important to measure self-perceived competence as a perceived deviation of the self from the norm. Observing a good or a poor performance can affect a participant's ratings by affecting her interpretation of particular scale values rather than by changing her self-perceived competence (see [Krantz & Campbell, 1961](#)), but this problem is virtually eliminated when one defines self-perceived competence not simply as "a high scale value" but as "a higher scale value than one gives the average person." There is ambiguity in the statement "A Ford is a good car" but not in the statement "A Ford is better than a Chevrolet." Ratings of both the self and the average person are shown in [Table 1](#) but, by our logic, it is the difference between these ratings and not the component ratings themselves that is critical.

## 2

The excluded participant was 36 years old (8.92 sigma units from the mean participant age of 18.83 years) and was thus considered a demographic outlier. Including her data would not change the pattern or significance level of any result.

Table 1  
*Self-Perceived Competence as a Function of Confederate  
 Performance and Cognitive Busyness in Experiment 1*

Rating of	Confederate's performance		
	Good	Poor	Difference
		Busy	
Self	2.46	3.23	-0.77
Average person	2.60	2.00	0.60
Self-perceived competence	0.14	1.23	-1.37 <sup>a</sup>
		Not busy	
Self	2.63	3.00	-0.37
Average person	2.69	2.65	0.04
Self-perceived competence	-0.06	0.35	-0.41

*Note.* Self-perceived competence was computed by subtracting the rating of the average person from the rating of the self. Larger values indicate greater self-perceived competence.

<sup>a</sup> Significant at  $p < .05$  or less by simple effects tests.

Table 2  
*Self-Perceived Competence and Changes in Affective State  
as a Function of Confederate's Task and Performance in  
Experiment 2*

Dependent measure	Confederate's performance		
	Good	Poor	Difference
	Same task		
Rating of self	4.69	6.31	-1.62*
Rating of average person	5.84	5.38	0.46
Self-perceived competence	-1.15	0.93	-2.08*
Other feedback measure (Time 2)	4.69	6.30	-1.61*
Self-feedback measure (Time 1)	5.31	5.77	0.46
Change in affective state	-0.62	0.53	-1.15*
	Different task		
Rating of self	5.50	5.17	0.33
Rating of average person	5.40	4.92	0.48
Self-perceived competence	0.10	0.25	-0.15
Other feedback measure (Time 2)	5.30	5.58	-0.28
Self-feedback measure (Time 1)	5.80	4.67	1.13*
Change in affective state	-0.50	0.91	-1.41*

*Note.* Self-perceived competence was computed by subtracting the rating of the average person from the rating of the self. Changes in affective state were computed by subtracting the self-feedback measure from the other-feedback measure. Higher values indicate (a) greater self-perceived competence or (b) changes toward more positive affect states over time.

\* Significant at  $p < .05$  or less by simple effects tests.