How Chronic Self-Views Influence (and Potentially Mislead) Estimates of Performance

Joyce Ehrlinger and David Dunning
Cornell University

An important source of people's perceptions of their performance, and potential errors in those perceptions, are chronic views people hold regarding their abilities. In support of this observation, manipulating people's general views of their ability, or altering which view seemed most relevant to a task, changed performance estimates independently of any impact on actual performance. A final study extended this analysis to why women disproportionately avoid careers in science. Women performed equally to men on a science quiz, yet underestimated their performance because they thought less of their general scientific reasoning ability than did men. They, consequently, were more likely to refuse to enter a science competition.

How'm I doin'?

Ed Koch, Mayor of New York City, 1978–1989

The knowledge of whether one is succeeding or failing at a task can be a precious piece of information. The degree to which individuals think that they performed well often has substantial consequences for the actions they take and the outcomes they encounter. A young person's impression of how well he or she did on a first date can determine if he or she pursues the relationship. A law student's impression of how well he has answered practice questions for a bar exam will determine whether he continues to study. A young actor's impression of her performance in the latest community play may determine whether she decides to hop on the next bus to Hollywood rather than enroll in the local college. Thus, much like Ed Koch did during his 12 years as mayor of New York City, people frequently ask and reflect on how they are doing at a particular task.

However, gaining accurate insight into one's performance can be an elusive goal. Abstract beliefs about one's ability, as well as beliefs about how well one is performing in particular instances, often correlate rather imperfectly with the reality of performance—and at times fail to correlate at all. For example, the accuracy with which people convey their feelings to others frequently does not correlate with their perception of how well they have conveyed them (Riggio, Widaman, & Friedman, 1985). Doctors' estimates of their knowledge about a variety of disorders fail to correlate with demonstrated knowledge (Tracey, Arroll, Richmond, & Barham, 1997); nurses' estimates of their basic life support skills are not related to their actual level of knowledge (Marteau, Johnston, Wynne, & Evans, 1989). For adolescent boys taking a quiz of condom use, their assessment of their knowledge correlates only slightly with their actual knowledge (Crosby & Yarber, 2001). Gun owners have little insight into how they have performed on a test of gun use, safety, and knowledge (Ehrlinger, Johnson, & Dunning, 2002). When attempting to discover who might be lying, the confidence people imbue in their judgments correlates only .04 with their accuracy (DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997). The confidence of eyewitnesses when making identifications correlates very modestly and unreliably with whether they have accurately identified the culprit out of the lineup (Sporer, Penrod, Read, & Cutler, 1995).

Although extreme, these examples are not necessarily isolated cases. In general, the perceptions people hold, of either their overall ability or specific performance, tend to be correlated only modestly with their actual performance. In 1982, Mabe and West surveyed 55 studies that had examined the relationship between perception and reality of performance across a wide variety of domains, ranging from managerial duties, clerical duties, scholastic performance to athletic skill, and found the average correlation to hover around .29. To be sure, such correlations did tend to be positive and statistically significant, but they also tended to be far from perfect. Whether people's perceptions were expressed as specific task predictions (e.g., what grade will you get in this

Editor's Note. James Olson served as the guest editor for this article.—PD

The research described herein was supported financially by National Institute of Mental Health Grant RO1-56072 awarded to David Dunning.

We thank Leah Doane, Kevin Van Aelst, and Nathalie Vizueta for assisting in the collection of data and Dick Darlington, Thomas Gilovich, and Dennis Regan for their comments on previous versions of this article. Elaine Wethington's help with AMOS is appreciated.

Correspondence concerning this article should be addressed to either Joyce Ehrlinger or David Dunning, Department of Psychology, Uris Hall, Cornell University, Ithaca, New York 14853-7601. E-mail: jme15@cornell.edu or dad6@cornell.edu

Journal of Personality and Social Psychology, 2003, Vol. 84, No. 1, 5–17
 0022-3514/03/$12.00 DOI: 10.1037/0022-3514.84.1.5

particular class?) or as more abstract assessments (e.g., their overall assessment of their organizational skill) mattered only slightly in influencing the correlation.[1]

Studies in the workplace also tend to find that perceptions of performance are only modestly related to more objective measures of performance. Across a number of studies, employees' perceptions of their job performance correlate, on average, only .35 with their supervisors' and .36 with their peers' opinion of them, even after correcting for unreliable measurement of ability. By contrast, the perceptions of peers and supervisors correlate roughly .62 (Harris & Schaubroeck, 1988). Peer perceptions of leadership ability have been shown to better predict such outcomes as recommendations for promotion among naval officers than do self-perceptions (Bass & Yammarino, 1991). Peer and supervisor perceptions also predict how well surgical residents will do on an objective test of surgical skill, whereas self-perceptions do not (Risucci, Tortolani, & Ward, 1989).

In short, people often possess an imperfect degree of what educational, cognitive, and social psychologists refer to as *metacognitive* insight, which, among its many meanings, refers to the skill of anticipating the likely accuracy and error of one's responses (Metcalfe & Shimamura, 1994; Yzerbyt, Lories, & Dardenne, 1998)—and this lack of insight often extends to situations in which people attempt to estimate their performance on a particular task or test. In recent years, research has increasingly focused on why metacognitive knowledge can be so imperfect. For example, Kruger and Dunning (1999) suggested that some individuals, namely, the incompetent, are just not in a position to judge their performances accurately. Not only does their lack of skill prevent the incompetent from forming current responses to situational demands, but it also prevents them from recognizing when judgments will be accurate and when they will be erroneous. In a series of studies, Kruger and Dunning demonstrated that incompetent individuals (i.e., those performing poorly relative to their peers) were the least able to assess the quality of their performance as well as the performances of others (see also Bem & Lord, 1979; Fagot & O'Brien, 1994; Kunkel, 1971; Maki, Jonas, & Kallod, 1994; Moreland, Miller, & Laucka, 1981; Shaughnessy, 1979; Sinkavich, 1995, for similar findings).

Sources of Performance Estimates: The Importance of Chronic Self-Views

Although demonstrating one reason why people fail to evaluate their performance accurately, Kruger and Dunning (1999) did not explore an equally important question: Where do people's evaluations of their performances come from, particularly instances in which they assess their performance on a specific test? If not strongly based on their actual achievement, what are people's perceptions of their performance based on?

In this article, we identify a potentially important source of people's impressions of their performance on a specific task, one that can be independent of their actual performance and that can potentially produce errors in estimation. In four studies, we suggest that people's evaluation of how well they have done at a particular task is not exclusively determined by their experience with the current task or the cues they encounter as they complete it. Rather, these evaluations stem from general views individuals chronically hold about their abilities, ones they hold before they even begin a task. People who think they have terrific skills at logical reasoning will think they did better on a test of logic than will those who think they do not have the skill, independent of the actual performance they post. In a phrase, performance evaluations on specific tasks may not be so much bottom-up as they are top-down, formed by referring to a person's chronic view about his or her abilities in the specific domain in question.

One could argue that it is proper to base performance estimates on chronic self-views because they are likely to be based on a lifetime of performances in the domain in question. Thus, differing self-views would largely reflect differences in objective competence, and relying on self-views would be an important strategy toward accurately assessing one's current performance at a task. In part, this is true. As Mabe and West (1982) discovered, chronic self-views did tend to predict performances, albeit modestly.

That said, there are many reasons to believe that self-views would not necessarily objectively reflect one's past performances. Self-views are likely to be influenced by a host of features that have little to do with true ability level. For example, assessments of ability tend to be self-serving (for reviews, see Dunning, 1993; Kunda, 1990). People do not dispassionately count up their success and failures to form a self-impression as much as they actively interpret them to fit chronic views, usually positive ones, of the self (Dunning, 2001; Kunda, 1990; Miller & Ross, 1985). Positive feedback is more likely to be accepted unquestioningly; negative feedback is placed under close scrutiny with an eye toward discounting it (Ditto & Lopez, 1992). Finally, people's memories of past events become distorted according to their self-views, and so the act of accurately recording some feedback provides little insurance that it will be accurately recalled in the future (Story, 1998). Given that self-views are unlikely to be a completely objective accounting of how one has performed in the past, they may not necessarily lead to accuracy over error in performance estimates.

Some fragments exist in the psychological literature indirectly suggesting that performance evaluations are top-down in nature. People rely heavily on self-views for specific domains when estimating how well they pass other "tests." Kenny and DePaulo (1993) discovered that people's beliefs about how they were viewed by others were more related to their own beliefs about

[1] Of course, one well-developed literature systematically contradicts this conclusion, finding that self-perceptions are tightly correlated with objective performance—namely, work on self-efficacy (Bandura, 1977, 1997). This contradiction may be instructive. In much of the work reviewed in the text, participants are faced with tasks in which good performance depends on knowledge, intelligence, wisdom, or common sense. In self-efficacy work, performance depends on something else—participants' ability to control or regulate their actions, such as controlling a phobia (Bandura, 1977, 1997). Self-assessment questions in self-efficacy work are also very tightly crafted to specifically match the tasks confronting participants. It might be the case that people can provide much more accurate self-assessments when the issue is what they can control versus what they know, and, given the lessons learned from self-efficacy research, when self-assessments of control are quite specific to the task at hand (Bandura, 1977, 1997). Given that Mabe and West (1982) found a wide range of self-view/performance correlations across the studies they reviewed, this speculation might be one well worth pursuing in determining when self-perceptions are accurate and when they are not.

themselves than they were to the actual impressions that others held. McFarland, Ross, and DeCourville (1989) discovered that women's memories of a menstrual cycle were biased toward their general theories about how they thought they were affected by menstrual cycles. With respect to performance situations, both self-esteem and depression have been shown to predict people's predictions about how well they will perform in the future (Dunning & Story, 1991; McFarlin & Blascovich, 1981). Self-esteem also predicts what personality traits were reflected in performances (Lewinsohn, Mischel, Chaplin, & Barton, 1980), as well as whether people will evaluate their performances as "good" or "bad" (Jussim, Coleman, & Nassau, 1987; Lindeman, Sundvik, & Rouhiainen, 1995).

However, to date, virtually no work has examined whether self-views influence what people think their performance objectively was rather than their satisfaction or dissatisfaction with a performance. That is, in a situation in which the quality of a person's performance can be objectively determined, do self-views influence what level of performance people think they have attained? And do self-views produce estimates of performance that are independent of the reality of the performance? To our knowledge, there is only one extant study showing that self-views influence estimates of performance independently of an actual performance that can be verified. In 1976, Shrauger and Terbovic asked participants who possessed high or low self-esteem to complete a conceptual learning task. Although high- and low-esteem participants did not differ in the number of items they answered correctly, high-esteem participants estimated that that they had gotten a higher number of items right than did their low-esteem peers. High-esteem participants also believed they outperformed a higher percentage of their peers than did low-esteem individuals.

In this article, we extend this initial finding by Shrauger and Terbovic (1976) to provide a more comprehensive examination of the role played by chronic self-views in estimates of performance. Like, Shrauger and Terbovic, we examine whether the relationship between self-views and performance estimates exists independent of any relationship to actual performance. Unlike Shrauger and Terbovic, we examine more specific self-views (e.g., perceptions of scientific talent), ones that might have a better chance of being related to a person's actual performance than more general impressions of self-esteem. But more than that, we examine whether self-views play a causal role in producing performance estimates that are independent of actual performance, focusing on self-views that are specific to a domain rather than a general sense of self-esteem. We do so by manipulating which specific self-view is seen as relevant to a task, or by raising or lowering a person's self-view, to see if we can alter performance estimates even when no change in actual performance is observed. We place the issue of accuracy and error under closer scrutiny, to see if self-views are related to whether people under- or overestimate their performances. Finally, we examine whether self-views, independent of actual performance, produce performance estimates that have an impact on behavior.

Overview of Present Research

In four studies, we examined the extent to which individuals rely on self-views of ability in a top-down fashion in order to estimate performance levels. In Study 1, we explored whether perceptions of performance were predicted by chronic self-views independent of how individuals actually performed. In Study 2, we manipulated which self-view was supposedly relevant to a task to see if we could alter participants' performance estimates independently of any change in performance. In Study 3, we manipulated the favorability of self-views directly to see if that influenced performance estimates independently of any difference in actual performance. Finally, in Study 4, we extended our analysis to behavioral choices, specifically, the opportunity to enter into a science competition. We predicted that perceptions of performance would be driven significantly by self-views. We expected that perceptions of performance would influence choices to participate in the contest, independent of actual performance. Thus, we compared the perceptions, performances, and choices of men and women in this experiment in an effort to contribute to an understanding of why women leave scientific careers more frequently than do men.

Study 1: Do Self-Views Matter?

Study 1 was designed to fulfill a two-fold task. The first was to examine the exact role played by chronic self-views in estimates of performance. We expected self-views to be related to performance estimates. However, that relationship could come in two forms. First, self-views might be related to performance estimates simply because self-views track differences in actual performance. Once that actual performance was controlled for, the relationship between self-views and estimates would disappear. Second, and actually what we predicted, self-views could be significantly related to performance estimates, independent of any relation to actual performance. Even after accounting for actual performance, more positive self-views could be associated with more optimistic performance estimates, more negative self-views with more pessimistic estimates.

The second task was to gauge how important self-views were in predicting performance estimates. We did this by comparing how well self-views predicted performance estimates with what should be predicting performance estimates, the actual performance participants had attained.

In Study 1, we first measured self-views of abstract reasoning ability. We then gave students a test of that ability and asked them to estimate how well they had done on the test, in terms of their raw score as well as their percentile ranking relative to their peers.

Method

Participants. Participants were 59 (44 women, 15 men) Cornell University undergraduate students who were recruited from large lecture psychology courses to take part in the study in exchange for extra credit toward their course grades.

Procedure. We described the first study as one about perceptions of strengths and weaknesses. Participants rated the extent to which they possessed 14 abilities on a scale ranging from 1 (*not at all*) to 9 (*to an extreme degree*). One item asked the degree to which they possessed an "ability to reason abstractly"; the others referred to irrelevant abilities.

After completing the scale, participants completed a supposed second study, described as focusing in on a single ability. Participants were asked to complete a 10-item multiple-choice test consisting of items taken from a Law School Aptitude Test (LSAT) preparation guide (Orton, 1993). We labeled the test as one measuring "logical reasoning ability" rather than "abstract reasoning" to reduce a demand for consistency by obscuring the relationship between the initial ability questionnaire and the test. Finally,

participants provided an estimate of how many of the 10 items they had answered correctly as well as a percentile estimate of their performance relative to other Cornell University students in the experiment (between 1 and 100).

Results and Discussion

Gender did not interact with any of the measures in this or the next two studies. Thus, all analyses are reported without mention of gender until Study 4.

Accuracy of performance estimates. Replicating past research, participants tended to overestimate how well they had done on the test relative to their peers. Participants on average believed their performance fell in the 61st percentile, significantly higher than the true average (50th percentile), $t(58) = 4.54$, $p < .0001$. Participants, however, did not overestimate the raw number of items they got right ($t < 1$). Estimates of performance were correlated with actual performance, but the strength of that correlation depended on the specific measure examined. Percentile estimates were only marginally correlated with the actual percentile participants attained, $r(58) = .22$, $p < .10$, but estimates of raw score were correlated with the actual raw score that participants achieved, $r(58) = .42$, $p < .001$.

The role of self-views. Our key hypothesis was that self-views would be a significant source of performance estimates, independent of any relationship with actual performance. Looking at zero-order correlations, we found, as predicted, that self-views did significantly predict performance estimates—$r(58) = .31$ and .42, ps $< .05$, for percentile and raw estimates, respectively.

To address whether self-views contribute to performance estimates when actual performance is held constant, we conducted multiple regression analyses predicting performance estimates from both self-views and actual performance. For percentile estimates of performance, the overall model was marginally significant, $F(2, 56) = 3.53$, $p < .10$. Of more importance, performance estimates turned out to be significantly related to self-views ($\beta = .27$, $p < .05$), but not to participants' actual percentile rankings ($\beta = .13$, *ns*). Estimates of raw scores showed a similar pattern. Self-views of abstract reasoning ability and actual performance (measured as raw score) predicted raw score estimates, $F(2, 56) = 9.81$, $p < .0005$. Taken separately, each also significantly predicted raw score estimates—and did so equally (both βs $= .31$, $p < .05$).

Summary. In sum, self-views of logical ability appeared to be a significant source of participants' perceptions of how well they had performed on the test. Indeed, depending on the measure, self-views were just as, and sometimes more, important a source as actual performance. Participants' estimates of their percentile ranking were significantly correlated with their self-views, but not with their actual percentile ranking. Participants' estimates of their raw score on the test were as related to their preconceived notions of self as they were with their actual raw scores.

Of key importance, self-views significantly influenced performance estimates after controlling for actual performance regardless of the measure. This suggests that self-views are partially responsible for the mistakes people make when they evaluate how well they have performed on a task. Relative to people with low opinions of their abilities, people with high opinions will estimate that they are doing much better, even when both groups are posting equal performances. Conversely, even when people with little confidence in their abilities perform just as well as their high self-view counterparts, they estimate that they are doing worse.

Study 2: Which Self-View to Apply?

Study 1 suggested that self-views are an important source of performance estimates, but did not do so conclusively. The design of Study 1 was correlational, and although it showed that self-views were related to performance estimates at times more strongly than actual performance was, it could not show that self-views are a causal influence over those estimates.

Thus, in Studies 2 and 3, we performed experiments to show more directly that chronic self-views influence performance estimates, and do so independently of actual performance. In Study 2, we did so by varying which of two possible self-views was purportedly relevant to a test that we asked participants to complete. We gave participants a test based on the Graduate Record Examination (GRE) analytical section, and capitalized on the fact that the test can be described as measuring several different abilities. We told one group of participants that the test measured abstract reasoning ability, an ability our participant population tended to think they had in abundance. We told another group that the test assessed computer programming ability, an ability our population tended to think they did not have.

We predicted, independent of actual performance, that participants in the abstract reasoning group, consulting self-views of abstract reasoning ability, would estimate they did much better than would participants in the computer programming group, consulting self-views of programming ability, even though each group confronted exactly the same test. Through mediational analyses, we also sought to show that these differences in performance estimates would arise precisely because students hold chronic notions that they possess more abstract reasoning ability than they do computer programming skills.

Method

Participants. Participants were 91 (64 women, 27 men) Cornell University students who participated in exchange for extra credit in large lecture psychology courses.

Procedure. Participants first rated the extent to which they possessed a number of abilities, including the ability to program a computer and the ability to think about abstract concepts, using the 9-point scale described for Study 1. Participants also rated the extent to which they thought each of these abilities was desirable on a scale ranging from 1 (*not at all desirable*) to 9 (*extremely desirable*). The order of these two sets of measures was counterbalanced.

The experimental task was then introduced. Participants were given a short test of 10 GRE analytical items taken from a GRE test preparation guide (Educational Testing Service, 1996). Those in the abstract reasoning group were told in written instructions that "the logic necessary to understand abstract concepts generally makes more sense to those who score highly on this test." Those in the computer programming condition were told "the logic of how to build a computer program generally makes more sense to those who score highly on the test." The test was exactly identical for both conditions except for what it was said to measure and a large label on the front page reading either "Abstract Reasoning Ability Exam" or "Computer Programming Ability Exam."

After completing the exam, participants provided percentile and raw-score estimates of performance, using measures similar to those in Study 1.

They also estimated how many questions they believed Cornell University students in the experiment correctly answered on average.

Results and Discussion

Self-views. As expected, participants overall rated their abstract reasoning ability more highly ($M = 6.6$) than their computer programming skills ($M = 2.7$), $t(90) = 16.38$, $p < .0001$. Participants also rated the ability to think abstractly as more desirable ($M = 7.2$) than computer skills ($M = 5.6$), $t(90) = 7.02$, $p < .0001$.

Performance estimates. These self-views mattered when it came to participants' performance estimates. Across participants, the self-view considered relevant to the task (ratings of abstract reasoning in the abstract reasoning condition and of computer ability in the computer programming condition) predicted both raw (partial $r = .33$, $p < .005$) and percentile (partial $r = .44$, $p < .0001$) estimates of performance even after controlling for the actual performance obtained. Thus, it was not surprising that estimates of performance differed across the abstract reasoning and computer programming conditions. When providing percentile estimates of their performance on the test, abstract reasoning participants (those who had been told that the test was a measure of abstract reasoning) rated themselves more highly ($M = 70.8$) than did their computer programming counterparts ($M = 58.4$), $t(89) = 2.80$, $p < .01$. They also estimated that they had achieved a higher raw score ($M = 7.7$) than their computer programming peers ($M = 6.7$), $t(89) = 2.34$, $p < .05$. Of importance, these different perceptions arose in the absence of a difference in actual performance on the test (Ms = 8.2 and 7.9 for the abstract reasoning and computer programming conditions, respectively), $t(89) = 0.81$, *ns.* In addition, we created a single index of performance estimates by standardizing participants' percentile and raw score estimates and then averaging them into one overall measure. A multiple regression predicting this index from condition and actual performance showed a significant effect of condition independent of actual score, $F(1, 88) = 7.42$, $p < .001$. Participants across conditions also did not differ in how many items they estimated the average Cornell University student in the experiment would get right (Ms in both conditions = 6.9).

Mediational analyses. In sum, despite equivalent performance, participants in the abstract reasoning group thought they had done much better than their counterparts in the computer programming condition. We conducted a mediational analysis, using the series of tests suggested by Kenny, Kashy, and Bolger (1998), to affirm that these different performance estimates arose because of a difference in the self-views participants had brought to bear in their performance estimates. We have already established that the condition to which participants were assigned influenced the proposed mediator, namely the favorability of the self-views they thought were relevant to the test ($\beta = .67$, $p < .0001$). We have also shown that the condition was related to the key dependent measure, our overall index of performance estimates ($\beta = .22$, $p < .05$). A final test revealed that self-views of ability continued to predict performance estimates even after controlling for the condition to which participants had been assigned ($\beta = .39$, $p < .005$). The impact of condition on performance estimates was significantly reduced by a Sobel test ($z = 3.04$, $p < .005$), and was, in fact, nonsignificant ($\beta = -.04$, *ns*).

Summary. In sum, Study 2 provided more conclusive evidence that people rely on self-views to determine how well they have performed on a task. In the study, we manipulated whether a high or low self-view was relevant on a test, and that manipulation had a significant impact on how people perceived their performance without impacting their actual performance. Mediational analyses provided further support that participants differed in their performance estimates as a direct result of the manipulation changing the self-view considered relevant.

Study 3: Does Altering the Self-View Alter Performance Estimates?

Study 3 was designed to further demonstrate that people rely on self-views when estimating their performance, independent of actual performance. We manipulated whether participants had a high or low opinion of their proficiency in a single domain, knowledge of U.S. geography. We then gave them a test of geographical knowledge and investigated whether our manipulation of self-views altered participants' estimates of performance.

Our double-barreled manipulation of self-views for Study 3 was inspired by two manipulations found in previous research. First, Salancik and Conway (1975; for a similar design, see Chaiken & Baldwin, 1981) found that participants could be led to believe they were more or less religious by the way they were asked about their previous religious behaviors. Participants given items for which it is easy to give the "religious" response (e.g., "Do you occasionally go to church?") afterwards described themselves as more religious than did participants given items for which it is hard to provide the religious response (e.g., "Do you go to church every week?"). Second, Schwarz et al. (1991) found that they could alter participants' self-views by varying the amount of behavioral evidence they asked participants to provide about those views. For example, after asking participants to provide 12 examples of their assertive behaviors, participants rated themselves as less assertive than did those who were asked to list only 4 examples.

In our study, we used both techniques to influence participants' views of their knowledge of U.S. geography. In one part of the manipulation, we asked participants whether they had visited a series of locations that we had determined, through pretesting, were likely or unlikely locations for our participants to have visited. We also asked them a series of questions that placed their knowledge of geography in a favorable light (e.g., could they name 1, 2–3, or more than 3 Canadian provinces) or an unfavorable one (e.g., could they name 1–5, 6–10, or more than 10 Canadian provinces). One group of participants received a questionnaire that was designed to give them a favorable impression of their geographical knowledge, in that they were asked whether they visited locations that in all likelihood they had, and also asked questions that placed their knowledge in a favorable light. The other group received a negative questionnaire that asked about locations they had likely not visited and presented them with questions that placed their geographical knowledge in some disrepute.

After this self-view manipulation, participants in both conditions were given the exact same geography quiz. They were given a blank map of North America and asked to place 15 U.S. cities. Independent of actual performance on this test, we expected participants who had received a positive self-view manipulation

would provide more favorable performance estimates than those who received a negative manipulation. In addition, we predicted that these differences in performance estimates would be mediated by the manipulated favorability of self-views.

Method

Participants. Participants were 55 (40 women, 15 men) Cornell University undergraduate students who took part in exchange for extra credit toward their grades in large lecture psychology courses.

Self-view manipulation. We sought to manipulate self-view of U.S. geographical knowledge by asking participants a series of questions purportedly designed to see how much they had traveled and how much geography they knew. The questions were designed to lead participants to provide favorable responses about themselves in one condition and unfavorable responses in the other. In the first part of the questionnaire, participants were asked whether they had ever visited six given locations. In the positive condition, the places were ones that, according to pretesting, Cornell undergraduates were likely to have visited (i.e., New York City, Connecticut, California, Pennsylvania, Florida, and Massachusetts). In the negative condition, the locations were ones likely not to have been visited (i.e., Mississippi, Missouri, Wyoming, North Dakota, Nebraska, and Oregon).

Participants were next asked six questions about their knowledge, for example, of state capitals and Canadian provinces. On each question, they were asked to rate their knowledge on a 3-point scale—but the responses indicated by each point in the scale differed across conditions. For example, when asked how many Canadian provinces they could name, participants in the positive condition confronted a 3-point scale that read *1, 2–3,* and *4 or more.* In the negative condition, the scale read *1–5, 6–10,* and *11 or more.*

Geography test. The geography test consisted of a blank map of the United States that was drawn to scale. The map contained a small amount of topological information (e.g., rivers, lakes) but contained no information about state or national borders. Participants were given a list of 15 cities and asked to indicate where each city was located on the map. The location of Boston, Massachusetts was marked on the map to serve as an example, and was surrounded by a circle of 2.1 cm in diameter indicating to participants the range that would be considered correct for each city. Participants were told that the circle corresponded to a 150-mile radius around the true location of each city.

Procedure. Upon arriving at the laboratory, participants first completed the self-view manipulation questionnaire and then completed a manipulation check, for which they rated their knowledge of U.S. geography on a scale ranging from 1 (*very weak*) to 10 (*very strong*). Participants then completed the geography test and, finally, estimated how well they had performed using the percentile and raw-score scales described in previous studies.

Results and Discussion

Self-views. Manipulation checks showed that the positive and negative surveys did influence responses. Participants reported having visited 83% of the locations listed in the positive manipulation survey and only 17% of those in the negative one, $t(52) = 6.83$, $p < .0001$. On the six questions testing geographical knowledge with use of a 3-point scale, the mean response of participants in the positive condition was at 2.51 scale points, whereas in the negative condition it was 1.32, $t(53) = 9.68$, $p < .0001$. These different responses then prompted participants to report different self-impressions when it came to their knowledge of U.S. geography (Ms = 6.0 and 4.4 for the positive and negative groups, respectively), $t(53) = 2.85$, $p < .01$.

Performance estimates. Across the sample, self-views predicted both raw (partial $r = .55$, $p < .005$) and percentile (partial $r = .68$, $p < .0001$) estimates of performance, controlling for actual performance. Further, as predicted, participants in the positive condition evaluated their performance on the geography quiz more favorably than did those in the negative condition. This was evident in their percentile estimates (Ms = 62 and 44 for positive and negative conditions, respectively), $t(52) = 2.37$, $p < .05$, as well as in their estimates of raw score (Ms = 9.4 and 6.7, for positive and negative conditions, respectively), $t(52) = 3.00$, $p < .005$.

These significant differences in performance estimates arose in the absence of a comparable difference in actual performance. Although we found that positive-manipulation participants answered a greater number of questions correctly than their negative-manipulation counterparts (Ms = 8.6 and 7.1, respectively), this difference achieved only marginal significance, $t(52) = 1.86$, $p < .10$. Further, performance estimates remained significantly different across the two conditions even after controlling for differences in actual performance. As in Study 2, we collapsed percentile and raw score estimates into a single performance estimate index. A multiple regression predicting this index from condition and actual performance showed a significant effect of condition independent of actual score, $F(1, 52) = 4.24$, $p < .05$.

Mediational analysis. We performed a mediational analysis to assess whether the self-view manipulation influenced performance estimates by virtue of its impact on self-views of geography knowledge. The combined performance estimate index, as computed above, was significantly influenced by our manipulation ($\beta = .23$, $p < .05$). In addition, the self-view manipulation had a significant impact on the self-ratings of knowledge of geography ($\beta = .24$, $p < .05$). Of key importance, self-rated knowledge of geography strongly predicted how participants would estimate their performance on the geography test ($\beta = .70$, $p < .0001$), even after controlling for the impact of the self-view manipulation. After controlling for self-views, the impact of the manipulation fell significantly ($z = 2.19$, $p < .05$) and, indeed, was nonsignificant ($\beta = -.06$).

Summary. The data from Study 3 showed that altering an individual's self-view had a measurable impact on his or her performance estimates that operated independently of how well that person had actually performed. Participants manipulated to hold a more positive view of their knowledge of U.S. geography, relative to those manipulated toward a more negative self-view, estimated that they had performed better on a geography test. A mediational analysis showed that the impact of the manipulation on performance estimates was a result of changed self-views.

Study 4: Do Differing Self-Views Underlie Gender Differences in the Pursuit of Science?

Study 4 was designed to fulfill a two-fold purpose. The first purpose dealt with a minor methodological point. In the first three studies, we measured or altered participants' self-views immediately before we asked them to take a test and estimate how well they had done. Perhaps self-views influenced later performance estimates only by virtue of having been brought to mind minutes before performance estimates were elicited. Thus, in Study 4, we examined whether performance estimates were related to chronic

self-views measured weeks before participants faced the relevant test.

The second purpose was more extensive in scope, to see whether performance estimates influenced by chronic self-views carry behavioral consequences. In particular, we wanted to examine the relevance of our analysis of performance estimates for a significant and enduring issue in social life—the persistent gender gap in those who pursue scientific careers. Despite recent improvements, there exists a discernible gender difference in participation in scientific and engineering careers (National Research Council, 1991). From early on, girls and women are less enthusiastic about taking science courses than are their male counterparts (DeBacker & Nelson, 2000; Farenga & Joyce, 1999). According to a recent National Science Foundation (2000) report, women consistently account for a minority of the bachelor's degrees awarded in chemistry (43%), computer science (28%), chemical engineering (32%), and earth sciences (35%)—accounting for less than 20% of those awarded in physics as well as electrical, aerospace, and mechanical engineering. This gender gap increases as people get older, with a smaller percentage of women earning postgraduate degrees than those earning their bachelor's. Women make up only 22% of the science and engineering labor force, despite making up 46% of the labor force overall. These differences arise despite no apparent difference in true ability (Seymour, 1992a, 1992b).

Researchers from a variety of perspectives have made contributions toward understanding this gender disparity, focusing, for example, on the influence of cultural norms and a lack of encouragement by teachers and family (Fox, Benbow, & Perkins, 1983), as well as the dual demands placed on women to start families as well as engage in careers (Tittle, 1986). It has been demonstrated that women are less likely to value science than men (Weinburgh, 1995), and that they think less of their science and mathematical abilities (Eccles, 1987). Eccles (1994) suggests that these factors lead women to avoid scientific and engineering careers.

We speculated that our analysis of performance estimates might, in part, explain the relative reluctance of women to pursue scientific activities. Specifically, because women tend to hold more negative self-views about their scientific ability, they may misestimate how well they perform on science-related tasks. As a consequence, they decide to avoid scientific activities even when they show just as much proficiency as their male counterparts. We decided to explore this by, first, surveying male and female college students on their self-views of scientific ability, and then inviting them to the laboratory where they take a science quiz. We predicted that women would hold more negative views of their scientific ability relative to men, and that this would lead them to estimate their performance on the quiz more negatively than would men. The relatively negative impressions of performance on the part of women would be mediated by the self-views reported on our preliminary survey.

After taking the quiz, we invited male and female participants to take part in a science competition taking place sometime in the future. We expected that women would decline this invitation more frequently than the men, and this reluctance to take part would be predicted more strongly by the perception they held of their performance on the science quiz than by the reality of their performance. In short, in linking gender to differential interest in scientific activities, we proposed a cascade. Women, holding more negative self-views of scientific ability than men, would provide more negative estimates of their performance on a test, which would then lead to more disinterest in taking part in a science competition. This cascade would take place even if male and female participants did not differ in actual performance.

Method

Participants. Participants were 119 (75 women, 44 men) Cornell University students who took part in exchange for extra credit in their large lecture psychology courses.

Materials and procedure. We asked several hundred students in large lecture psychology courses to rate their ability in a number of domains, including their ability to "reason about science," on a scale ranging from 1 (*I do not possess this ability at all*) to 9 (*I possess this ability to an extreme degree*). Several weeks after answering this question, we e-mailed an invitation to participate in our study, with little description of the study's focus. A total of 116 students responded and participated in the study. Data from an additional 3 participants were excluded because of suspicion.

Upon arriving at the laboratory, participants were given a test consisting of 10 scientific reasoning items taken from an American College Test (ACT) preparation guide (American College Test, 1999). The test required participants to interpret scientific passages, as well as tables and graphs, to answer multiple-choice questions. As in past studies, after completing the test, participants estimated their performance by providing percentile ranks and estimates of their raw score. They also reported the extent to which they thought scientific reasoning ability was desirable on a scale ranging from 1 (*not at all*) to 9 (*extremely desirable*).

Participants were next given a flyer advertising a Science Jeopardy contest being organized by several science departments on campus in collaboration with the psychology department. Prizes were said to include $200 in cash and a free dinner for two at a local expensive restaurant. Participants were asked three questions with respect to this contest. They were asked whether they would like to receive an e-mail with more information, like to sign up for the contest, or be added to an e-mail list that would ask weekly science jeopardy questions for smaller prizes. Participants' responses were coded as to whether they expressed interest in the context on any of these three queries. If a participant responded "yes" to any of the questions, he or she was categorized as having some interest in the contest. If a participant answered "no" to all questions, he or she was coded as uninterested.

Results and Discussion

Gender and self-views. As predicted, women harbored more negative views of their scientific ability than did men (Ms = 6.5 and 7.6, respectively), $t(114) = -3.95$, $p < .0001$. Women, to a marginal degree, also considered scientific skills to be less desirable than did men (Ms = 6.5 and 7.2, respectively), $t(114) = 1.69$, $p < .10$.

Performance estimates. Also as expected, women evaluated their performance on the science quiz more negatively than men did, both on the percentile measure (Ms = 56.0 and 73.4, respectively), $t(114) = -3.82$, $p < .0005$, and on the raw-score estimate (Ms = 5.8 and 7.1), $t(114) = -2.83$, $p < .01$. These differential estimates arose even though women scored just as well on the test as men (Ms = 7.5 and 7.9, respectively), $t(114) = -1.03$, *ns.* We should note that both men and women underestimated their raw score on the test—but women did so more to a significant degree, $t(114) = 2.83$, $p < .01$.

Multiple regressions showed that, for men and women together, self-views predicted estimates of both raw (partial $r = .51$, $p < .0001$) and percentile (partial $r = .49$, $p < .0001$) score, indepen-

dently of actual performance. A mediational analysis showed that male and female participants provided different performance estimates because they walked into the laboratory with different views of their scientific ability. To conduct the mediational analysis, we first dummy-coded gender (women = 1, men = 0) and constructed the combined performance estimate index described in previous studies. We then conducted regression analyses that showed that (a) gender predicted performance estimates ($\beta = -.15$, $p < .05$) and (b) gender predicted the self-views of scientific ability measured weeks before participants arrived in the laboratory ($\beta = -.27$, $p < .001$). Finally, even after controlling for the influence of gender, self-views still predicted performance estimates ($\beta = .22$, $p < .005$), suggesting that these views mediated the link between gender and estimates of performance on the science quiz. Further affirming that conclusion, after controlling for self-views, the impact of gender on performance estimates was significantly reduced ($z = -3.34$, $p < .001$) and, indeed, was no longer significant ($\beta = -.09$, *ns*).

Interest in scientific activities. Women also were less enthusiastic about signing up for the science jeopardy competitions. Although 71% of the men showed some interest in signing up, only 49% of the women did ($z = 2.26$, $p < .025$). This lack of interest among women was linked more to their perceptions of how well they had done on the quiz than it was to reality. In a series of simple logistic regression analyses, we found that signing up for a science competition was related to perceptions of performance, as measured by the combined performance estimate index ($b = .67$, $p < .005$), but not to actual performance ($b = .07$, *ns*).

Further analyses revealed that perceptions of performance on the science quiz were the crucial link connecting gender to lack of interest. In a multiple logistic regression analysis, we examined how well interest in the science competitions was predicted from the combined performance estimate index, actual score on the test, ratings of the desirability of science ability, participant gender, and prerating of scientific ability. Only perceptions of performance predicted interest ($b = .99$, $p < .05$), and controlling for perceptions of performance caused the relationship between gender and interest to become nonsignificant ($b = -.38$, *ns*). Actual performance and perceptions of desirability did not significantly predict interest in further scientific activities.

Linking gender to interest in scientific activities: A path model. Our theoretical framework suggests a specific path model linking gender to interest in the science jeopardy contest. Gender is linked to chronic self-views of scientific ability, which affects perceptions of performance on the science quiz, which in turn influences interest in the science competition. We assessed this account with the structural equation modeling (SEM) program within the AMOS procedure (Arbuckle & Wothke, 1999). Five measures were included. Four were suggested by our theoretical framework: gender, chronic self-views of scientific ability, perceptions of performance (indexed by the combined raw score and percentile measure described above), and interest in the science jeopardy contest (a binary variable indexed, as above, by whether participants indicated any interest in participating on any of the three chances we gave them). We also included actual performance on the science quiz.

The model is depicted in the left half of Figure 1. In the model, the cascade from gender to self-views to performance estimates to interest in the scientific contest is depicted. Actual performance is also allowed to predict performance estimates. As can be seen in the figure, all paths were significant (Zs > 2.4, all ps $< .01$). In fact, the model fit the data rather well (comparative fit index = .99). Adding a direct link from gender to interest in the science competition did not significantly increase the fit to the data, $\chi^2(1,$

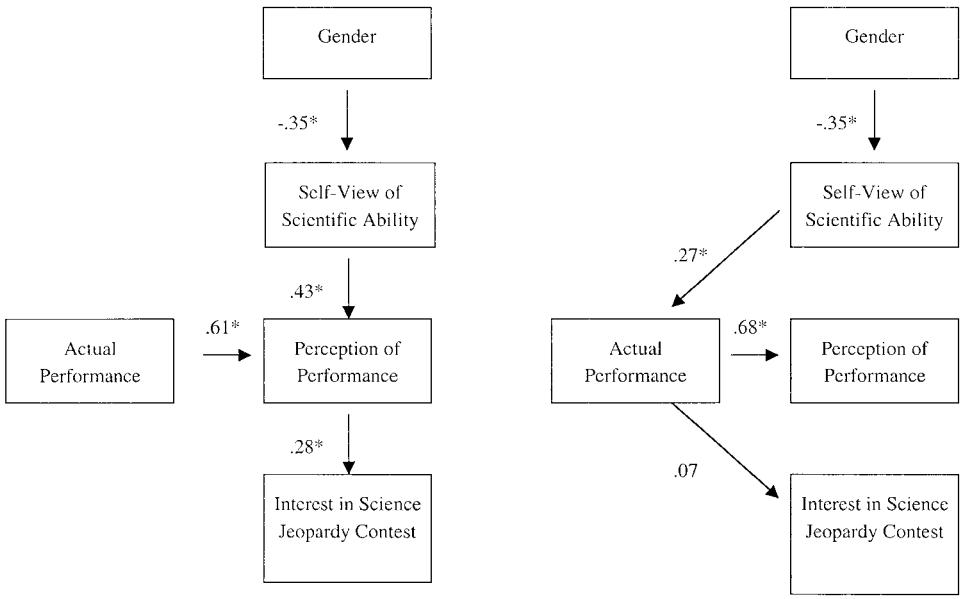

Figure 1. Two path models linking gender to interest in the science jeopardy contest. The model on the left-hand side is the predicted model in which gender is linked to interest through perceived performance on the science quiz. The alternative model of the right-hand side connects gender to interest through actual performance. All correlations marked with an asterisk are significant at $p < .005$.

$N = 119) = 2.09$, *ns*, suggesting that the cascade accounted for the link between gender and interest.

This model also fit the data better than an alternative model, depicted in the right half of Figure 1. In this model, the cascade from gender to interest in the science jeopardy competition runs through actual performance, not perceived performance. This model, $\chi^2(6, N = 119) = 59.07$, did not fit the data as well as the one above, $\chi^2(6, N = 119) = 20.14$, which is not surprising given that actual performance was not significantly related to interest in the science competition.[2,3]

Summary. In sum, for women, lack of interest in pursuing science-related activities could be traced to different perceptions of how well they had performed on a science test. Women provided much more pessimistic self-estimates than men, even though women performed just as well as men. These less favorable estimates could, in turn, be traced back to the lowered self-views of scientific ability that women brought to the laboratory. This pattern of data suggests, by extension, that women might disproportionately avoid scientific pursuits because their self-views lead them to mischaracterize how well they are objectively doing on any given scientific task. Because they think they are doing more poorly than do men, they are more likely than men to avoid science when given an option.

General Discussion

We began this article by noting how perceptions of performance are often loosely tethered to actual performance, and sought an explanation of what leads these perceptions astray from actual performance. Four studies suggested that a significant source of performance estimates were participants' chronic views of their ability within the relevant domain. Independent of actual performance, participants with high opinions of their ability thought they had done well on any given test; those with unfavorable self-views thought they had done poorly. Study 1 demonstrated this pattern by showing that previously held self-views of logical reasoning ability predicted performance estimates as strongly, and sometimes more strongly, than did actual performance on a logic test.

Study 2 went further to reveal the causal role played by these self-views. When participants were told that a test measured abstract reasoning skills—an ability they believed they had in abundance—they thought they had done much better than when told that the test assessed computer programming skills—an ability they did not believe they had. These changes in performance estimates arose even though participants confronted the exact same test and achieved the same performance regardless of what the test was said to measure. Study 3 supplied convergent evidence. Inducing participants to hold more or less favorable views of their knowledge about U.S. geography caused them to raise or lower, respectively, their estimate of how well they had done on a test of geographic knowledge, independent of any actual change in performance. Finally, Study 4 demonstrated how overreliance on inaccurate self-views can have behavioral consequences. When deciding whether to participate in a science competition, participants' decisions were based on perceptions of their performance, which were significantly linked to their chronic self-views, but not to their actual performance.

Participants across our four experiments displayed a fair degree of accuracy in predicting how well they have performed, yet the influence that self-views imposed on performance estimates existed independently of any relationship to actual performance. In fact, the relation of self-views to performance estimates, independent of actual performance, was quite strong. Across the four studies, self-views predicted estimates of one's raw score on a test as strongly as did actual raw score (mean βs = .40 and .43 for self-views and actual raw scores, respectively). Self-views predicted perceptions of one's percentile standing more strongly than actual percentile scores did (mean βs = .55 and .17 for self-views and actual percentile scores, respectively).

How Do Self-Views Influence Estimates?

A potentially valuable area for future research would be the specific pathways by which chronic self-views influence performance estimates. There seem to be several possible mechanisms linking self-views to performance estimates, independent of actual performance. For example, self-views may potentially serve as an anchor on which people base their performance estimates, away from which they adjust (insufficiently) to take into account any specific experiences they have had with the task.

Or, as people review their experience with the task, their memories may be biased by their self-views. As research on memory

---

[2] We also ran models comparing how well the cascade from gender to interest flowed through the participants' beliefs about the desirability of having scientific talent rather than through perceptions of performance on the test. These comparisons revealed that models involving desirability as part of the cascade were not as good fits to the data compared with those involving perceptions of performance. Desirability did marginally predict interest in the science jeopardy contest ($\beta = .19$, $p < .06$), but did not do so after performance perceptions were controlled for ($\beta = -.03$, *ns*).

[3] We also examined the potential role played by mood in the link between gender and interest in scientific activities. Past research has shown mood congruency in a wide variety of judgments (for reviews, see Clore, Schwarz, & Conway, 1994; Forgas, 1995). Thus, it might be possible that the mood of female participants was more negative than that of male participants, and this difference carried over to performance estimates. We collected the Positive and Negative Affect Schedule measure of positive and negative affect (Watson, Clark, & Tellegen, 1988) after completion of the exam to determine if facing a test in a domain for which one holds, for example, a negative self-view creates a negative mood, which then leads one to make negative performance estimates. In fact, there was no gender difference in either positive or negative mood (ts < 1). Further, neither positive nor negative mood correlated with estimates of performance (rs $<$ .13). Thus, it seems transitory mood did not moderate the proposed path between gender and performance evaluations in our data. On further reflection, the lack of a mood effect in this domain is not surprising. Mood tends to influence judgments of global life satisfaction, but is much less common for judgments that are familiar and regarding specific domains like the one we presented to participants (e.g., Levine, Wyer, & Schwarz, 1994; Schwarz et al., 1991, reviewed in Forgas, 1995).

A second alternative explanation worthy of attention concerns the role of self-esteem, which has been shown to predict self-views (e.g., Dutton & Brown, 1997), as well as performance evaluations (Shrauger & Terbovic, 1976). As such, it might be the case that a gender difference in self-esteem led to the reported effects. The Rosenberg Self-Esteem Scale was included within the larger pretest questionnaire from which we recruited participants (Rosenberg, 1965). We matched self-esteem scores to 69 participants (59%) in our sample. The present results occur in the absence of any gender difference in self-esteem ($t < 1$) or correlation between level of self-esteem and performance evaluations ($r = .01$, *ns*).

has shown from the classic work of Carmichael, Hogan, and Walter (1932) and Bartlett (1932), all the way to more recent work (Halberstadt & Niedenthal, 2001), memories can be significantly altered by the conceptual labels people apply to them or the abstract expectations they have. A third possible mechanism might be that the self-views provide an initial hypothesis of how one might perform. Individuals commonly seek out information that will confirm a favored hypothesis (e.g., Mynatt, Doherty, & Tweeney, 1977; Wason, 1966, 1968) and, as such, interpret new information as more supportive of a favored hypothesis than might objectively be the case (Russo, Medvec, & Meloy, 1996; Russo, Meloy, & Medvec, 1998). As a result, individuals will seem to have considerable evidence supporting an estimate of performance that is consistent with the preexisting self-view, even if one consulting the same evidence in the absence of an initial hypothesis would estimate performance quite differently. Future research would look closer at the process through which self-views influence estimates of performance.

Appropriateness of Using Self-Views in Performance Estimates

A careful reader at this point may wish to return to the issue of whether it is normatively correct to rely on self-views when forming performance estimates. It would seem plausible that using self-views to estimate performance would be a quite useful strategy, simply because such self-views would have been formed by noting feedback one has received from past performances in a domain. If past performances influence self-views, and also predict current performance, then self-views would be a valuable predictor of performance.

This analysis is correct, but it rests on two assumptions. The first is that people receive unambiguous and unbiased feedback about their performances. This assumption is not necessarily true in every case. People might not always be in a position to receive complete and unambiguous feedback about their failures and successes (Einhorn, 1982). In addition, the feedback that other people provide might be distorted by many different influences, such as the reluctance to transmit bad news (Blumberg, 1972; Tesser & Rosen, 1975).

Even when objective complete feedback is readily available, a second assumption must be fulfilled—that people dispassionately note and incorporate the feedback into their self-views. On this, the literature contains a few difficulties. Feedback often has no measurable effect on future predictions and later estimates of performance (Fischer, 1982; Keren, 1988). Further, those cases in which feedback has been shown to influence performance estimates are quite circumscribed, such as for underconfident but not overconfident judgments (Subbotin, 1996), when feedback is delivered by computer rather than by a person (Zakay, 1992), or in the presence of an explicit directive to consult feedback when estimating confidence (Winman & Juslin, 1993).

Even with these difficulties, it is still possible that self-views contain enough valid information about one's abilities and capacities to aid in providing accurate assessments. Thus, what is there to say about the value of using self-views to inform performance estimates? Answering this question is not a simple matter. Self-views may have some value in informing performance estimates, but they can also lead people astray.

First, the optimistic side. Consistent with past research, self-views, on average, did significantly predict actual performance on the tasks we gave to participants. As the top half of Table 1 shows, chronic self-views on average, tended to significantly predict the performances that participants attained (collapsed across condition, where relevant), measured both in raw score and percentile terms (mean $r = .27$, weighted by the number of participants in each study, for both performance estimate measures). The relationship between self-views and performance was modest in magnitude for Studies 1, 2, and 4 (hovering between .00 and .39) and quite strong for Study 3 (in excess of .60). Thus, overall, there was some usefulness in considering self-views when predicting performance (mean rs $= .29$ and .38, zs $= 4.83$ and 4.82, for raw-score and percentile estimates, respectively, ps $< .001$).

However, there also is a pessimistic side. Although chronic self-views accurately anticipated actual performance, such self-views also led to systematic errors in performance appraisals. Analyses we have described for each study already suggest this: Holding actual performance constant, self-views still predicted performance estimates, indicating that such self-views exerted "unwarranted" influences that could lead to mistaken performance estimates. A more conservative test, depicted in the bottom half of Table 1, also shows that self-views were related to errors in estimation (collapsed across conditions, where relevant). Table 1 shows how strongly errors in overestimation or underestimation (estimated performance minus actual performance) were predicted by chronic self-views. For Studies 2 and 4, self-views significantly predicted errors in performance estimates (all ps $< .05$). The more positive the self-view, the more participants overestimated their performance; the more negative their self-views, the more they underestimated them. Study 3 showed a similar, albeit nonsignificant, trend. Study 1 was mixed. Across the four studies, this pattern was significant (mean $r = .29$, $Z = 5.68$, $p < .001$ for raw-score estimates, and mean $r = .18$, $Z = 3.41$, $p < .001$ for percentile estimates, weighted by number of participants in each study).

In addition, even when self-views correlate with objective performance, they still might lead more to error than to accuracy. They will do so when the self-views people hold overall are either too high or too low relative to objective performance. If people, for example, believe their skill on average hovers around the 70th percentile, when the true average is the 50th percentile, then

Table 1
Relationship of Chronic Self-Views to Actual Performance and Over- or Underestimation of Performance

	Study			
Correlation between self-views and	1	2	3	4
Actual performance				
Raw score	.34**	.00	.66**	.27**
Percentile score	.31*	.39**	.64**	.29**
Over-/underestimate of performance				
Raw score	.08	.30*	.19	.44**
Percentile score	−.10	.31*	.13	.23*

Note. For each study, the correlations represent the overall data, collapsing across conditions.
* $p < .05$. ** $p < .01$.

relying on those positive self-views will prompt people in general to overestimate their performances, even if there is some underlying correlation between self-view and objective performance. Indeed, a wealth of research has shown that people tend to hold self-views that are too favorable given objective criteria, lending credence to this concern. More people, for example, claim to have above-average talents and capacities than is possible given the logic of descriptive statistics (Alicke, 1985; Dunning, Meyerowitz, & Holzberg, 1989; Kruger & Dunning, 1999; Weinstein, 1980).

An examination of these issues, as well as the psychological literature in general, suggests that the value of self-views for assessing performance will most likely depend on the domain under consideration. In some domains, the correlation between self-views and objective performance is more substantial than it is in others. For athletics, the correlation between self-view and performance tends to be higher (roughly .48) than it is for domains involving interpersonal (roughly .17) or managerial skills (roughly .04) (Mabe & West, 1982). Indeed, in this article, self-views correlated with both actual performance and error in estimates, but the strength of those correlations varied across the four domains represented. In addition, the extent to which self-views overestimate objective performance overall also varies by domain (Alicke, 1985; Dunning et al., 1989; Kruger, 1999; Weinstein, 1980). Although psychological research has provided some pointers about which domains generate more accurate self-assessments and which generate more erroneous ones, the extant work is hardly comprehensive. Once a greater understanding of the domains in which self-views tend to be accurate is achieved, more can be said about exactly when one is well-advised to base judgments on said self-views.

Women in Science

Perhaps the lesson of this article is most important when applied to Study 4, which demonstrated the potential behavioral consequences of the link between self-views and performance estimates. Women, relative to men, walked into our laboratory with more unfavorable beliefs about their level of talent at scientific reasoning. As a consequence, their performance estimates on a science quiz were lower then those of men, even though in actuality they did just as well. Furthermore, these lowered perceptions had consequences. When asked whether they would like to enter into a science competition, women were more likely than men to decline. Further, this higher rate of refusal among women was linked to their perceptions of their performance on the test we gave them, not to the reality of their performance.

The findings of Study 4 highlight the importance of learning how clear and unambiguous feedback might be incorporated to improve the accuracy of self-views. Left to their own devices, people can be inaccurate about how well they are doing, as were our female participants with respect to the quiz we gave them. Feedback from an outside agent might have gone a long way toward bringing their perceptions into line, which could have important behavioral consequences, but only if that feedback influences future perceptions and estimates. One wonders, for example, if female participants in Study 4 would have expressed more interest in the science competition if we had told them, before they made the choice, how well they had actually done on the quiz.

To be sure, the issue of why women disproportionately avoid activities involving science is a complex one. There are undoubtedly many forces—psychological, sociological, and structural—that cause women to avoid the small science competition we suggested to our participants and to withdraw from careers in science and engineering. Past research, for example, has highlighted the specific hurdles that women face when it comes to science, such as discrimination as well as the difficulty of starting a family life at the difficult first stages of one's career. Research has also shown how the threat of societal stereotypes leave women, but not men, with special psychological hurdles that they have to deal with and that might inhibit their performance (Spencer, Steele, & Quinn, 1999; Steele, 1997).

That said, what Study 4 suggests is that even when women perform just as well as men, psychological processes may still prevent them from recognizing that fact. Perhaps as women tackle engineering and scientific tasks, they come to evaluate their performances more negatively than men do, even if their performances are every bit as good as the men's. Reviewing what appears to be a more lackluster series of achievements, women decide that it is best to pursue some other livelihood. Conversely, perhaps men, convinced of the sterling quality of their achievements, decide to pursue a career at which they are not as uniquely talented as they believe themselves to be.

Concluding Remarks

We are fairly certain that, with a few minutes of reflection, any reader can generate examples of people whose perceptions of their performances fails to match a plausible form of reality. Perhaps it is that stiff and gangly dancer who prides himself on his grace, or the brilliant student who is convinced that every single comment she makes is banal. In this research, we sought to provide insight into how such disparities between perception and reality are generated, and perhaps more importantly, why they are maintained. Even when people desire to know themselves accurately, their performance estimates may vary from the truth because those estimates are importantly influenced by inaccurate self-views of ability.

Thus, when gaining an accurate impression of one's current performance is important, it may be wise to assume that one does not already know the answer but to seek out information from the outside world—and then to pay it heed. In sum, it may pay to be a little more like Ed Koch, making sure to ask the question "How'm I doin'?" when accuracy is important, or at least not to assume that one already knows the answer.

References

American College Test. (1997). *Getting into the ACT: Official guide to the ACT assessment.* San Diego, CA: Harcourt Brace.

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49,* 1621–1630.

Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 User's Guide.* Chicago, IL: SmallWaters Corporation.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191–215.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. New York: Macmillan.

Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology, 5,* 437–454.

Bem, D. J., & Lord, C. G. (1979). Template matching: A proposal for probing the ecological validity of experimental settings in social psychology. *Journal of Personality and Social Psychology, 37,* 833–846.

Blumberg, H. H. (1972). Communication of interpersonal evaluations. *Journal of Personality and Social Psychology, 23,* 157–162.

Carmichael, L. C., Hogan, H. P., & Walter, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology, 15,* 73–86.

Chaiken, S., & Baldwin, M. W. (1981). Affective-cognitive consistency and the effect of salient behavioral information on the self-perception of attitudes. *Journal of Personality and Social Psychology, 41,* 1–12.

Clore, G. L., Schwarz, N., Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 1, pp. 323–419). Hillsdale, NJ: Erlbaum.

Crosby, R. A., & Yarber, W. L. (2001). Perceived versus actual knowledge about correct condom use among U.S. adolescents: Results from a national study. *Journal of Adolescent Health, 28,* 415–420.

DeBacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: Differences related to gender, class type, and ability. *The Journal of Educational Research, 93,* 245–254.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1,* 346–357.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63,* 568–584.

Dunning, D. (1993). Words to live by: The self and definitions of social concepts and categories. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 4, pp. 99–126). Hillsdale, NJ: Erlbaum.

Dunning, D. (2001). On the motives underlying social cognition. In N. Schwarz & A. Tesser (Eds.), *Blackwell handbook of social psychology, Volume 1: Intraindividual processes* (pp. 348–374). New York: Blackwell Publishers.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57,* 1082–1090.

Dunning, D., & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology, 61,* 521–532.

Dutton, K. A., & Brown, J. D. (1997). Global self-esteem and specific self-views as determinants of people's reactions to success and failure. *Journal of Personality and Social Psychology, 73,* 139–148.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11,* 135–172.

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly, 18,* 585–609.

Educational Testing Service. (1996). *GRE: Practicing to take the general test: Big book.* Princeton, NJ: Educational Testing Service.

Ehrlinger, J., Johnson, K. L., & Dunning, D. (2002). *Are the unskilled really unaware? Further explorations of self-assessments among the incompetent.* Manuscript in preparation, Cornell University, Ithaca, New York.

Einhorn, H. J. (1982). Learning from experience and suboptimal rules in decision making. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 268–286). New York: Cambridge University Press.

Fagot, B. I., & O'Brien, M. (1994). Activity level in young children: Cross-age stability, situational influences, correlates with temperament, and the perception of problem behaviors. *Merrill Palmer Quarterly, 40,* 378–398.

Farenga, S. J., & Joyce, B. A. (1999). Intentions of young students to enroll in science courses in the future: An examination of gender differences. *Science Education, 83,* 55–75.

Fischer, G. W. (1982). Scoring-rule feedback and the overconfidence syndrome in subjective probability forecasting. *Organizational Behavior and Human Performance, 29,* 352–369.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39–66.

Fox, L. H., Benbow, C. P., & Perkins, S. (1983). An accelerated mathematics program for girls: A longitudinal evaluation. In C. P. Benbow & J. C. Stanley (Eds.), *Academic precocity: Aspects of its development* (pp. 133–139). Baltimore, MD: Johns Hopkins University Press.

Halberstadt, J. B., & Niedenthal, P. M. (2001). Effects of emotion concepts on perceptual memory for emotional expressions. *Journal of Personality and Social Psychology, 81,* 587–598.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology, 41,* 43–62.

Jussim, L., Coleman, L., & Nassau, S. (1987). The influence of self-esteem on perceptions of performance and feedback. *Social Psychology Quarterly, 50,* 95–99.

Kenny, D. A., & DePaulo, B. M. (1993). Do people know how others view them? An empirical and theoretical account. *Psychological Bulletin, 114,* 145–161.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). New York: McGraw-Hill.

Keren, G. (1988). On the ability of monitoring non-veridical perceptions and uncertain knowledge: Some calibration studies. *Acta Psychologica, 67,* 95–119.

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

Kruger, J. M., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121–1134.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480–498.

Kunkel, E. (1971). On the relationship between estimate of ability and driver qualification. *Psychologie und Praxis, 15,* 73–80.

Levine, S. R., Wyer, R. S., & Schwarz, N. (1994). Are you what you feel? The affective and cognitive determinants of self-judgments. *European Journal of Social Psychology, 24,* 63–78.

Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology, 89,* 203–212.

Lindeman, M., Sundvik, L., & Rouhiainen, P. (1995). Under- or overestimation of self? Person variables and self-assessment accuracy in work settings. *Journal of Social Behavior and Personality, 10,* 123–134.

Mabe, P. A., III, & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67,* 280–296.

Maki, R. H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin and Review, 1,* 126–129.

Marteau, T. M., Johnston, M., Wynne, G., & Evans, T. R. (1989). Cognitive factors in the explanation of the mismatch between confidence and

competence in performing basic life support. *Psychology and Health, 3,* 173–182.

McFarland, C., Ross, M., & DeCourville, N. (1989). Women's theories of menstruation and biases in recall of menstrual symptoms. *Journal of Personality and Social Psychology, 57,* 522–531.

McFarlin, D. B., & Blascovich, J. (1981). Effects of self-esteem and performance feedback on future affective preferences and cognitive expectations. *Journal of Personality and Social Psychology, 40,* 521–531.

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition.* Cambridge, MA: MIT Press.

Miller, D. T., & Ross, M. (1985). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82,* 213–225.

Moreland, R., Miller, J., & Laucka, F. (1981). Academic achievement and self-evaluations of academic performance. *Journal of Educational Psychology, 73,* 335–344.

Mynatt, C. R., Doherty, M. E., & Tweeney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology, 29,* 85–95.

National Research Council: Committee on Women in Science and Engineering. (1991). *Women in science and engineering: Increasing their numbers in the 1990s.* Washington, DC: National Academy Press.

National Science Foundation. (2000). *Women, minorities, and persons with disabilities in science and engineering.* Arlington, VA: National Science Foundation.

Orton, P. Z. (1993). *Cliffs law school admission test preparation guide.* Lincoln, NE: Cliffs Notes Inc.

Riggio, R. E., Widaman, K. F., & Friedman, H. S. (1985). Actual and perceived emotional sending and personality correlates. *Journal of Nonverbal Behavior, 9,* 69–83.

Risucci, D. A., Tortolani, A. J., & Ward, R. J. (1989). Ratings of surgical residents by self, supervisors, and peers. *Surgical and Gynecological Obstetrics, 169,* 519–526.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Russo, J. E., Medvec, V. H., & Meloy, M. G. (1996). The distortion of information during decisions. *Organizational Behavior and Human Decision Processes, 66,* 102–110.

Russo, J. E., Meloy, M. G., & Medvec, V. H. (1998). Predecisional distortion of product information. *Journal of Marketing Research, 35,* 438–452.

Salancik, G., & Conway, M. (1975). Attitude inference from salient and relevant cognitive content about behavior. *Journal of Personality and Social Psychology, 32,* 829–840.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195–202.

Seymour, E. (1992a). "The problem iceberg" in science, mathematics, and engineering education: Student explanations for high attrition rates. *Journal of College Science Teaching, 21,* 230–232.

Seymour, E. (1992b). Undergraduate problems with teaching and advising in SME majors: Explaining gender differences in attrition rates. *Journal of College Science Teaching, 21,* 284–292.

Shaughnessy, J. J. (1979). Confidence judgment accuracy as a predictor of test performance. *Journal of Research in Personality, 13,* 505–514.

Shrauger, J. S., & Terbovic, M. L. (1976). Self-evaluation and assessments of performance by self and others. *Journal of Consulting and Clinical Psychology, 44,* 564–572.

Sinkavich, F. J. (1995). Performance and metamemory: Do students know what they don't know? *Instructional Psychology, 22,* 77–87.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin, 118,* 315–327.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Story, A. L. (1998). Self-esteem and memory for favorable and unfavorable personality feedback. *Personality and Social Psychology Bulletin, 24,* 51–64.

Subbotin, V. (1996). Outcome feedback effects on under- and overconfident judgments (general knowledge tasks). *Organizational Behavior and Human Decision Processes, 66,* 268–276.

Tesser, A., & Rosen, S. (1975). The reluctance to transmit bad news. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 193–232). New York: Academic Press.

Tittle, C. K. (1986). Gender research and education. *American Psychologist, 41,* 1161–1168.

Tracey, J. M., Arroll, B., Richmond, D. E., & Barham, P. M. (1997). The validity of general practitioners' self assessment of knowledge. *British Medical Journal, 315,* 1426–1428.

Wason, P. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). London: Penguin.

Wason, P. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20,* 273–281.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of the brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970–1991. *Journal of Research in Science Teaching, 32,* 387–398.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 58,* 806–820.

Winman, A., & Juslin, P. (1993). Calibration of sensory and cognitive judgments: Two different accounts. *Scandinavian Journal of Psychology, 34,* 135–148.

Yzerbyt, V., Lories, G., & Dardenne, B. (Eds.). (1998). *Metacognition: Cognitive and social dimensions.* London: Sage.

Zakay, D. (1992). The influence of computerized feedback on overconfidence in knowledge. *Behaviour and Information Technology, 11,* 329–333.

Received January 17, 2002
Revision received July 22, 2002
Accepted July 24, 2002 ■