Journal of Personality and Social Psychology
1999, Vol. 77, No. 6, 1121-1134

0022-3514/99/$3.00

Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments

Justin Kruger and David Dunning
Cornell University

People tend to hold overly favorable views of their abilities in many social and intellectual domains. The authors suggest that this overestimation occurs, in part, because people who are unskilled in these domains suffer a dual burden: Not only do these people reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the metacognitive ability to realize it. Across 4 studies, the authors found that participants scoring in the bottom quartile on tests of humor, grammar, and logic grossly overestimated their test performance and ability. Although their test scores put them in the 12th percentile, they estimated themselves to be in the 62nd. Several analyses linked this miscalibration to deficits in metacognitive skill, or the capacity to distinguish accuracy from error. Paradoxically, improving the skills of participants, and thus increasing their metacognitive competence, helped them recognize the limitations of their abilities.

> It is one of the essential features of such incompetence that the person so afflicted is incapable of knowing that he is incompetent. To have such knowledge would already be to remedy a good portion of the offense. (Miller, 1993, p. 4)

In 1995, McArthur Wheeler walked into two Pittsburgh banks and robbed them in broad daylight, with no visible attempt at disguise. He was arrested later that night, less than an hour after videotapes of him taken from surveillance cameras were broadcast on the 11 o'clock news. When police later showed him the surveillance tapes, Mr. Wheeler stared in incredulity. "But I wore the juice," he mumbled. Apparently, Mr. Wheeler was under the impression that rubbing one's face with lemon juice rendered it invisible to videotape cameras (Fuocco, 1996).

We bring up the unfortunate affairs of Mr. Wheeler to make three points. The first two are noncontroversial. First, in many domains in life, success and satisfaction depend on knowledge, wisdom, or savvy in knowing which rules to follow and which strategies to pursue. This is true not only for committing crimes, but also for many tasks in the social and intellectual domains, such as promoting effective leadership, raising children, constructing a solid logical argument, or designing a rigorous psychological study. Second, people differ widely in the knowledge and strategies they apply in these domains (Dunning, Meyerowitz, & Holzberg, 1989; Dunning, Perie, & Story, 1991; Story & Dunning, 1998), with varying levels of success. Some of the knowledge and theories that people apply to their actions are sound and meet with favorable results. Others, like the lemon juice hypothesis of McArthur Wheeler, are imperfect at best and wrong-headed, incompetent, or dysfunctional at worst.

Perhaps more controversial is the third point, the one that is the focus of this article. We argue that when people are incompetent in the strategies they adopt to achieve success and satisfaction, they suffer a dual burden: Not only do they reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the ability to realize it. Instead, like Mr. Wheeler, they are left with the mistaken impression that they are doing just fine. As Miller (1993) perceptively observed in the quote that opens this article, and as Charles Darwin (1871) sagely noted over a century ago, "ignorance more frequently begets confidence than does knowledge" (p. 3).

In essence, we argue that the skills that engender competence in a particular domain are often the very same skills necessary to evaluate competence in that domain—one's own or anyone else's. Because of this, incompetent individuals lack what cognitive psychologists variously term *metacognition* (Everson & Tobias, 1998), *metamemory* (Klin, Guizman, & Levine, 1997), *metacomprehension* (Maki, Jonas, & Kallod, 1994), or *self-monitoring* skills (Chi, Glaser, & Rees, 1982). These terms refer to the ability to know how well one is performing, when one is likely to be accurate in judgment, and when one is likely to be in error. For example, consider the ability to write grammatical English. The skills that enable one to construct a grammatical sentence are the same skills necessary to recognize a grammatical sentence, and thus are the same skills necessary to determine if a grammatical mistake has been made. In short, the same knowledge that underlies the ability to produce correct judgment is also the knowledge

Justin Kruger and David Dunning, Department of Psychology, Cornell University.

We thank Betsy Ostrov, Mark Stalnaker, and Boris Veysman for their assistance in data collection. We also thank Andrew Hayes, Chip Heath, Rich Gonzalez, Ken Savitsky, and David Sherman for their valuable comments on an earlier version of this article, and Dov Cohen for alerting us to the quote we used to begin this article. Portions of this research were presented at the annual meeting of the Eastern Psychological Association, Boston, March 1998. This research was supported financially by National Institute of Mental Health Grant RO1 56072.

Correspondence concerning this article should be addressed to Justin Kruger, who is now at the Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820, or to David Dunning, Department of Psychology, Uris Hall, Cornell University, Ithaca, New York 14853-7601. Electronic mail may be sent to Justin Kruger at jkruger@s.psych.uiuc.edu or to David Dunning at dad6@cornell.edu.

that underlies the ability to recognize correct judgment. To lack the former is to be deficient in the latter.

Imperfect Self-Assessments

We focus on the metacognitive skills of the incompetent to explain, in part, the fact that people seem to be so imperfect in appraising themselves and their abilities.[1] Perhaps the best illustration of this tendency is the "above-average effect," or the tendency of the average person to believe he or she is above average, a result that defies the logic of descriptive statistics (Alicke, 1985; Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; Brown & Gallagher, 1992; Cross, 1977; Dunning et al., 1989; Klar, Medding, & Sarel, 1996; Weinstein, 1980; Weinstein & Lachendro, 1982). For example, high school students tend to see themselves as having more ability in leadership, getting along with others, and written expression than their peers (College Board, 1976–1977), business managers view themselves as more able than the typical manager (Larwood & Whittaker, 1977), and football players see themselves as more savvy in "football sense" than their teammates (Felson, 1981).

We believe focusing on the metacognitive deficits of the unskilled may help explain this overall tendency toward inflated self-appraisals. Because people usually choose what they think is the most reasonable and optimal option (Metcalfe, 1998), the failure to recognize that one has performed poorly will instead leave one to assume that one has performed well. As a result, the incompetent will tend to grossly overestimate their skills and abilities.

Competence and Metacognitive Skills

Several lines of research are consistent with the notion that incompetent individuals lack the metacognitive skills necessary for accurate self-assessment. Work on the nature of expertise, for instance, has revealed that novices possess poorer metacognitive skills than do experts. In physics, novices are less accurate than experts in judging the difficulty of physics problems (Chi et al., 1982). In chess, novices are less calibrated than experts about how many times they need to see a given chessboard position before they are able to reproduce it correctly (Chi, 1978). In tennis, novices are less likely than experts to successfully gauge whether specific play attempts were successful (McPherson & Thomas, 1989).

These findings suggest that unaccomplished individuals do not possess the degree of metacognitive skills necessary for accurate self-assessment that their more accomplished counterparts possess. However, none of this research has examined whether metacognitive deficiencies translate into inflated self-assessments or whether the relatively incompetent (novices) are systematically more miscalibrated about their ability than are the competent (experts).

If one skims through the psychological literature, one will find some evidence that the incompetent are less able than their more skilled peers to gauge their own level of competence. For example, Fagot and O'Brien (1994) found that socially incompetent boys were largely unaware of their lack of social graces (see Bem & Lord, 1979, for a similar result involving college students). Mediocre students are less accurate than other students at evaluating their course performance (Moreland, Miller, & Laucka, 1981). Unskilled readers are less able to assess their text comprehension than are more skilled readers (Maki, Jonas, & Kallod, 1994). Students doing poorly on tests less accurately predict which questions they will get right than do students doing well (Shaughnessy, 1979; Sinkavich, 1995). Drivers involved in accidents or flunking a driving exam predict their performance on a reaction test less accurately than do more accomplished and experienced drivers (Kunkel, 1971). However, none of these studies has examined whether deficient metacognitive skills underlie these miscalibrations, nor have they tied these miscalibrations to the above-average effect.

Predictions

These shards of empirical evidence suggest that incompetent individuals have more difficulty recognizing their true level of ability than do more competent individuals and that a lack of metacognitive skills may underlie this deficiency. Thus, we made four specific predictions about the links between competence, metacognitive ability, and inflated self-assessment.

Prediction 1. Incompetent individuals, compared with their more competent peers, will dramatically overestimate their ability and performance relative to objective criteria.

Prediction 2. Incompetent individuals will suffer from deficient metacognitive skills, in that they will be less able than their more competent peers to recognize competence when they see it—be it their own or anyone else's.

Prediction 3. Incompetent individuals will be less able than their more competent peers to gain insight into their true level of performance by means of social comparison information. In particular, because of their difficulty recognizing competence in others, incompetent individuals will be unable to use information about the choices and performances of others to form more accurate impressions of their own ability.

Prediction 4. The incompetent can gain insight about their shortcomings, but this comes (paradoxically) by making them more competent, thus providing them the metacognitive skills necessary to be able to realize that they have performed poorly.

The Studies

We explored these predictions in four studies. In each, we presented participants with tests that assessed their ability in a domain in which knowledge, wisdom, or savvy was crucial: humor (Study 1), logical reasoning (Studies 2 and 4), and English grammar (Study 3). We then asked participants to assess their ability

[1] A few words are in order about what we mean by *incompetent.* First, throughout this article, we think of incompetence as a matter of degree and not one of absolutes. There is no categorical bright line that separates "competent" individuals from "incompetent" ones. Thus, when we speak of "incompetent" individuals we mean people who are less competent than their peers. Second, we have focused our analysis on the incompetence individuals display in specific domains. We make no claim that they would be incompetent in any other domains, although many a colleague has pulled us aside to tell us a tale of a person they know who is "domain-general" incompetent. Those people may exist, but they are not the focus of this research.

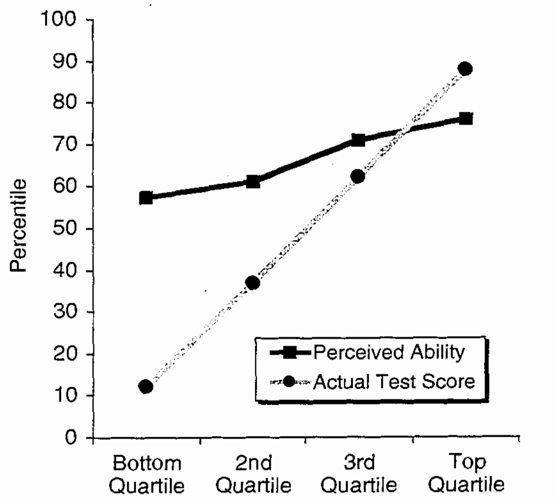

Figure 1. Perceived ability to recognize humor as a function of actual test performance (Study 1).

top quartile actually underestimated their ability relative to their peers, paired $t(15) = -2.20$, $p < .05$.

Summary

In short, Study 1 revealed two effects of interest. First, although perceptions of ability were modestly correlated with actual ability, people tended to overestimate their ability relative to their peers. Second, and most important, those who performed particularly poorly relative to their peers were utterly unaware of this fact. Participants scoring in the bottom quartile on our humor test not only overestimated their percentile ranking, but they overestimated it by 46 percentile points. To be sure, they had an inkling that they were not as talented in this domain as were participants in the top quartile, as evidenced by the significant correlation between perceived and actual ability. However, that suspicion failed to anticipate the magnitude of their shortcomings.

At first blush, the reader may point to the regression effect as an alternative interpretation of our results. After all, we examined the perceptions of people who had scored extremely poorly on the objective test we handed them, and found that their perceptions were less extreme than their reality. Because perceptions of ability are imperfectly correlated with actual ability, the regression effect virtually guarantees this result. Moreover, because incompetent participants scored close to the bottom of the distribution, it was nearly impossible for them to underestimate their performance.

Despite the inevitability of the regression effect, we believe that the overestimation we observed was more psychological than artifactual. For one, if regression alone were to blame for our results, then the magnitude of miscalibration among the bottom quartile would be comparable with that of the top quartile. A glance at Figure 1 quickly disabuses one of this notion. Still, we believe this issue warrants empirical attention, which we devote in Studies 3 and 4.

Study 2: Logical Reasoning

We conducted Study 2 with three goals in mind. First, we wanted to replicate the results of Study 1 in a different domain, one focusing on intellectual rather than social abilities. We chose logical reasoning, a skill central to the academic careers of the participants we tested and a skill that is called on frequently. We wondered if those who do poorly relative to their peers on a logical reasoning test would be unaware of their poor performance.

Examining logical reasoning also enabled us to compare perceived and actual ability in a domain less ambiguous than the one we examined in the previous study. It could reasonably be argued that humor, like beauty, is in the eye of the beholder.[2] Indeed, the imperfect interrater reliability among our group of professional comedians suggests that there is considerable variability in what is considered funny even by experts. This criterion problem, or lack of uncontroversial criteria against which self-perceptions can be compared, is particularly problematic in light of the tendency to define ambiguous traits and abilities in ways that emphasize one's own strengths (Dunning et al., 1989). Thus, it may have been the tendency to define humor idiosyncratically, and in ways favorable to one's tastes and sensibilities, that produced the miscalibration we observed—not the tendency of the incompetent to miss their own failings. By examining logical reasoning skills, we could circumvent this problem by presenting students with questions for which there is a definitive right answer.

Finally, we wanted to introduce another objective criterion with which we could compare participants' perceptions. Because percentile ranking is by definition a comparative measure, the miscalibration we saw could have come from either of two sources. In the comparison, participants may have overestimated their own ability (our contention) or may have underestimated the skills of their peers. To address this issue, in Study 2 we added a second criterion with which to compare participants' perceptions. At the end of the test, we asked participants to estimate how many of the questions they had gotten right and compared their estimates with their actual test scores. This enabled us to directly examine whether the incompetent are, indeed, miscalibrated with respect to their own ability and performance.

Method

Participants. Participants were 45 Cornell University undergraduates from a single introductory psychology course who earned extra credit for their participation. Data from one additional participant was excluded because she failed to complete the dependent measures.

Procedure. Upon arriving at the laboratory, participants were told that the study focused on logical reasoning skills. Participants then completed a 20-item logical reasoning test that we created using questions taken from a Law School Admissions Test (LSAT) test preparation guide (Orton, 1993). Afterward, participants made three estimates about their ability and test performance. First, they compared their "general logical reasoning ability" with that of other students from their psychology class by providing their percentile ranking. Second, they estimated how their score on the test would compare with that of their classmates, again on a percentile scale. Finally, they estimated how many test questions (out of 20) they thought they had answered correctly. The order in which these questions were asked was counterbalanced in this and in all subsequent studies.

[2] Actually, some theorists argue that there are universal standards of beauty (see, e.g., Thornhill & Gangestad, 1993), suggesting that this truism may not be, well, true.

mean of 50, one-sample $ts(83) = 5.90$ and 5.13, respectively, $ps < .0001$. Participants also overestimated the number of items they answered correctly, $M = 15.2$ (perceived) versus 13.3 (actual), $t(83) = 6.63$, $p < .0001$. Although participants' perceptions of their general grammar ability were uncorrelated with their actual test scores, $r(82) = .14$, *ns*, their perceptions of how their test performance would rank among their peers was correlated with their actual score, albeit to a marginal degree, $r(82) = .19$, $p < .09$, as was their direct estimate of their raw test score, $r(82) = .54$, $p < .0001$.

As Figure 3 illustrates, participants scoring in the bottom quartile grossly overestimated their ability relative to their peers. Whereas bottom-quartile participants ($n = 17$) scored in the 10th percentile on average, they estimated their grammar ability and performance on the test to be in the 67th and 61st percentiles, respectively, $ts(16) = 13.68$ and 15.75, $ps < .0001$. Bottom-quartile participants also overestimated their raw score on the test by 3.7 points, $M = 12.9$ (perceived) versus 9.2 (actual), $t(16) = 5.79$, $p < .0001$.

As in previous studies, participants falling in other quartiles overestimated their ability and performance much less than did those in the bottom quartile. However, as Figure 3 shows, those in the top quartile once again underestimated themselves. Whereas their test performance fell in the 89th percentile among their peers, they rated their ability to be in the 72nd percentile and their test performance in the 70th percentile, $ts(18) = -4.73$ and -5.08, respectively, $ps < .0001$. Top-quartile participants did not, however, underestimate their raw score on the test, $M = 16.9$ (perceived) versus 16.4 (actual), $t(18) = 1.37$, *ns*.

Study 3 (Phase 2): It Takes One to Know One

Thus far, we have shown that people who lack the knowledge or wisdom to perform well are often unaware of this fact. We attribute this lack of awareness to a deficit in metacognitive skill. That is, the same incompetence that leads them to make wrong choices also deprives them of the savvy necessary to recognize competence, be it their own or anyone else's.

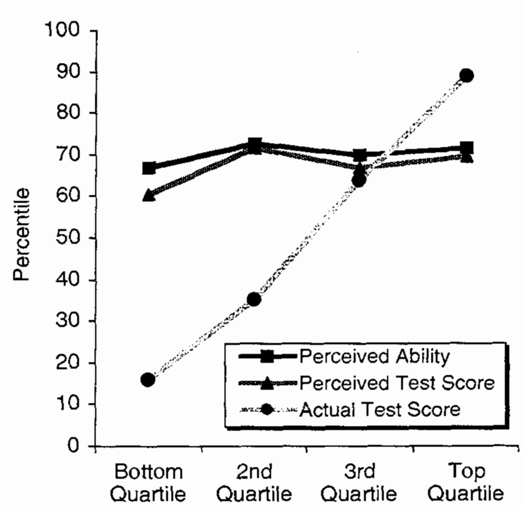

Figure 3. Perceived grammar ability and test performance as a function of actual test performance (Study 3).

We designed a second phase of Study 3 to put the latter half of this claim to a test. Several weeks after the first phase of Study 3, we invited the bottom- and top-quartile performers from this study back to the laboratory for a follow-up. There, we gave each group the tests of five of their peers to "grade" and asked them to assess how competent each target had been in completing the test. In keeping with Prediction 2, we expected that bottom-quartile participants would have more trouble with this metacognitive task than would their top-quartile counterparts.

This study also enabled us to explore Prediction 3, that incompetent individuals fail to gain insight into their own incompetence by observing the behavior of other people. One of the ways people gain insight into their own competence is by comparing themselves with others (Festinger, 1954; Gilbert, Giesler, & Morris, 1995). We reasoned that if the incompetent cannot recognize competence in others, then they will be unable to make use of this social comparison opportunity. To test this prediction, we asked participants to reassess themselves after they have seen the responses of their peers. We predicted that despite seeing the superior test performances of their classmates, bottom-quartile participants would continue to believe that they had performed competently.

In contrast, we expected that top-quartile participants, because they have the metacognitive skill to recognize competence and incompetence in others, would revise their self-ratings after the grading task. In particular, we predicted that they would recognize that the performances of the five individuals they evaluated were inferior to their own, and thus would raise their estimates of their percentile ranking accordingly. That is, top-quartile participants would learn from observing the responses of others, whereas bottom-quartile participants would not.

In making these predictions, we felt that we could account for an anomaly that appeared in all three previous studies: Despite the fact that top-quartile participants were far more calibrated than were their less skilled counterparts, they tended to underestimate their performance relative to their peers. We felt that this miscalibration had a different source then the miscalibration evidenced by bottom-quartile participants. That is, top-quartile participants did not underestimate themselves because they were wrong about their own performances, but rather because they were wrong about the performances of their peers. In essence, we believe they fell prey to the *false-consensus effect* (Ross, Greene, & House, 1977). In the absence of data to the contrary, they mistakenly assumed that their peers would tend provide the same (correct) answers as they themselves—an impression that could be immediately corrected by showing them the performances of their peers. By examining the extent to which competent individuals revised their ability estimates after grading the tests of their less competent peers, we could put this false-consensus interpretation to a test.

Method

Participants. Four to six weeks after Phase 1 of Study 3 was completed, we invited participants from the bottom- ($n = 17$) and top-quartile ($n = 19$) back to the laboratory in exchange for extra credit or $5. All agreed and participated.

Procedure. On arriving at the laboratory, participants received a packet of five tests that had been completed by other students in the first phase of Study 3. The tests reflected the range of performances that their peers had achieved in the study (i.e., they had the same mean and standard

The best acid test of our proposition, however, is to manipulate competence and see if this improves metacognitive skills and thus the accuracy of self-appraisals (Prediction 4). This would not only enable us to speak directly to causality, but would help rule out the regression effect alternative account discussed earlier. If the incompetent overestimate themselves simply because their test scores are very low (the regression effect), then manipulating competence after they take the test ought to have no effect on the accuracy of their self-appraisals. If instead it takes competence to recognize competence, then manipulating competence ought to enable the incompetent to recognize that they have performed poorly. Of course, there is a paradox to this assertion. It suggests that the way to make incompetent individuals realize their own incompetence is to make them competent.

In Study 4, that is precisely what we set out to do. We gave participants a test of logic based on the Wason selection task (Wason, 1966) and asked them to assess themselves in a manner similar to that in the previous studies. We then gave half of the participants a short training session designed to improve their logical reasoning skills. Finally, we tested the metacognitive skills of all participants by asking them to indicate which items they had answered correctly and which incorrectly (after McPherson & Thomas, 1989) and to rate their ability and test performance once more.

We predicted that training would provide incompetent individuals with the metacognitive skills needed to realize that they had performed poorly and thus would help them realize the limitations of their ability. Specifically, we expected that the training would (a) improve the ability of the incompetent to evaluate which test problems they had answered correctly and which incorrectly and, in the process, (b) reduce the miscalibration of their ability estimates.

Method

Participants. Participants were 140 Cornell University undergraduates from a single human development course who earned extra credit toward their course grades for participating. Data from 4 additional participants were excluded because they failed to complete the dependent measures.

Procedure. Participants completed the study in groups of 4 to 20 individuals. On arriving at the laboratory, participants were told that they would be given a test of logical reasoning as part of a study of logic. The test contained ten problems based on the Wason selection task (Wason, 1966). Each problem described four cards (e.g., *A, 7, B,* and *4*) and a rule about the cards (e.g., "If the card has a vowel on one side, then it must have an odd number on the other"). Participants then were instructed to indicate which card or cards must be turned over in order to test the rule.[4]

After taking the test, participants were asked to rate their logical reasoning skills and performance on the test relative to their classmates on a percentile scale. They also estimated the number of problems they had solved correctly.

Next, a random selection of 70 participants were given a short logical-reasoning training packet. Modeled after work by Cheng and her colleagues (Cheng, Holyoak, Nisbett, & Oliver, 1986), this packet described techniques for testing the veracity of logical syllogisms such as the Wason selection task. The remaining 70 participants encountered an unrelated filler task that took about the same amount of time (10 min) as did the training packet.

Afterward, participants in both conditions completed a metacognition task in which they went through their own tests and indicated which problems they thought they had answered correctly and which incorrectly. Participants then re-estimated the total number of problems they had answered correctly and compared themselves with their peers in terms of their general logical reasoning ability and their test performance.

Results and Discussion

Pretraining self-assessments. Prior to training, participants displayed a pattern of results strikingly similar to that of the previous three studies. First, participants overall overestimated their logical reasoning ability (M percentile = 64) and test performance (M percentile = 61) relative to their peers, paired $ts(139) = 5.88$ and 4.53, respectively, $ps < .0001$. Participants also overestimated their raw score on the test, $M = 6.6$ (perceived) versus 4.9 (actual), $t(139) = 5.95$, $p < .0001$. As before, perceptions of raw test score, percentile ability, and percentile test score correlated positively with actual test performance, $rs(138) = .50$, .38, and .40, respectively, $ps < .0001$.

Once again, individuals scoring in the bottom quartile ($n = 37$) were oblivious to their poor performance. Although their score on the test put them in the 13th percentile, they estimated their logical reasoning ability to be in the 55th percentile and their performance on the test to be in the 53rd percentile. Although neither of these estimates were significantly greater than 50, $t(36) = 1.49$ and 0.81, they were considerably greater than their actual percentile ranking, $ts(36) > 10$, $ps < .0001$. Participants in the bottom quartile also overestimated their raw score on the test. On average, they thought they had answered 5.5 problems correctly. In fact, they had answered an average of 0.3 problems correctly, $t(36) = 10.75$, $p < .0001$.

As Figure 4 illustrates, the level of overestimation once again decreased with each step up the quartile ladder. As in the previous studies, participants in the top quartile underestimated their ability. Whereas their actual performance put them in the 90th percentile, they thought their general logical reasoning ability fell in the 76th percentile and their performance on the test in the 79th percentile, $ts(27) < -3.00$, $ps < .001$. Top-quartile participants also underestimated their raw score on the test (by just over 1 point), but given that they all achieved perfect scores, this is hardly surprising.

Impact of training. Our primary hypothesis was that training in logical reasoning would turn the incompetent participants into experts, thus providing them with the skills necessary to recognize the limitations of their ability. Specifically, we expected that the training packet would (a) improve the ability of the incompetent to monitor which test problems they had answered correctly and which incorrectly and, thus, (b) reduce the miscalibration of their self-impressions.

Scores on the metacognition task supported the first part of this prediction. To assess participants' metacognitive skills, we summed the number of questions each participant accurately identified as correct or incorrect, out of the 10 problems. Overall, participants who received the training packet graded their own tests more accurately ($M = 9.3$) than did participants who did not receive the packet ($M = 6.3$), $t(138) = 7.32$, $p < .0001$, a difference even more pronounced when looking at bottom-quartile participants exclusively, $Ms = 9.3$ versus 3.5, $t(36) = 7.18$, $p < .0001$. In fact, the training packet was so successful that those who had originally scored in the bottom quartile were just as accurate in monitoring their test performance as were those who had ini-

[4] *A* and *4*.

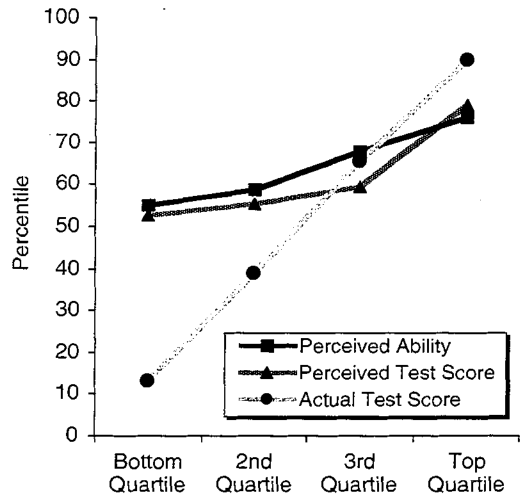

Figure 4. Perceived logical reasoning ability and test performance as a function of actual test performance (Study 4).

tially scored in the top quartile, *M*s = 9.3 and 9.9, respectively, $t(30) = 1.38$, *ns*. In other words, the incompetent had become experts.

To test the second part of our prediction, we examined the impact of training on participants' self-impressions in a series of 2 (training: yes or no) × 2 (pre- vs. postmanipulation) × 4 (quartile: 1 through 4) mixed-model analyses of variance (ANOVAs). These analyses revealed the expected three-way interactions for estimates of general ability, $F(3, 132) = 2.49$, $p < .07$, percentile score on the test, $F(3, 132) = 8.32$, $p < .001$, and raw test score, $F(3, 132) = 19.67$, $p < .0001$, indicating that the impact of training on self-assessment depended on participants' initial test performance. Table 2 displays how training influenced the degree of miscalibration participants exhibited for each measure.

To examine these interactions in greater detail, we conducted two sets of 2 (training: yes or no) × 2 (pre- vs. postmanipulation) ANOVAs. The first looked at participants in the bottom quartile, the second at participants in the top quartile. Among bottom-quartile participants, we found the expected interactions for estimates of logical reasoning ability, $F(1, 35) = 6.67$, $p < .02$, percentile test score, $F(1, 35) = 14.30$, $p < .002$, and raw test score, $F(1, 35) = 41.0$, $p < .0001$, indicating that the change in participants' estimates of their ability and test performance depended on whether they had received training.

As Table 2 depicts, participants in the bottom quartile who had received training ($n = 19$) became more calibrated in every way. Before receiving the training packet, these participants believed that their ability fell in the 55th percentile, that their performance on the test fell in the 51st percentile, and that they had answered 5.3 problems correctly. After training, these same participants thought their ability fell in the 44th percentile, their test in the 32nd percentile, and that they had answered only 1.0 problems correctly. Each of these changes from pre- to posttraining was significant, $t(18) = -2.53$, -5.42, and -6.05, respectively, *p*s < .03. To be sure, participants still overestimated their logical reasoning ability, $t(18) = 5.16$, $p < .0001$, and their performance on the test relative to their peers, $t(18) = 3.30$, $p < .005$, but they were considerably more calibrated overall and were no longer miscalibrated with respect to their raw test score, $t(18) = 1.50$, *ns*.

No such increase in calibration was found for bottom-quartile participants in the untrained group ($n = 18$). As Table 2 shows,

Table 2
Self-Ratings in Percentile Terms of Ability and Performance for Trained and Untrained Participants (Study 4)

	Untrained				Trained			
Rating	Bottom ($n = 18$)	Second ($n = 15$)	Third ($n = 22$)	Top ($n = 15$)	Bottom ($n = 19$)	Second ($n = 20$)	Third ($n = 18$)	Top ($n = 13$)
				Self-ratings of percentile ability				
Before	55.0	58.5	67.2	78.3	54.7	59.3	68.6	73.4
After	55.8	56.3	68.1	81.9	44.3	52.3	68.6	81.4
Difference	0.8	−2.1	0.9	3.6	−10.4*	−7.0*	0.1	8.0
Actual	11.9	32.2	62.9	90.0	14.5	41.0	69.1	90.0
				Self-ratings of percentile test performance				
Before	55.2	57.9	57.5	83.1	50.5	53.4	61.9	74.8
After	54.3	58.8	59.8	84.3	31.9	46.8	69.7	86.8
Difference	−0.8	0.9	2.3	1.3	−18.6***	−6.6*	7.8	12.1*
Actual	11.9	32.2	62.9	90.0	14.5	41.0	69.1	90.0
				Self-ratings of raw test performance				
Before	5.8	5.4	6.9	9.3	5.3	5.4	7.0	8.5
After	6.3	6.1	7.5	9.6	1.0	4.1	8.2	9.9
Difference	0.6*	0.7	0.6*	0.3	−4.3***	−1.4**	1.2**	1.5*
Actual	0.2	2.7	6.7	10.0	0.4	3.3	7.9	10.0

Note. "Bottom," "Second," "Third," and "Top" refer to quartiles on the grading task.
* $p < .05$. ** $p < .01$. *** $p < .001$.

they initially reported that both their ability and score on the test fell in the 55th percentile, and did not change those estimates in their second set of self-ratings, all $ts < 1$. Their estimates of their raw test score, however, did change—but in the wrong direction. In their initial ratings, they estimated that they had solved 5.8 problems correctly. On their second ratings, they raised that estimate to 6.3, $t(17) = 2.62$, $p < .02$.

For individuals who scored in the top quartile, training had a very different effect. As we did for their bottom-quartile counterparts, we conducted a set of 2 (training: yes or no) × 2 (pre- vs. postmanipulation) ANOVAs. These analyses revealed significant interactions for estimates of test performance, $F(1, 26) = 6.39$, $p < .025$, and raw score, $F(1, 26) = 4.95$, $p < .05$, but not for estimates of general ability, $F(1, 26) = 1.03$, *ns*.

As Table 2 illustrates, top-quartile participants in the training condition thought their score fell in the 78th percentile prior to receiving the training materials. Afterward, they increased that estimate to the 87th percentile, $t(12) = 2.66$, $p < .025$. Top-quartile participants also raised their estimates of their percentile ability, $t(12) = 1.91$, $p < .09$, and raw test score, $t(12) = 2.99$, $p < .025$, although only the latter difference was significant. In contrast, top-quartile participants in the control condition did not revise their estimates on any of these measures, $ts < 1$. Although not predicted, these revisions are perhaps not surprising in light of the fact that top-quartile participants in the training condition received validation that the logical reasoning they had used was perfectly correct.

The mediational role of metacognitive skills. We have argued that less competent individuals overestimate their abilities because they lack the metacognitive skills to recognize the error of their own decisions. In other words, we believe that deficits in metacognitive skills mediate the link between low objective performance and inflated ability assessment. The next two analyses were designed to test this mediational relationship more explicitly.

In the first analysis, we examined objective performance, metacognitive skill, and the accuracy of self-appraisals in a manner suggested by Baron and Kenny (1986). According to their procedure, metacognitive skill would be shown to mediate the link between incompetence and inflated self-assessment if (a) low levels of objective performance were associated with inflated self-assessment, (b) low levels of objective performance were associated with deficits in metacognitive skill, and (c) deficits in metacognitive skill were associated with inflated self-assessment even after controlling for objective performance. Focusing on the 70 participants in the untrained group, we found considerable evidence of mediation. First, as reported earlier, participants' test performance was a strong predictor of how much they overestimated their ability and test performance. An additional analysis revealed that test performance was also strongly related to metacognitive skill, $\beta(68) = .75$, $p < .0001$. Finally, and most important, deficits in metacognitive skill predicted inflated self-assessment on the all three self-ratings we examined (general logical reasoning ability, comparative performance on the test, and absolute score on the test)—even after objective performance on the test was held constant. This was true for the first set of self-appraisals, $\beta s(67) = -.40$ to $-.49$, $ps < .001$, as well as the second, $\beta s(67) = -.41$ to $-.50$, $ps < .001$.[5]

Given these results, one could wonder whether the impact of training on the self-assessments of participants in the bottom quartile was similarly mediated by metacognitive skills. To find out, we conducted a mediational analysis focusing on bottom quartile participants in both trained and untrained groups. Here too, all three mediational links were supported. As previously reported, bottom-quartile participants who received training (a) provided less inflated self-assessments and (b) evidenced better metacognitive skills than those who did not receive training. Completing this analysis, regression analyses revealed that metacognitive skills predicted inflated self-assessment with participants' training condition held constant, $\beta(34)s = -.68$ to $-.97$, $ps < .01$. In fact, training itself failed to predict miscalibration when bottom-quartile participants' metacognitive skills were taken into account, $\beta s(34) = .00$ to $.25$, *ns*. These analyses suggest that the benefit of training on the accuracy of self-assessment was achieved by means of improved metacognitive skills.[6]

Summary. Thomas Jefferson once said, "he who knows best best knows how little he knows." In Study 4, we obtained experimental support for this assertion. Participants scoring in the bottom quartile on a test of logic grossly overestimated their test performance—but became significantly more calibrated after their logical reasoning skills were improved. In contrast, those in the bottom quartile who did not receive this aid continued to hold the mistaken impression that they had performed just fine. Moreover, mediational analyses revealed that it was by means of their improved metacognitive skills that incompetent individuals arrived at their more accurate self-appraisals.

General Discussion

In the neurosciences, practitioners and researchers occasionally come across the curious malady of anosognosia. Caused by certain types of damage to the right side of the brain, anosognosia leaves people paralyzed on the left side of their body. But more than that, when doctors place a cup in front of such patients and ask them to pick it up with their left hand, patients not only fail to comply but also fail to understand why. When asked to explain their failure, such patients might state that they are tired, that they did not hear the doctor's instructions, or that they did not feel like responding—but never that they are suffering from paralysis. In essence, anosognosia not only causes paralysis, but also the inability to realize that one is paralyzed (D'Amasio, 1994).

In this article, we proposed a psychological analogue to anosognosia. We argued that incompetence, like anosognosia, not only causes poor performance but also the inability to recognize that one's performance is poor. Indeed, across the four studies, participants in the bottom quartile not only overestimated themselves, but thought they were above-average, $Z = 4.64$, $p < .0001$. In a

[5] A mediational analysis of the absolute miscalibration (independent of sign) in participants' self-appraisals revealed a similar pattern: Controlling for objective performance on the test, deficits in metacognitive skill predicted absolute miscalibration on the all three self-ratings we examined for both the first set of self-appraisals, $\beta s(67) = -.53$ to $-.78$, $ps < .001$, and the second, $\beta s(67) = -.60$ to $-.79$, $ps < .0001$

[6] An analysis of the absolute miscalibration (independent of sign) revealed a similar pattern. Controlling for objective performance on the test, deficits in bottom-quartile participants' metacognitive skill predicted their absolute miscalibration on all three of the self-ratings, $\beta s(34) = -.79$ to $-.98$, $ps < .01$.

Limitations of the Present Analysis

We do not mean to imply that people are always unaware of their incompetence. We doubt whether many of our readers would dare take on Michael Jordan in a game of one-on-one, challenge Eric Clapton with a session of dueling guitars, or enter into a friendly wager on the golf course with Tiger Woods. Nor do we mean to imply that the metacognitive failings of the incompetent are the only reason people overestimate their abilities relative to their peers. We have little doubt that other factors such as motivational biases (Alicke, 1985; Brown, 1986; Taylor & Brown, 1988), self-serving trait definitions (Dunning & Cohen, 1992; Dunning et al., 1989), selective recall of past behavior (Sanitioso, Kunda, & Fong, 1990), and the tendency to ignore the proficiencies of others (Klar, Medding, & Sarel, 1996; Kruger, 1999) also play a role. Indeed, although bottom-quartile participants accounted for the bulk of the above-average effects observed in our studies (overestimating their ability by an average of 50 percentile points), there was also a slight tendency for the other quartiles to overestimate themselves (by just over 6 percentile points)—a fact our metacognitive analysis cannot explain.

When can the incompetent be expected to overestimate themselves because of their lack of skill? Although our data do not speak to this issue directly, we believe the answer depends on the domain under consideration. Some domains, like those examined in this article, are those in which knowledge about the domain confers competence in the domain. Individuals with a great understanding of the rules of grammar or inferential logic, for example, are by definition skilled linguists and logicians. In such domains, lack of skill implies both the inability to perform competently as well as the inability to recognize competence, and thus are also the domains in which the incompetent are likely to be unaware of their lack of skill.

In other domains, however, competence is not wholly dependent on knowledge or wisdom, but depends on other factors, such as physical skill. One need not look far to find individuals with an impressive understanding of the strategies and techniques of basketball, for instance, yet who could not "dunk" to save their lives. (These people are called coaches.) Similarly, art appraisers make a living evaluating fine calligraphy, but know they do not possess the steady hand and patient nature necessary to produce the work themselves. In such domains, those in which knowledge about the domain does not necessarily translate into competence in the domain, one can become acutely—even painfully—aware of the limits of one's ability. In golf, for instance, one can know all about the fine points of course management, club selection, and effective "swing thoughts," but one's incompetence will become sorely obvious when, after watching one's more able partner drive the ball 250 yards down the fairway, one proceeds to hit one's own ball 150 yards down the fairway, 50 yards to the right, and onto the hood of that 1993 Ford Taurus.

Finally, in order for the incompetent to overestimate themselves, they must satisfy a minimal threshold of knowledge, theory, or experience that suggests to themselves that they can generate correct answers. In some domains, there are clear and unavoidable reality constraints that prohibits this notion. For example, most people have no trouble identifying their inability to translate Slovenian proverbs, reconstruct an 8-cylinder engine, or diagnose acute disseminated encephalomyelitis. In these domains, without even an intuition of how to respond, people do not overestimate their ability. Instead, if people show any bias at all, it is to rate themselves as worse than their peers (Kruger, 1999).

Relation to Work on Overconfidence

The finding that people systematically overestimate their ability and performance calls to mind other work on calibration in which people make a prediction and estimate the likelihood that the prediction will prove correct. Consistently, the confidence with which people make their predictions far exceeds their accuracy rates (e.g., Dunning, Griffin, Milojkovic, & Ross, 1990; Vallone, Griffin, Lin, & Ross, 1990; Lichtenstein, Fischhoff, & Phillips, 1982).

Our data both complement and extend this work. In particular, work on overconfidence has shown that people are more miscalibrated when they face difficult tasks, ones for which they fail to possess the requisite knowledge, than they are for easy tasks, ones for which they do possess that knowledge (Lichtenstein & Fischhoff, 1977). Our work replicates this point not by looking at properties of the task but at properties of the person. Whether the task is difficult because of the nature of the task or because the person is unskilled, the end result is a large degree of overconfidence.

Our data also provide an empirical rebuttal to a critique that has been leveled at past work on overconfidence. Gigerenzer (1991) and his colleagues (Gigerenzer, Hoffrage, & Kleinbölting, 1991) have argued that the types of probability estimates used in traditional overconfidence work—namely, those concerning the occurrence of single events—are fundamentally flawed. According to the critique, probabilities do not apply to single events but only to multiple ones. As a consequence, if people make probability estimates in more appropriate contexts (such as by estimating the total number of test items answered correctly), "cognitive illusions" such as overconfidence disappear. Our results call this critique into question. Across the three studies in which we have relevant data, participants consistently overestimated the number of items they had answered correctly, $Z = 4.94$, $p < .0001$.

Concluding Remarks

In sum, we present this article as an exploration into why people tend to hold overly optimistic and miscalibrated views about themselves. We propose that those with limited knowledge in a domain suffer a dual burden: Not only do they reach mistaken conclusions and make regrettable errors, but their incompetence robs them of the ability to realize it. Although we feel we have done a competent job in making a strong case for this analysis, studying it empirically, and drawing out relevant implications, our thesis leaves us with one haunting worry that we cannot vanquish. That worry is that this article may contain faulty logic, methodological errors, or poor communication. Let us assure our readers that to the extent this article is imperfect, it is not a sin we have committed knowingly.

References

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.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology, 68,* 804–825.

Allen, W. (1975). *Without feathers.* New York, NY: Random House.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research. *Journal of Personality and Social Psychology, 51,* 1173–1182.

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.

Bobrow, J., Nathan, N., Fisher, S., Covino, W. A., Orton, P. Z., Bobrow, B., & Weber, L. (1989). *Cliffs NTE preparation guide.* Lincoln, NE: Cliffs Notes, Inc.

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition, 4,* 353–376.

Brown, J. D., & Gallagher, F. M. (1992). Coming to terms with failure: Private self-enhancement and public self-effacement. *Journal of Experimental Social Psychology, 28,* 3–22.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293–328.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 17–76). Hillsdale, NJ: Erlbaum.

College Board. (1976–1977). *Student descriptive questionnaire.* Princeton, NJ: Educational Testing Service.

Cross, P. (1977). Not *can* but *will* college teaching be improved? *New Directions for Higher Education, 17,* 1–15.

D'Amasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York: Putnam.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35,* 867–881.

Darwin, C. (1871). *The descent of man.* London: John Murray.

Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities in social judgment. *Journal of Personality and Social Psychology, 63,* 341–355.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58,* 568–581.

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., Perie, M., & Story, A. L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology, 61,* 957–968.

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.

Everson, H. T., & Tobias, S. (1998). The ability to estimate knowledge and performance in college: A metacognitive analysis. *Instructional Science, 26,* 65–79.

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.

Felson, R. B. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly, 44,* 64–69.

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

Frankin, A. (1992). *Deep Thoughts by Jack Handy.* New York: Berkley Publishing Group.

Fuocco, M. A. (1996, March 21). Trial and error: They had larceny in their hearts, but little in their heads. *Pittsburgh Post-Gazette,* p. D1.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases." *European Review of Social Psychology, 2,* 83–115.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswickian theory of confidence. *Psychological Review, 98,* 506–528.

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

Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes, 18,* 213–231.

Klar, Y., Medding, A., & Sarel, D. (1996). Nonunique invulnerability: Singular versus distributional probabilities and unrealistic optimism in comparative risk judgments. *Organizational Behavior and Human Decision Processes, 67,* 229–245.

Klin, C. M., Guizman, A. E., & Levine, W. H. (1997). Knowing that you don't know: Metamemory and discourse processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1378–1393.

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

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

Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology, 62,* 194–198.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? The calibration of probability judgments. *Organizational Behavior and Human Performance, 20,* 159–193.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Ed.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

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

Matlin, M., & Stang, D. (1978). *The Pollyanna principle.* Cambridge, MA: Schenkman.

McPherson, S. L., & Thomas, J. R. (1989). Relation of knowledge and performance in boys' tennis: Age and expertise. *Journal of Experimental Child Psychology, 48,* 190–211.

Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality and Social Psychology Review, 2,* 100–110.

Miller, W. I. (1993). *Humiliation.* Ithaca, NY: Cornell University Press.

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

Orton, P. Z. (1993). *Cliffs Law School Admission Test preparation guide.* Lincoln, NE: Cliffs Notes Incorporated.

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attributional processes. *Journal of Experimental Social Psychology, 13,* 279–301.

Rovin, J. (1996). *More Really Silly Pet Jokes.* Boca Raton, FL: Globe Communications Corp.