Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Experimental Social Psychology xxx (2005) xxx–xxx
Journal of Experimental Social Psychology
www.elsevier.com/locate/jesp

What you don't know: The role played by errors of omission in imperfect self-assessments☆

Deanna Caputo, David Dunning*

Department of Psychology, Cornell University, Ithaca, NY 14853, USA

Received 21 March 2004

Abstract

Perceptions of ability often bear little relationship to objective performance. We suggest that people fail to judge their ability more accurately because they have little or no insight into their *errors of omission* (i.e., solutions they could have generated to problems but missed), although they can be perfectly aware of solutions found. Across five studies with tasks involving, for example, word games and research methodology, we found that participants gave weight to the number of solutions found when making self-evaluations, but not to solutions missed. When given explicit information about these errors of omission, participants gave them just as much weight as they did solutions found, and thus provided more accurate self-evaluations.

Keywords: Errors of omission; Self-perception; Self-assessment; Self-evaluation; Ill-defined problems

I am not ashamed to confess I am ignorant of what I do not know.

Cicero, Roman Orator

Cicero not withstanding, people do have difficulty recognizing and admitting the limits of their expertise. A good deal of psychological research over the past 25 years has shown that people frequently overestimate their skill to a logically impossible degree (Alicke, 1985; College-Board, 1976–1977; Dunning, Meyerowitz, & Holzberg, 1989; Epley & Dunning, 2000; Weinstein, 1980) and express too much confidence in their judgments and predictions (Dunning, Griffin, Milojkovic, & Ross, 1990; Dunning & Story, 1991; Lichtenstein & Fischhoff, 1977). Indeed, it is exactly when people are most unknowledgeable that they are the least successful at identifying, and thus reporting, the breadth and depth of their ignorance. People who are incompetent—that is, performing poorly relative to their peers—tend not to be aware that their skills are so impoverished. Instead, they tend to think they are doing just fine, and are almost as confident in their abilities and performance as are their more competent peers (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Haun, Zeringue, Leach, & Foley, 2000; Hodges, Regehr, & Martin, 2001; Kruger & Dunning, 1999).

People also have a difficult time separating their strengths from their shortcomings, in that the perceptions they hold of their talents and performances often correlate modestly, and sometimes not at all, with the truth. In a comprehensive review, Mabe and West (1982) discovered that people's impressions of their skills and performances tended to correlate roughly .29 with reality, a finding that has been more or less replicated in edu-

☆ Many thanks to Dennis Regan, Melissa Ferguson, and the Dunning Lab for assistance with material design. We also thank Amy Tsai, David Chen, and Tarik Scott for all their assistance in data collection. This research was supported financially by National Institute of Mental Health Grant RO1 56072. Portions of this research were presented at the annual meeting of the Society for Personality and Social Psychology, Savannah, January, 2002, and the annual meeting of the Society for Personality and Social Psychology, Austin, 2004.

* Corresponding author. Fax: +1 607 255 8433.
E-mail addresses: ddc22@cornell.edu (D. Caputo), dad6@cornell.edu (D. Dunning).

0022-1031/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2004.09.006

cational settings (Falchikov & Boud, 1989; Hansford & Hattie, 1982) and organizational ones (Harris & Schaubroeck, 1988). Certainly, depending on the domain, correlations between perception and reality can be quite high, such as they are with athletes' estimates of their physical skills ($r = .47$), but in many important domains such correlations barely sustain a pulse. With managerial tasks, the typical correlation is .04; with interpersonal skills tasks it is .17 (Mabe & West, 1982).

In this paper, we examine why impressions of performance tend to hold only modest relationships with actual achievement. We suggest that the act of judging one's performance is an intrinsically difficult one, in that people tend not to have—indeed, cannot be expected to have—all the relevant information they need to reach accurate self-impressions. It is not that people are deficient in their self-evaluation skills; they would give far more accurate appraisals if they had all the information in hand.

Consider the following example. A popular way to keep children (and some boisterous adults) busy on road trips or in summer camp is to give them an object and ask them to find as many legitimate uses for the object as possible. For instance, how many household uses are there for WD-40, a common oil lubricant? WD-40 can be used for many purposes. It can be used to lubricate bicycle chains or remove squeaks from doors, but it can also be used to unclog showerheads. Suppose one spent an hour thinking of uses for WD-40 and produced 15 legitimate uses for the oil. How good is this performance?

Evaluating a performance at this task clearly involves taking note of how many uses someone has found, but it also importantly depends on another piece of information—how many legitimate uses are possible that the person failed to generate. If the person misses only two possible uses, his or her performance is much more impressive than if he or she missed 50. But here is the rub—how is a person, evaluating their own performance, to estimate how many solutions he or she has missed? This WD-40 game illustrates a basic difficulty facing people as they strive to evaluate their performances. They can be quite aware of the solutions they find to apply to some task, but by definition they are unaware of the solutions they could have generated but missed. People simply cannot be expected to be aware of this important type of information, although its relevance for reaching accurate self-judgment is obvious.

In a phrase, we propose that people have no magical access to the number and magnitude of their *errors of omission*, and thus cannot be expected to make accurate self-evaluations. To be sure, they would act like Cicero and acknowledge their errors of omission—and incorporate them into their self-evaluations—if they only knew of them. They simply do not know, nor can they be expected to know. In reality, there are over 2000 legitimate household uses for WD40 (Wheelmen, 2003). WD-40, for example, can remove Kool-Aid and lipstick stains from carpet and fabrics, camouflage scratches in ceramic tile or linoleum, as well as, repel pigeons from around household gardens and driveways.

This problem of taking into account and incorporating errors of omission into self-evaluation has not gone unnoticed by others. Fatsis (2002), for example, said in his book *Word freak*, which is about the atypical world of people trying to make a living playing the board game Scrabble:

> In a way, the living-room player is lucky … He has no idea how miserably he fails with almost every turn, how many possible words or optimal plays slip by unnoticed. The idea of Scrabble greatness doesn't exist for him (p. 128).

In the research reported in this paper, we examined the role played by errors of omission in imperfect self-assessments. We presented participants with tasks that could elicit a number of possible solutions and then asked them to evaluate how well they had done. We expected that participants' evaluations would give weight to the number of solutions they had found but would show little, if any, insight into the number of solutions they had missed. Because of this ignorance, their self-evaluations would not reflect their errors of omission. Undoubtedly, if participants were told about their errors of omission, they would consider those omissions relevant and would change their self-evaluations in a more accurate direction. However, left to their own devices, participants would not be able to anticipate how many errors of omission they had made.

Relation to previous work on judgments of skill

A careful reader may wonder about the relation of this analysis to other recent research suggesting that people are at times not in a position to evaluate their skills adequately because they lack the requisite knowledge or information. In particular, Kruger and Dunning (1999; see also Dunning et al., 2003) argued that incompetent individuals should not be expected to recognize their poor performances because their deficits present them with two problems. First, their deficits prevent them from reaching accurate decisions. Second, their deficits prevent them from recognizing when accurate—and erroneous—decisions have been made. Because they cannot recognize the poor quality of their decisions, they tend to think they are doing just fine.

The present research bears some similarity to this past work on incompetence (Dunning et al., 2003; Kruger & Dunning, 1999), in that it focuses on barriers that

prevent people from reaching accurate self-assessments. However, the present research differs in two important ways in its analysis. First, the Kruger and Dunning (1999) analysis described a circumstance in which some individuals, namely poor performers, fail to recognize correct responses even when those responses are laid out in front of them. In essence, poor performers, because of their deficits, fail to have the metacognitive wherewithal to differentiate accurate from inaccurate responses. In the present research, we explored situations in which people have adequate metacognitive skill and, thus, could recognize correct solutions when they were presented to them. The barrier they faced to accurate self-perception was just that they had no awareness of these correct responses until some outside agent pointed them out.

Second, Kruger and Dunning (1999) focused their analysis on poor performers. In the current research, we proposed that the difficulty of omission errors would prompt people all along the dimension of competence to have difficulty correctly assessing their level of skill. Not knowing the extent of their omission errors, poor performers will overestimate their skill and performance. High performers, not knowing how few errors of omission they made, may underrate theirs. Indeed, if high performers only knew how comprehensive their performances were, they would have a more favorable, and more appropriate, view of their ability.

Predictions

Thus, we made five specific predictions about the links between peoples' self-evaluations, errors of omission, and actual performance.

Prediction 1

Peoples' evaluations of their ability and performance would largely reflect the number of solutions that they find at a task and would largely fail to reflect the number of solutions they miss (i.e., their errors of omission).

Prediction 2

People would be largely unable to predict how many errors of omission they had made.

Prediction 3

In evaluations of their performance, people would consider errors of omission to be relevant and would give weight to omission errors, if they only knew about them. Thus, peoples' evaluations of their ability and performance would correlate with their errors of omission if they were informed of them.

Prediction 4

Informing people of their errors of omission would reduce how positively they viewed their skills and ability. If people underestimated the number of their omission errors, informing people of those errors would reduce how positively they viewed their skills and ability. However, if they overestimated those errors, the same information would prompt them to raise their self-evaluations.

Prediction 5

If informed of their errors of omission, peoples' evaluations of their ability and performance will be more accurate, relative to objective performance criteria that they themselves cite as appropriate.

The Studies

We explored all five predictions by presenting participants with tasks that had multiple solutions, such as a popular word game (Studies 1 and 5), a visual search task (Study 2), a grammar exercise (Study 3), or a methodological critique of scientific research (Study 4). In Studies 1–4 we then asked participants to assess their general ability to perform the task as well as their specific performance. We predicted that participants' self-evaluations would largely reflect the number of solutions they found but not the number that they had missed (Prediction 1). When asked, participants would not be able to predict with great accuracy the number of solutions they had missed (Prediction 2).

Participants were then informed of the total number of possible solutions to the task—and thus their errors of omission—and were asked again to assess their general ability and specific performance on the task. We predicted that participants' evaluations would give significant weight to their errors of omission once they were informed of them, thus significantly revising their self-evaluations (Predictions 3 and 4). Their evaluations would, thus, be more tightly correlated with objective measures of performance that they themselves considered appropriate (Prediction 5).

In addition, we also wanted to show that participants would use their self-evaluations to guide their actions even when they did not have knowledge about their errors of omission. Thus, in Study 5, we asked participants whether they would bet money based on their performance, with half of the participants knowing about their errors of omission and half not. We expected participants would bet more if they did not know about their errors of omission, but would place more conservative bets once they knew.

Study 1: Word play

In Study 1, we wanted to explore people's self-evaluations on a task that had a large number of solutions that would not be immediately evident to them. However, it was important that the task have a finite set of solutions so that we could test our prediction related to their actual performance. We used the board game Boggle in which one is given a 4×4 array of alphabetical letters and asked to find as many words as possible, within a short time limit, by chaining adjacent letters together.

We asked participants to complete three Boggle puzzles. Then, we asked them to rate their ability at this type of task as well as their specific performance on the task. Subsequently, we gave them information about the total words actually contained in each puzzle and asked again the same questions about ability and performance. Our main interest was whether participants' self-evaluations would change once they were given the answers to the puzzles, and whether these changes would be reflected in the weight given to the number of words missed versus found. We were also interested in whether the association between self-evaluations and objective performance would increase once they had access to their errors of omission.

Method

Participants

Participants were 69 Cornell University undergraduates from a variety of courses in psychology and human development who earned extra credit for their participation.

Procedure

Participants were run in groups and all experimenters were blind to condition and the specific hypotheses of the study. Upon arrival participants were given a packet with three different puzzles from the classic board game Boggle. Boggle is a hidden word game made by Parker Brothers. Each puzzle contained a 4×4 array of letters, and participants were asked to identify as many English words that could be found by connecting adjacent letters. Participants were instructed about the rules of Boggle and given a few examples. The rules state that words can only be formed from adjoining letters. Letters must join in the proper sequence to spell a word, but they may join horizontally, vertically, or diagonally, to the left, right, or up-and-down. However, no letter cube may be used more than once within a single word. Words must be at least three letters long, and cannot be proper names. Participants were given 3 min for each puzzle.

Participants were randomly assigned to one of three different sets of puzzles. One set contained a relatively low number of total solutions across three puzzles (100 possible words, $n = 23$), another had a medium number (179 possible words, $n = 22$), and one had a high number (254 possible words, $n = 24$). These different puzzle sets were provided to ensure that the number of words that participants found would not artifactually correlate with the number of words they missed across the study.

Following completion of the three puzzles, participants evaluated their performance on two questions: "How would you rate your overall ability at finding words in a Boggle puzzle?" and "How would you rate your performance on the three puzzles you just confronted?" Participants responded on Likert scales from 1 (extremely bad) to 9 (extremely good). Participants next estimated the number of words they found for each puzzle, as well as the number of words they missed in each puzzle. They also rated on 9-point Likert scales how relevant these two pieces of information were to assessing how well a person had done on the Boggle puzzle. Next, participants were presented with four possible criteria for evaluating how well a person had done on the boggle puzzles and were asked to rank order the criteria from the most to the least appropriate. The criteria included: number of words found, the number of words found compared with the total number of words actually in the puzzle, the number of words found compared with the number the highest scorer found, and the total number of words found compared with the average number of words found by people in the experiment.

After indicating that they had completed these questions, participants were given a "final information packet" that contained information about the total number of words that were actually in each of the puzzles; this answer sheet did not include the actual words in the Boggle puzzle, just the total possible. After looking over the information participants answered again the two self-evaluation questions on their overall Boggle ability and performance on the three puzzles. Participants were then debriefed and thanked.

Results and discussion

Our first question was whether participants showed any insight into their errors of omission. We predicted they would show little or no insight—that evaluations of performance would reflect the number of solutions participants found in the puzzles, but not solutions missed.[1] To test this prediction, we performed multiple regression analyses in which we predicted self-evaluations of ability and performance from number of solutions found and number of solutions missed. As expected, self-evaluations reflected the number of solutions found ($bs = .63$ and .28, $ps < .01$, for ability and performance, respec-

[1] In supplemental analyses in each of the studies reported in the paper, we examined whether performance or self-ratings differed systematically across low, medium, and high solution experimental conditions. We found no reliable differences that replicated across studies, and so do not discuss potential group differences further.

tively) but not the number missed ($bs = -.00$ and .06, ns, for ability and performance, respectively).

But did this neglect of omission errors arise because participants did not know of their omission errors? Perhaps instead they did have a sense of how many solutions they had missed but decided not to give this datum weight when they evaluated themselves. Additional data suggested that it was lack of insight, and not lack of caring, that led participants not to give weight to their omission errors when providing self-evaluations. Participants' predictions of how many words they missed in the puzzles were not correlated with how many they actually missed, $r\,(68) = -.15$, ns, therefore, participants showed no insight into how many errors of omission they had made.

In addition, when we informed participants about the total number of solutions included in each puzzle, and thus notified them of their errors of omission, they did give weight to those errors in their final self-evaluations. In a multiple regression analysis that included number of solutions found and number missed as independent variables, we found that participants gave significant weight to their errors of omission ($bs = -.29$, $p < .02$, and $-.23$, $p < .06$, for ability and performance, respectively), as well as to the number of solutions they had found ($bs = .59$ and .36, $ps < .01$, for ability and performance, respectively).

This greater weight given to errors of omission, once known, had two significant influences on self-evaluations. First, overall, informing participants of their errors of omission caused them to significantly lower their self-evaluations. The average rating of ability fell from 5.8 to 3.8, $t(68) = -10.55$, $p < .001$, and the average performance rating fell from 5.3 to 2.9, $t(68) = -12.33$, $p < .001$. This decrease is not surprising, given that participants on average underestimated their omission errors by nearly 136 words. They thought they had missed 18.3 words across the three puzzles, when in fact they missed 153.9, $t(68) = -17.74$, $p < .0001$.

Second, participants' self-ratings were more accurate, given the objective criterion that they endorsed as the most appropriate for judging their performance. In the questionnaire participants completed before receiving the puzzle answers, they were asked to rank order four possible criteria for assessing how well someone had done on the boggle puzzles. The questionnaire data showed that the majority of participants (57%) thought that the percent correct (e.g., total words found divided by total words possible) was the most appropriate measure of performance for this task and 78% of participants ranked actual percent correct as one of their top two choices for determining performance. There was significant variability in performance based on this criterion, some got 3.9% correct while others got as high as 35% correct. Participants' self-ratings correlated with this measure more tightly after being informed of their

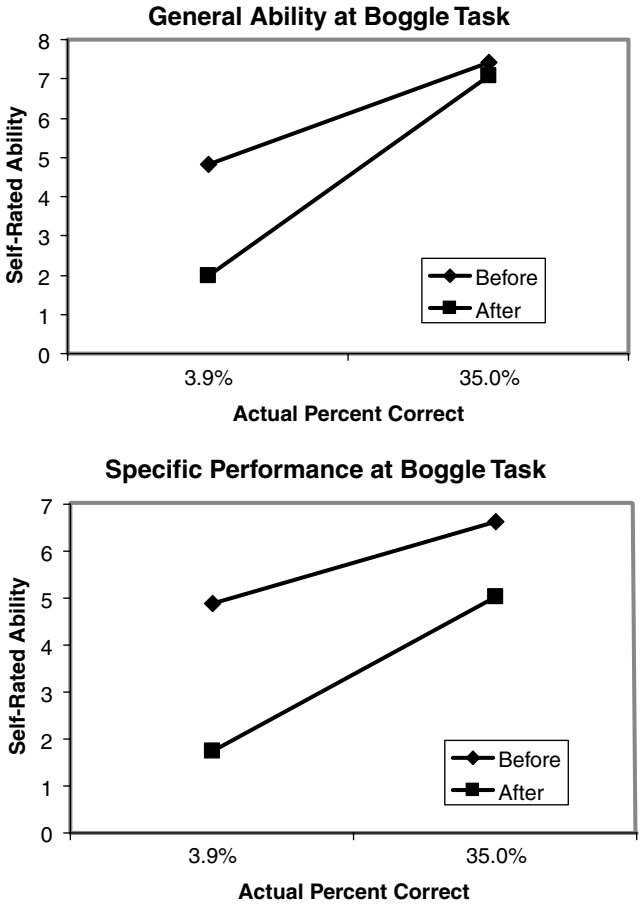

Fig. 1. The association between self-evaluations and actual performance before and after the total number of solutions were revealed (Study 1).

omission errors than those ratings did before. Correlational analyses showed that the association between participants' actual percentage correct and their self-evaluations increased from their pre-evaluations ($r = .43$, $p < .001$; $r = .21$, $p < .08$, for ability and performance) to their post-evaluations ($r = .63$, $p < .001$; $r = .47$, $p < .001$, for ability and performance).

More importantly, additional analyses revealed that this improvement in accuracy was statistically significant. Fig. 1 depicts the relationship between self-ratings and actual performance both before and after learning about the total number of solutions present in the puzzles. As seen in the figure, the regression slope between self-perception and actual performance is steeper after participants were informed than it was before. The difference between these two regression slopes is significant. This was most directly shown by the fact that differences in self-assessments (e.g., post-ratings minus pre-ratings) were positively correlated with actual performance, indicating that post-ratings shifted significantly to be more tightly related to actual performance ($r = .38$, $p < .001$ for general ability; $r = .29$, $p < .02$ for specific performance). This figure also shows that

although participants decreased their self-evaluations in general, participants at the bottom end changed more than those at the top.

Summary

In sum, Study 1 revealed that participants failed to take into account their errors of omission when making self-evaluations. They showed little insight about the entire set of solutions that they could have generated in a task when determining how well they performed on that task. However, when participants were confronted with their errors of omission they had no difficulty giving weight to that information in further evaluations of their performance. This caused them to provide more humble self-evaluations. It also prompted them to provide evaluations that more closely correlated with the metric they nominated as an appropriate measure of objective performance, actual percent correct.

Taken together, these findings suggest that it is not that people do give weight to errors of omission in their self-evaluations—once they know about them. However, without being explicitly informed of those errors, they are not necessarily able to assess how many of them they have made. Once informed of their omission errors, they give such errors weight, and their self-evaluations correlate much higher with the metric of performance that they see as most valuable to the task.

Study 2: Visual search

In Study 2, we explored self-evaluations on a visual task that did not have a large number of solutions but that contained solutions that would not be immediately evident to participants. Again, it was important that the task have a finite set of solutions so that we could test our prediction related to their actual performance. We used caricatures from a collection of Al Hirschfeld's works in *The world of Hirschfeld* (1970). In order to add something different to his caricatures of celebrities, Broadway stars, and Hollywood scenes, Mr. Hirschfeld imbedded the name of his daughter "NINA" in the background all of his drawings after 1945. For example, the flowing lines of the NINA were occasionally worked into the folds of drapery or strands of hair. The insertion of NINA's became a game he played with his audience, and in 1956 he actually appended a numeral after his signature to tell his fans how many NINA's he had hidden.

In our study, participants were given some of these caricatures and asked to find as many NINA's as possible in the time allotted. After completing four such pictures, participants were asked a question about their general ability at this type of task and a question about their specific performance on the four pictures we had presented to them. Then, they were told how many total NINA's were contained in each picture and asked again the same questions about ability and performance. The actual NINA's were circled on the answer sheets.

As a replication, our main focus was whether participants' initial self-evaluations would reveal any awareness of their errors of omission. However, in this study we wanted to replicate the findings of Study 1 with a visual search task instead—something a bit less novel than the Boggle game. Again, would participants' self-evaluations give appropriate, or any weight, to the number of words they had missed, or would their evaluations reflect only the number of solutions they had found? Would they give weight to these omission errors after receiving them? Would it change their self-evaluations in general? Would it make their self-evaluations more accurate relative to an objective criterion they cited as appropriate?

Method

Participants

Participants were 55 Cornell University undergraduates from a variety of courses in psychology and human development who earned extra credit for their participation.

Procedure

Participants arrived in groups but were seated far enough apart so they could not see other participants' packets, and all experimenters were blind to condition and the specific hypotheses of the study. Participants were given a packet with four different Hirschfeld caricatures. They were given 15 min to find as many "NINA's" as possible in all four pictures. Participants were randomly assigned to either a low (8 NINA's, $n = 19$), a medium (14 NINA's, $n = 18$), or a high (20 NINA's, $n = 18$) condition. This was, again, to ensure that the number of words found would not correlate with the number of words missed in the analyses of the data. Before beginning, participants were given a few examples of NINA's circled in a sample picture.

Following the puzzles, participants were given the same packet of questions as in Study 1. They rated their overall ability at finding words in a picture and their performance on the four pictures they had just confronted, using 9-point Likert scales modeled after those used in Study 1. The next page had participants answer a number of other questions concerning the task (i.e., number of total words found and missed, rank ordering criteria for assessing ability at task). When participants reached the bottom of the third page they were asked to inform the experimenter that they were ready for the "final information packet," which contained the four pictures with all the NINA's circled, and the total number of words that were actually in each of the pictures was written on each picture. After looking over the information,

participants turned to the last page and answered exactly the same two questions on the first page of the packet, one question on overall ability at finding words in pictures and a second question on performance on the four pictures confronted. Participants were then debriefed and thanked.

Results and discussion

We were again primarily interested in seeing whether participants would show any insight into their errors of omission. Multiple regression analyses revealed that self-evaluations, surprisingly, did not reflect the number of solutions found ($bs = .26$, $p < .06$ and .13, ns, for ability and performance, respectively), although the correlation was in the right direction. However, as expected, self-evaluations did not reflect the number missed ($bs = .03$ and .02, ns, for ability and performance, respectively).

Was this neglect of omission errors because participants did not know or because they did not consider them important? Additional data, again, suggested that participants did consider errors of omission to be important—they just could not accurately anticipate the number of such errors. Participants' predictions of how many words they missed in the puzzles did not correlate with how many they actually missed ($r(45) = .12$, ns). In addition, when we informed participants about the total number of solutions included in each puzzle, thus notifying them of their errors of omission, they did give those errors weight in their self-evaluations. In a multiple regression analysis, we found that participants gave significant weight to their errors of omission ($bs = -.51$ and $-.50$, $ps < .001$, for ability and performance, respectively), but no weight to the number of solutions they had found ($bs = .07$ and .10, ns, for ability and performance, respectively).

This greater weight given to errors of omission had two significant influences on self-evaluations. First, informing participants of their errors of omission caused them to significantly *increase* their self-evaluations. The average rating of ability went from 6.2 to 7.0, $t(54) = 3.72$, $p < .001$; the average performance rating jumped from 5.7 to 7.1, $t(54) = 6.80$, $p < .001$. This result apparently follows from the fact that they tended to overestimate their errors of omission, thinking that they had missed 6.5 NINA's on average when they, in fact, missed only 2.3, $t(55) = 4.15$, $p < .001$.

Questionnaire data, as in Study 1, showed that the majority of participants (60%) thought that percent correct (e.g., total words found divided by total words possible) was the most appropriate measure of performance for this task, and 81% of participants ranked actual percent correct as one of their top two choices for determining performance. There was significant variability in performance based on this criterion, percent correct ranged from 50 to 100% correct. Correlational analyses showed that the association between participants' actual percent correct and their self-evaluations increased from their pre-evaluations ($r = .05$, ns; $r = .04$, ns, for ability and performance, respectively) to their post-evaluations ($r = .37$, $p < .006$; $r = .38$, $p < .005$, for ability and performance, respectively). Most importantly, as Fig. 2 shows, the relationship between self-ratings and objective performance was significantly stronger in post-evaluations. This is most directly revealed by the fact that differences in self-assessments (e.g., post-ratings minus pre-ratings) were at least marginally or significantly associated with actual performance as measured by the percent of words correctly found ($r = .25$, $p < .07$ for general ability; $r = .28$, $p < .04$ for specific performance). In this study the regression slopes show that overall these participants increased their self-evaluations. However, opposite of Study 1, participants at the top changed their evaluations more than those at the bottom.

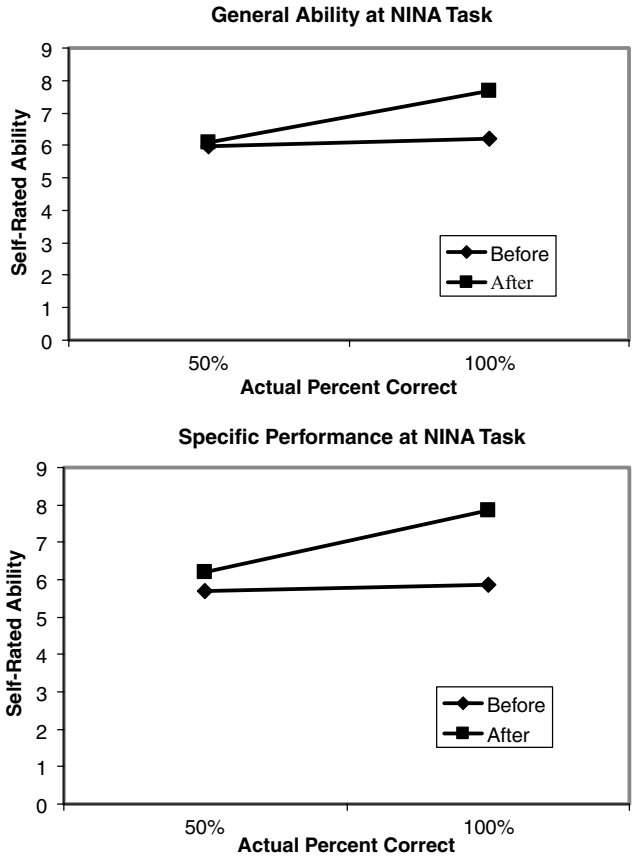

Fig. 2. The association between self-evaluations and actual performance before and after the total number of solutions were revealed (Study 2).

Summary

In sum, Study 2 successfully replicated Study 1 and revealed that people fail to take into account their errors

of omission when making self-evaluations. Participants showed little insight into the number of omission errors they had made and, thus, overestimated the number of possible solutions and underestimated their performance. In this study, participants performed better than they thought they had. We believe that the ease and simplicity of the task made participants feel that they were not doing well because they completed the task so quickly.

Once participants were confronted with their errors of omission they had no difficulty giving weight to this information when making further evaluations of their performance. This caused them to increase their self-evaluations, a result that seems appropriate given the high performance of participants with this task. It also prompted them to provide evaluations that more closely correlated with the metric they nominated as an appropriate measure of objective performance. Again, these findings and the findings from Study 1 suggest that it is not that people are uninterested in accurately self-evaluating, nor that they are disinterested in their errors of omission. People just do not have access to this information.

Study 3: Grammatical correction

In Study 3, we explored self-evaluations on another task involving errors of omission. It could be argued that the tasks used in both Studies 1 and 2 were somewhat unfamiliar to most participants. In Study 3 we wanted to use a task that participants would definitely have had experience with, finding grammatical errors in text. We took a report from the National Commission on Excellence and Education and inserted grammatical errors throughout the four pages of text. We were careful to only insert grammatical errors that would be marked as such on exams such as the NTE, GRE, or SAT. After identifying and correcting errors in the document, participants were asked a question about their general grammar ability and a question about their specific performance on the document we had presented to them. Then they were told how many total errors were contained in the document and asked again the same questions about ability and performance.

Our main focus was whether participants' initial self-evaluations would reveal any awareness of their errors of omission. Would their self-evaluations give appropriate, or any weight, to the number of words they had missed? We again explored whether their evaluations would change once they were actually informed of their errors of omission.

Method

Participants

Participants were 189 Cornell University undergraduates from a variety of courses in psychology and human development who earned extra credit for their participation.

Procedure

Participants were again run in groups, and all experimenters were blind to condition and the specific hypotheses of the study. Participants were given a packet with a four-page document titled: *An open letter to the American people, a nation at risk: The imperative for educational reform*, by the National Commission on Excellence in Education (Department of Education, 1983). They were given 30 min to find as many grammatical errors as possible in the document, and were asked to circle and correct the errors using a red pen. Participants were randomly assigned either to a low (15 errors, $n = 112$), a medium (30 errors, $n = 55$), or a high (60 errors, $n = 29$) condition. Four different versions of the low condition document were created, so that all the errors contained in the 60-error version would appear at least once in one of the low-error versions. Similarly, two different versions of the 30-error document were created.

Before beginning, all participants were given an instruction sheet that instructed them to "Focus on grammar, sentence structure, and punctuation; DO NOT consider the style of the writing (i.e., passive voice, word order). When considering punctuation, you should not be looking for preferred locations of punctuation but for the presence or absence of a punctuation mark that makes the sentence grammatically incorrect." In addition, all participants were given a sample paragraph that had four grammatical errors highlighted and corrected. These examples included the use of "as if" instead of "like," and conjugations such as "knowing" instead of "to know."

Following completion of the document, participants were given the same packet of questions as in Studies 1 and 2. They were asked to rate their overall ability at detecting grammatical errors as well as their performance on the document they had just confronted, both assessments used 9-point Likert scales. The questionnaire asked participants to estimate how many errors they had found and missed. It also asked them to rank order a number of criteria concerning how appropriate they were for assessing performance at the task.

After completing these questions, participants received a "final information packet" that contained the four pages of document with the errors highlighted, but the errors were not corrected. The total number of errors was written on the top of the first page. After looking over the information, participants answered exactly the same two questions on the first page of the packet, one question on overall ability at finding grammatical errors and a second question on performance on the document confronted. Participants were then debriefed and thanked.

Results and discussion

As in the first two studies, we performed multiple regression analyses in which we predicted self-evaluations of ability and performance from number of solutions found and number of solutions missed. This analysis presented us with a surprising pattern of findings. Self-evaluations reflected both the number of solutions found ($bs = .38$ and .37, $ps < .001$, for ability and performance, respectively) and the number missed ($bs = -.23$, $p < .07$ and $-.25$, $p < .04$, for ability and performance, respectively). Based on these findings, perhaps when the task is less novel participants are more aware of their errors of omission. However, that speculation would imply a significant correlation between the number of errors participants predicted they missed and the number they actually missed. No such correlation existed, $r(184) = -.003$, ns.

Further analyses indicated that participants' knowledge of their errors of omission was far from perfect. When we informed participants about the total number of errors in the document, and thus notified them of their errors of omission, they gave those errors greater weight in their self-evaluations. In a multiple regression analysis that included number of solutions found and number missed as independent variables, we found that participants gave significant weight to their errors of omission ($bs = -.31$ and $-.48$, $ps < .001$, for ability and performance, respectively), as well as to the number of solutions they had found ($bs = .54$ and .64, $ps < .001$, for ability and performance, respectively). Furthermore, number of omission errors predicted changes in self-evaluations between initial and final self-ratings ($bs = -.23$, $p < .07$ and $-.25$, $p < .04$, for ability and performance, respectively).

The greater weight given to omission errors had two significant influences on self-evaluations. First, informing participants of their errors of omission caused them to significantly lower their self-evaluations. The average rating of ability fell from 5.7 to 4.8, $t(188) = -7.61$, $p < .001$, and the average performance rating fell from 4.6 to 3.6, $t(188) = -7.74$, $p < .001$. This result follows from the fact that participants tended to underestimate how many omission errors they had made, thinking that they had made 12.7 errors when, in fact, they had made 18.1, $t(170) = -4.01$, $p < .001$.

Second, participants' self-ratings became more accurate, given the objective criterion that they endorsed as the most appropriate for judging their performance. The majority of participants (70%) thought that the percent correct (e.g., total words found divided by total words possible) was the most appropriate measure of performance for this task, and 86% of participants ranked actual percent correct as one of their top two choices for determining performance. The association between participants' objective performance using this measure and

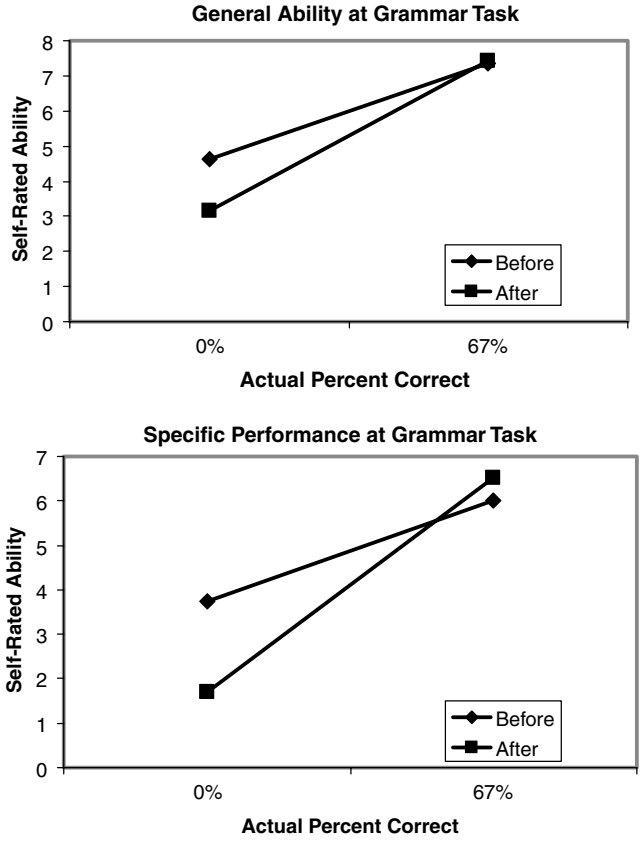

Fig. 3. The association between self-evaluations and actual performance before and after the total number of solutions were revealed (Study 3).

their self-evaluations increased from their pre-evaluations ($r = .33$, $p < .001$; $r = .30$, $p < .001$) to their post-evaluations ($r = .45$, $p < .001$; $r = .55$, $p < .001$). Echoing the previous studies, differences between post- and pre-evaluations were significantly associated with objective performance ($r = .22$, $p < .01$ for general ability; $r = .30$, $p < .01$ for specific performance) (see Fig. 3). Similar to the findings in Study 1, in general participants decreased their self-evaluations but participants who performed poorly changed more than those who had performed well on the task.

Summary

In sum, Study 3 successfully replicated most of the findings from Studies 1 and 2 using a more familiar task and revealed that people fail to fully take into account their errors of omission when making self-evaluations. Encouragingly, participants in this study showed a modicum of insight about their errors of omission, although that insight was inconsistent and imperfect. On the plus side, their initial self-evaluations were related to the number of errors they had missed. But on the negative side, participants could not accurately estimate how many omission errors they had made. Furthermore,

when participants were confronted with their errors of omission they had no difficulty giving substantially more weight to the information when making further evaluations of their performance. This caused them to lower their self-evaluations. It also prompted them to provide evaluations that more closely correlated with the metric they nominated as an appropriate measure of objective performance.

Study 4: Research methods

In Study 4, we explored self-evaluations on a task that psychologists and researchers encounter every day—searching for methodological flaws in experimental designs. We created four descriptions of hypothetical experiments and inserted methodological errors in each experiment. We then recruited advanced graduate student participants via email to identify and describe the errors. After completing the task, participants were asked about their methodological knowledge and performance at the task. They returned their responses to the experimenter via email and then they received our list of methodological flaws embedded in the studies. They were asked again to judge their knowledge and performance.

Our continued focus was whether participants' initial self-evaluations would reveal any awareness of their errors of omission. Even after significant training on how to design experiments, would their self-evaluations give appropriate, or any weight, to the number of errors they had missed, or would those evaluations reflect only the number of errors they had found? Beyond this, we pursued exploring whether their evaluations would change once they were actually informed of their errors of omission. Would they give weight to these omission errors and change their self-evaluations in general, or would they just not consider those omissions?

Method

Participants

Participants were 44 graduate students from several universities across the United States and eastern Canada who received $15 (US) gift certificates to Amazon.com for their participation. Participants were contacted via email and asked to participate; they received all materials as email attachments.

Procedure

After indicating a willingness to take part, participants were given a Microsoft Word file containing a consent form, instructions, the studies, and a final questionnaire. The instructions said "Please set aside enough time in one sitting and read through each study and for each *list and describe as many methodological flaws* that would call into question the experimenters' conclusions. I have provided a bulleted response page following each experiment, so write within this document, save, and email back when you are finished." Participants were further admonished not to consult with anyone else while providing critiques of the four experiments.

Each of the four experimental scenarios the participants read provided information on the hypotheses, participants, procedures, results, and conclusions. For example, one study scenario tested the hypothesis that males are more creative than females—but contained such errors as confounding age with gender and asking participants to think of creative uses of objects that were primarily male-oriented (e.g., a hammer). Other studies contained such errors as including no control group, failing to account for regression to the mean, differential attrition between conditions, or selection effects. Half ($n = 22$) of the participants were in the high-error condition and read descriptions that contained 24 total errors across the four studies. The other half ($n = 22$) were in one of two low conditions and read descriptions that contained 13 errors. After reading the instructions participants read each study and described the errors they found. Once they finished all four studies, participants were asked to rate: their ability to spot methodological problems in studies, their performance on the studies in our packet, their ability to spot methodological problems in their own research, and their skill at spotting methodological problems in the research of others. Participants provided their responses on 9-point Likert scales similar to those used in the first three studies. Participants then saved their responses and emailed the file back to the experimenter. Unlike Studies 1–3 the final questionnaire in this study did not ask participants to estimate how many errors they had found and missed before giving them their errors of omission and asking for their final performance evaluations. In addition, participants were not asked to rank order the criteria concerning how appropriate they were for assessing performance at the task. These questions were removed in order to rule out any possible experimenter demand effects due to the salience of having to estimate their errors found and missed before receiving their omissions and being asked to re-evaluate their performance.

Once the first part of the experiment was returned, the experimenter emailed the participant the second part. The second part contained the study descriptions once again along with a list of actual errors that we had inserted into each study. Participants were instructed that the list served "as the answer sheet to the task you completed." They were to read the list over and respond to follow-up questions, which were the same four ability and performance questions as they previously completed. Participants then returned the second part via email, were sent a debriefing and thanked, and were emailed their gift certificates a week later.

Results and discussion

Our continuing question was whether participants would show any insight into their errors of omission. We predicted that even doctoral students in psychology would show little or no insight—that evaluations of performance would reflect the number of solutions participants found in the puzzles, but would not reflect the solutions they missed.

Because we were really interested in whether participants' self-evaluations would take into account their errors of omission and because the four self-evaluation questions are not independent of each other, we decided to assess whether the four self-evaluation questions we used could be summed into an overall composite self-evaluation measure. Principal component analyses of our four self-evaluation questions (separately for initial and final responses) revealed that all measures loaded onto a single factor save one, the item about spotting problems in one's own research. The other three items (ability to spot methodological problems in studies, performance on the studies in our packet, and skill at spotting methodological problems in the research of others) formed a reliable composite scale (both α's $> .82$). Adding the fourth item reduced reliability. Thus, it was analyzed separately.

To test our prediction that participants' self-evaluation would not initially give weight to their errors of omission, we predicted participants' composite self-evaluation from number of solutions found and number missed. The composite self-evaluation did not reflect the number of solutions found ($b = .09$, ns) nor the number missed ($b = -.13$, ns). Since we did not ask them how many participants thought they missed, as in Studies 1–3, we could not look at the correlation between their predicted missed and their actual missed.

Again, was this neglect of omission errors because participants did not possess any sense of errors of omission or because they did not consider them valuable indicators of performance? As in past studies, informing participants of their omission errors did cause them to incorporate them into their self-evaluations. After informing participants of all the methodological problems we had embedded in the materials, we found that participants gave weight to their errors of omission ($b = -.43$, $p < .003$), as well as the number of solutions they had found ($b = .30$, $p < .04$).

This greater weight given to errors of omission had two significant influences. First, informing participants of their errors of omission caused them to lower their self-evaluations. The average composite self-evaluation rating fell from 6.7 to 6.4, $t(41) = -2.05$, $p < .05$. Second, participants' self-ratings were more accurate when we again used percentage of problems found as the metric of objective performance, as in the three previous studies. There was significant variability in performance

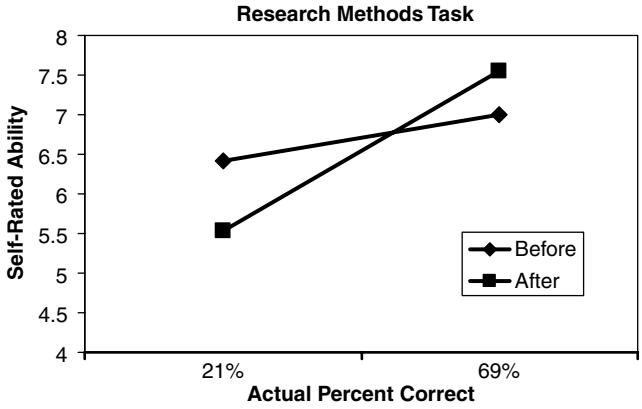

Fig. 4. The association between self-evaluations and actual performance before and after the total number of solutions were revealed (Study 4).

based on this criterion. Some participants found only 21% of the problems we had inserted whereas others found as many as 69%. Participants' composite self-rating correlated with this metric more tightly after being informed of their omission errors ($r(41) = .53$, $p < .001$) than before ($r(41) = .17$, ns), as depicted in Fig. 4. Differences in self-assessment from pre- to post-ratings were significantly associated with actual performance, $r(41) = .43$, $p < .005$. Fig. 4 shows that both participants who did well and those who did poorly adjusted their evaluations upon consideration of their errors of omission.

Perhaps not surprising, when participants were asked to make evaluations considering their capability at spotting methodological flaws in their own research—the item we found did not correlate with the rest—the results were rather inconsistent. We performed the same multiple regression analysis predicting their initial "own research" evaluation from number of solutions found and number of solutions missed and found that neither predicted their ratings ($b = -.18$, and $b = -.15$, ns; for found and missed, respectively). However, after informing participants about the total number of solutions included in each puzzle, they still did not give any weight to their errors in their self-evaluation ($b = -.07$, ns). Curiously, they did give weight to the number of solutions they had found to a marginal degree—but in a negative direction ($b = -.26$, $p < .09$). In addition, their average self-evaluation on the "own research" item changed negligibly from 6.3 to 6.5, $t(40) = 1.06$, ns.

Thus, participants easily separated their performance on the task presented to them from their performance on their own research studies. In the face of their errors of omission, participants decreased their self-evaluations of their ability in general, performance on the task, and ability with others' research, but maintained that they were still skilled at finding methodological flaws in their own personal research projects—a finding not surprising

to those who do research on the self and its elasticity against negative or inconsistent feedback (Kunda, 1990).

Summary

To summarize, Study 4 revealed that people fail to take into account their errors of omission when making self-evaluations, even concerning performances that they value and abilities that they have received training in. Doctoral students showed little insight about the entire set of the solutions that they could have generated in a task when determining how well they performed on that task. However, when they were confronted with their errors of omission they had no difficulty giving weight to the information when making further evaluations of their performance. This caused them to report more modest self-evaluations.

Taken together, these findings suggest once again that people are not disinterested in their errors of omission. Researchers do strive to find methodological flaws in their experiments before they started running participants. However, researchers may not have access to this information left to their own devices. When they do have access to all the relevant information, including errors of omission, their self-evaluations become more accurate.

Meta-analysis: Studies 1–4

The studies so far provided results that were generally supportive of our analysis concerning the role of omission errors in self-evaluations. However, not all critical statistical comparisons are consistent across all studies with our hypotheses. We did uncover a few surprises. Thus, to provide a sharper overall portrait of our results so far, we decided to conduct a meta-analysis to examine which of our predictions had received overall support, and at what level of significance. This analysis also allowed us to look at the effect sizes, to gauge, for example, how much weight people gave to solutions found versus solutions missed before being explicitly informed of their errors of omission. Looking at effect sizes also allowed us to see how much weight participants gave to omissions errors once they were informed of them in their self-evaluations. For purposes of this meta-analysis, we collapsed the ability and performance measures used in Studies 1–3 into one measure, and examined only the composite measure discussed in Study 4. Thus, for each study, we had one self-evaluation measure taken before we revealed errors of omission to participants, and one taken afterward.

We then turned out attention to how much weight participants gave across studies to solutions found and missed before they were informed of their omission errors. Participants' self-evaluations, as assessed via multiple regression analyses across studies, revealed that they gave significant weight to the number of solutions they had found (mean $r_{\text{partial}} = .30$, $Z = 5.84$, $p < .0001$), although there is a marginal hint of heterogeneity in this correlation across studies, $Q(3) = 6.91$, $p < .10$. Participants, however, did not give weight to the number of solutions they had missed across the studies (mean $r_{\text{partial}} = -.07$, $Z = -1.60$, ns), a result that was homogeneous across studies despite the appearance of one significant correlation in Study 3 (grammar), $Q(3) = 2.99$, ns. This lack of weight given to omission errors should not be surprising, given that participants did not successfully anticipate the number of omission errors they made in the three studies in which we asked them to (mean $r = .01$, $Z = .18$, ns).

Providing participants with information about omission errors had a consistent, and significant, impact on self-ratings, providing evidence that participants would incorporate these errors into self-evaluations if they only knew about them. In three studies, participants significantly lowered their self-estimates once the number of their omission errors was revealed to them (all ps $< .05$). In two of these studies, we discovered that participants had underestimated the number of their omission errors (this estimate was not collected in the third), and so the rationale for this shift in self-evaluation is straightforward. In the remaining study (Study 2; on visual search), alerting participants to their omission errors *raised* their self-evaluations, which was not a surprise, given that participants did so well on the task and had overestimated their omission errors.

Participants also gave significant weight to omission errors in their self-evaluations once they knew about them. Once notified of their omission errors, participants' self-ratings were significantly related to the number of those errors (mean $r_{\text{partial}} = -.43$, $Z = -7.87$, $p < .0001$), a finding that held true in each study. Indeed, the weight given to omission errors across studies equaled the weight given to the number of solutions found (mean $r_{\text{partial}} = .40$)! In addition, changes in self-evaluations from the initial to final ratings were significantly related to the number of omission errors that participants made (mean $r_{\text{partial}} = -.38$, $Z = -6.11$, $p < .0001$), a finding that held true in each study.

The net effect of this was that participants' final self-ratings were more highly correlated with the objective performance measure that participants described as appropriate-the percentage of solutions they had found relative to all the solutions possible (pre-evaluation mean $r = .23$, $Z = 4.59$, $p < .0001$; post-evaluation mean $r = .52$, $Z = 10.34$, $p < .0001$). Most importantly, these changes in self-assessments were significantly associated with actual performance (mean $r = .34$, $Z = 6.04$, $p < .0001$), indicating once again that participants were changing their self-evaluations in a more accurate direction.

One note before moving on, across all studies, it is clear that participants gave more weight to the number of solutions they had found after being told of their errors of omission, in that changes in self-evaluation also correlated with number of solutions found (mean $r_{partial} = .15$, $Z = 3.09$, $p < .005$), and this greater reliance on solutions found would also have produced a higher correlation between self-evaluations and objective performance. However, this relationship between self-evaluation changes and solutions found is more apparent than real. Self-evaluations appeared to show a tighter relationship with solutions found simply because more variance is explained elsewhere by omission errors.

The best way to show that the changes in self-ratings between initial and final questionnaires are due to the greater weight given to omission errors is to explore the zero-order correlations between (a) solutions found, (b) solutions missed, and (c) self-evaluation measures. These zero-order correlations reveal that, across studies, participants were not giving more weight to solutions found in their self-evaluations when we informed them about their omission errors (mean r increased only from .27 to .29). However, participants were giving quite a bit more weight to omission errors (mean r goes from $-.00$ to $-.32$) when they found out about those errors.

In sum, what seemed to be missing in participants' initial self-evaluations was awareness and knowledge about errors of omission. Providing participants with information about these errors caused them to significantly revise their self-evaluations in predictable ways, and also prompted them to provide more accurate evaluations when compared against an objective measure of performance.

Study 5: Implications for behavior

In Study 5, we examined potential implications for behavior. Participants again had the opportunity to play the popular board game Boggle, and then were asked to bet on the quality of their performance with money given to them by the experimenter. Roughly half the participants were given information about the total number of solutions present in the puzzle prior to betting and the other half were not. Our first question was whether participants not informed of their errors of omission would give these errors any weight in their betting decisions. Our second question was whether providing that information would influence betting decisions in an appropriate way. We expected participants informed of their errors of omission to bet less and pay more attention to the number of words they had missed.

Method

Participants

Participants were 94 Cornell University undergraduates from a variety of courses in psychology and human development who earned extra credit for their participation.

Procedure

The procedure of Study 5 followed closely that of Study 1. Again, participants were run in groups and all experimenters were blind to condition and the specific hypotheses of the study. Participants were given a packet with three different Boggle puzzles. Participant groups were randomly assigned to either a low (100 possible word, $n = 31$), a medium (179 possible words, $n = 29$), or a high (254 possible words, $n = 34$) condition, using the puzzle sets from Study 1. Following completion of the three puzzles, participants were given two questions answered on 9-point Likert scales, taken from Study 1, that asked them to rate their Boggle-playing ability and their specific performance on the three puzzles they confronted.

Participants were then told that they would have the opportunity to bet on how well they did on the *last* of the three puzzles they completed (i.e., the competition puzzle). All participants were handed $2 in quarters and told that they were going to be paired up with one other person in their group, after which they would place a bet on whether they found more words in the last puzzle then their partner. If they won the competition the experimenter would match what they bet, if they lost the competition they would have to give the experimenter the amount they bet.

Participants were then randomly put into one of two conditions. In the "answers" condition ($n = 47$), the packets had the cover sheet with the instructions, then the answer sheet, and then a place where they could indicate their bet. The answer sheet told participants the total number of possible words for each boggle puzzle, highlighting the third puzzle; it did not list all the words that could be found in the puzzles but merely gave a total for each puzzle. In the "no answers" condition ($n = 47$), information was not provided about the number of words in each puzzle. Participants were not aware that the packets were different for other people in the same room, their seats were spread out and packets were randomly distributed.

After participants placed their bets the packets were collected, participants were paired up randomly by drawing a piece of paper out of a bag that included pairs of letters, the competition puzzles were scored, and bets were called (i.e., money was collected and distributed to winners and losers). Participants were then debriefed and thanked.

Results and discussion

Our first question was whether participants would show any insight into their errors of omission. As expected, in multiple regression analyses, self-evaluations reflected the number of solutions found in both conditions ($bs = .50$, $p < .001$ and .38, $p < .011$, for ability and performance of answers group; $bs = .44$, $p < .002$ and .18, ns, for ability and performance of no answers group). Self-evaluations for both groups did not reflect the number of solutions missed ($bs = -.16$ and $-.18$, ns, for ability and performance of answers group; $bs = -.17$ and $-.20$, ns, for ability and performance of no answers group).

In all the studies until now we have found that it is not that participants do not consider errors of omission a valuable source of evaluation information, it is just that they do not have access to them. Thus, when we informed half the participants about the total number of solutions included in each puzzle, we predicted that they would give those errors weight in their bets on the competition puzzle. And, as predicted, informing participants of their errors of omission caused participants in the answers condition to bet significantly less. The average bet was \$.88 in the answers condition and \$1.13 in the no answers condition $t(92) = -2.10$, $p < .04$. However, in a multiple regression analysis that included number of solutions found and number missed as independent variables, looking separately at answers and no answers conditions, we found that neither condition gave significant weight to their errors of omission ($bs = -.04$ and $-.83$, ns, for answers and no answers conditions, respectively), but both gave either marginal or significant weight to number of solutions found ($bs = .26$, $p < .08$ and .38, $p < .007$, for answers and no answers, respectively), in their bets.

Summary

To summarize, Study 5 provided further evidence that people fail to take into account their errors of omission when making self-evaluations. When rating their ability and performance, participants gave weight to the number of solutions they had found but not to the number they had missed. Furthermore, informing participants of their omission errors produced behavioral consequences. In a betting situation, participants bet less when they knew the number of total solutions present in the puzzles than when they did not know. In truth, the size of their bets were not correlated with the number of their omission errors, but being informed of those errors caused participants to bet more conservatively. One additional aspect of this study is worth mentioning. In more consequential situations, such as in our betting task, it could be predicted that people would more actively seek out information on their errors of omission. However, not one participant in the no-answers condition expressed concern over betting money before having access to the number of solutions to the puzzle.

General discussion

In mid-2002, soon after the tragic events of September 11, 2001, Secretary of Defense Donald Rumsfeld was asked about the state of the U.S. government's war on terrorism. As he discussed the range of threats against the U.S., he remarked that:

> There are things that we know that we know. There are [also] known unknowns—that is to say, there are things that we now know we don't know. But there are also unknown unknowns. There are things we do not know we don't know. So when we do the best we can and we pull all this information together, and we then say well that's basically what we see as the situation, that is really only the known knowns and the known unknowns (Kamen, 2002, p. A25).

In his remarks, Donald Rumsfeld implied that it was the "unknown unknowns" that presented the greatest danger to the U.S. and its allies.

The unknown unknowns of everyday life

In this paper, we proposed that similar "unknown unknowns" present an intrinsic difficulty for people as they assess their skills and talents in everyday life. People may be aware of solutions they have found to problems, and might at times have some dim awareness of the solutions they did not identify, but by definition they are not generally aware of their gaps in knowledge, what they "don't know," their unknown unknowns, or their errors of omission. We argued that peoples' evaluations of their performance and ability would largely reflect the number of solutions they had found on a task and would largely fail to reflect the number of solutions that they missed, that is, their errors of omission. In fact, across the five studies, participants paid attention to the number of solutions they had found to the puzzles we presented them as they provided initial evaluations of their ability and performance. This reliance on solutions found was significant in three studies (Studies 1, 3, and 5), as well as highly significant when averaged across all studies. The impact of omission errors on self-ratings was much less evident, if evident at all. In the three studies in which we asked, participants' estimates of their omission errors failed to correlate with the reality. And in only one study out of five (Study 3) did their self-ratings significantly correlate with their errors of omission. Collapsing across all five studies, we did not

find a significant tendency for participants to give weight to their errors of omission when left to their own devices.

Do people give omission errors so little weight because they view them to be unimportant to their self-evaluations? In four studies people continually demonstrated that they consider errors of omission important by giving weight to these errors after being made aware of them. After receiving information about omission errors, participants gave those errors significant weight in their self-evaluations in the four studies in which we assessed the relationship. Perhaps more important, informing participants of their errors of omission reduced how positively they viewed their skills and abilities, except in the one study in which participants on average performed almost perfectly (Study 2), in which case they raised their performance evaluations. Also, informing participants of their omission errors prompted their self-evaluations to be more accurate, relative to objective performance criteria that they themselves cited as appropriate. In addition, this research also sought to demonstrate that a lack of awareness about errors of omission would have important implications for behavior. Study 5 revealed that participants who were informed of their errors of omission bet significantly less money on their performance than those who did not have access to their omissions.

Taken together, these studies suggest why well-intentioned people, motivated solely by the pursuit of an accurate self-image can still come to erroneous conclusions of their skill and talent. When judging themselves, people often do not have all the information they need to render a veridical judgment. They may know all the ways that they approach a problem to solve it, but they, of course, are not aware of alternative solutions (and sometimes better ones). Not cognizant of these other solutions, they may be left with an erroneous view of their skills even when they diligently consider all the evidence they have in front of them.

In this paper, we considered how this process played out in some laboratory puzzles (e.g., word and visual search games), but also in some other domains of potential real world consequence (e.g., grammar, the ability to spot methodological problems among those doing research). We doubt that the problem of unknown unknowns is constrained to the tasks we presented participants but extends to a host of real life tasks of significant consequence—tasks like looking for fire hazards in a new home, or child proofing a house before a baby starts crawling. The fundamental difficulty with these tasks and tasks like them is that although people are quite aware of the hazards that were anticipated, they are by definition completely unaware of the hazards that they could have identified but failed to. These types of errors of omission can be most costly.

Ill-defined problems

All of the studies in this paper showed how the inability to recognize errors of omission led to imperfect performance evaluations, but these studies also may have underestimated the magnitude and pervasiveness of this problem in real life. In real life, the problem is not that people are not aware that they make errors of omission, it is that often such errors are impossible to define and enumerate.

Errors of omission can be impossible to enumerate because of the important distinction between "well-defined" and "ill-defined" problems. Well-defined problems are like the Boggle puzzles we gave to participants. With each puzzle, it is possible to define all the possible solutions, as well as develop systematic ways to generate them all. Indeed, computers can be programmed to find all possible solutions. Ill-defined problems, however, are much different, in that they do not have a finite set of solutions or even a definitive or universally accepted procedure for reaching them (Newell, 1969; Reitman, 1965; Simon, 1973). For example, it is impossible to catalogue all the solutions to such ill-defined tasks as designing an architecturally significant building, composing a classic country and western song, or writing the poem that resurrects the sonnet, nor can a computer be outfitted with algorithms to produce them. There are strategies people can try to complete each task, but there is no one perfect approach for any of these problems, nor any absolute way to know when one has constructed the best of all possible solutions. Further confusing issues, the boundary between well-structured and ill-structured problem solving can be a vague and fluid one (Simon, 1973). Therefore, with ill-structured tasks, one makes errors of omission, but the set of omission errors is not well-defined. As a consequence, people are never really in a position to know just how well they have done because it is difficult to know all the alternative solutions they could have arrived at.

Familiarity

However, a reader could disagree with our observations about the difficulties that omission errors create for self-evaluation. Perhaps at completely novel tasks—Boggle task, for example—people may not be aware of their errors of omission. Maybe if our tasks were more familiar, and our participants had more experience with them, they would have had a more accurate sense of all the omission errors they were making. Perhaps this is true: Participants' initial self-evaluations along grammar skills, a more familiar task, were correlated with errors of omission. Nonetheless, other aspects of the data contradict this assertion, participants did not accurately estimate how many omission errors they made in the

grammar task. In addition, in Study 4, evaluations along another familiar task (looking for methodological problems) were not related to omission errors. But still, should not people be better able to identify their errors of omission with more familiar tasks? At its heart, this is an empirical question, and future research could examine the extent to which experience and familiarity with a task informs people of their omission errors.

However, even if future research reveals that people are aware of omission errors more with familiar tasks than unfamiliar ones, we think the import of this problem would not be diminished, for two reasons. First, the most important actions we take in life we often take only once or a few times, thus making those tasks unfamiliar. Examples of "unfamiliar," yet important, tasks would include buying a house, getting married, raising children, starting a new career, and negotiating the salary and perks for a job. None of these tasks come with an instruction manual, and so we can make costly errors of omission that we remain unaware of.

Second, in some sense, all tasks in life at one point are unfamiliar. People go out on the first date of their life, try to sell their first used car, or give a good public speech. To the extent that people are not aware of omission errors when they are novices at these tasks, they may not realize how much they can—and perhaps should—improve. As a result, they may remain novices for a much longer period than they might otherwise. Being unable to identify one's errors of omission would surely delay or prevent people from becoming experts.

However, given the importance of omission errors for self-evaluation, it would be potentially profitable to examine when people might have some awareness of their omission errors, as well as the psychology that allows them to gain that awareness.

Lack of feedback

Although the focus of this paper is on why people are poor at assessing their abilities and performances, we do not mean to imply that people never become better self-evaluators. If lack of awareness concerning omission errors leads to poor self-evaluation, then feedback from others about one's omissions should improve assessments of ability and performance. There is research outside of social psychology that suggests that this might be the case.

In organizational settings, *benchmarking* is a method through which organizations establish their position in an industry and learn from competitors with an aim toward achieving greater success in all aspects of their business. In benchmarking, organizations observe the behaviors of similar organizations to learn how those organizations solve typical problems (Campbell, 2002). The organizations goal "is not to find a model to be copied, but rather to stimulate reflection on a particular model and practices through analysis of models developed" (Boisvert, 2001, p. 33).

At a more individual level, benchmarking has been a tool used to inform students about their competencies—and the lack thereof. In medical education, for example, students watch how other students handle the same situation that they have recently been exposed to, such as interviewing a patient in a difficult situation. Researchers have found that putting medical students through benchmarking exercises improves the accuracy of their self-evaluations (Martin, Regehr, Hodges, & Mcnaughton, 1998), although follow-up research suggests that benchmarking helps the self-assessments of competent students more than it does poor performers (Hodges et al., 2001). In a similar vein, organizational researchers have discovered that letting employees observe the performances of their co-workers prompts them to provide self-evaluations that better match the evaluations provided by their supervisors (Farh & Dobbins, 1989). In the context of our analysis of performance estimates, we can presume that benchmarking might improve the accuracy of self-evaluations because it is a means of identifying one's own errors of omission.

In addition, one potentially crucial way one can learn about one's errors of omissions is to have a mentor. Mentoring is often thought of as an adult and youth relationship but it can also play an important role in adult social life, with a different emphasis. A peer mentor can be someone with whom it is okay to ask naïve questions, make uninformed suggestions, or offer incomplete ideas and plans. This mentor can point out mistakes without judging and offer advice and guidance in solving difficult problems or assessing complex situations (Green, 2002). This guidance could easily include identifying the mentee's errors of omission in order to increase their performance.

There is much research showing that mentoring increases skill in a number of domains, such as academic performance (Gordon, 2000; Johnson, 1997; Linnehan, 2001; Tekian, Jalovecky, & Hruska, 2001), professional performance (Boakes, Gardner, Yuen, & Doyle, 2000; Halai, 1998; Rotondo & Perrewe, 2000; Weinstein, 1999), work place socialization (Allen, McManus, & Russell, 1999; Weinstein, 1999), development of leadership skills (Murray, 1991), stabilization of problem behaviors (Grossman & Tierney, 1998; Moffatt, 2000), and meta-cognitive awareness (Asquith, 1996; Osana, 1998). Although there was little work explaining the specific processes involved in the success of mentoring, mentors can be in a unique position to fill someone in on the errors of omission they are making.

Concluding remarks

In sum, this paper is an exploration into why people make imperfect self-assessments of their performances

and abilities. We assert that people in many areas of life cannot be expected to reach accurate views because they face significant informational barriers. We proposed and showed across five studies that people largely possessed little or no magical insight into their errors of omission, even when they admitted that such errors were relevant to their performance, and were motivated to factor such errors into their performance evaluations once they knew about them.

It seems an obvious ending, but we cannot resist the temptation to apply our analysis to our own efforts in this paper. Although, in the conduct of this research, we attempted to do all that we could do to provide empirical support for our analysis of imperfect self-evaluations, there is bound to be some efforts we could have made but could not identify. We concede at this point that we do not know of them.

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.

Allen, T. D., McManus, S. E., & Russell, J. E. A. (1999). Newcomer socialization and stress: Formal peer relationships as a source of support. *Journal of Vocational Behavior, 54*(3), 453–470.

Asquith, P. S. (1996). Relations between metacognition and instruction: Development of metacognitive awareness via instruction of reflective research skills in an academic context for middle school students. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 57*(4-A), 1479.

Boakes, J., Gardner, D., Yuen, K., & Doyle, S. (2000). General practitioner training in palliative care: An experiential approach. *Journal of Palliative Care, 16*(2), 11–19.

Boisvert, H. (2001). Using benchmarking to strategically deploy the finance function profiling. *CMA Management, May*, 30–33.

Campbell, C. (2002). Understanding your industry inside and out. *Pulp and Paper, 76*(12), 56.

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

Department of Education (1983). *An open letter to the American people. A nation at risk: The Imperative for Educational Reform: A Report to the Nation and the Secretary of Education*. United States Department of Education, April, 65.

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58*(4), 568–581.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*(3), 83–87.

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*(6), 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*(4), 521–532.

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

Falchikov, N., & Boud, D. (1989). Student self assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395–430.

Farh, J., & Dobbins, G. H. (1989). Effects of comparative performance information on the accuracy of self-ratings and agreement between self- and supervisor ratings. *Journal of Applied Psychology, 74*, 606–610.

Fatsis, S. (2002). *Word freak: Heartbreak, triumph, genius, and obsession in the world of competitive scrabble players*. New York: Penguin Books.

Gordon, D. M. (2000). Mentoring urban Black students: Implications for academic achievement, ethnic/racial identity development, racial socialization, and academic disidentification. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61*(3-B), 1634.

Green, T. (2002). The advantages of mentoring. *Florida Alliance of Virtual Assistants, May*, Retrieved on August 28, 2003 from http://www.flavas.org/articles.

Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters Program. *Evaluation Review, 22*(3), 403–426.

Halai, A. (1998). Mentor, mentee and mathematics: A story of professional development. *Journal of Mathematics Teacher Education, 1*(3), 295–315.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement performance measures. *Review of Behavioral Research, 52*, 123–142.

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

Haun, D. E., Zeringue, A., Leach, A., & Foley, A. (2000). Assessing the competence of specimen-processing personnel. *Laboratory Medicines, 31*, 633–637.

Hirschfeld, A. (1970). *The world of Hirschfeld*. Hillsdale, NJ: Erlbaum.

Hodges, B., Regehr, G., & Martin, D. (2001). Difficulties in recognizing one's own incompetence: Novice physicians who are unskilled and unaware of it. *Academic Medicine, 76*, S87–S89.

Johnson, A. W. (1997). Mentoring at-risk youth: A research review and evaluation of the impacts of the Sponsor-A-Scholar Program on student performance. *Dissertation Abstracts International: Section A: Humanities and Social Sciences, 58*(3-A), 0813.

Kamen, A. (2002). Tent rent: $10 a day. *Washington Post*, A25.

Kruger, J., & 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*(6), 1121–1134.

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

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Decision Processes, 20*(2), 159–183.

Linnehan, F. (2001). The relation of a work-based mentoring program to the academic performance and behavior of African American students. *Journal of Vocational Behavior, 59*(3), 310–325.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Personality and Social Psychology, 67*(3), 280–296.

Martin, D., Regehr, G., Hodges, B., & Mcnaughton, N. (1998). Using videotaped benchmarks to improve the self-assessment ability of family practice residents. *Academic Medicine, 73*, 1201–1206.

Moffatt, M. J. (2000). The affective bond between preschool-aged children and mentors. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 60*(8-B), 4238.

Murray, M. (1991). *Beyond the myths and magic of mentoring: How to facilitate an effective mentoring program*. San Francisco, CA: Jossey-Bass/Pfeiffer.

Newell, A. (1969). Heuristic programming: Ill structured problems. In J. Aronofsy (Ed.), *Progress in Operations Research* (Vol. 3, pp. 360–414). New York: Wiley.

Osana, H. P. (1998). Designing effective instruction toward the development of correlational reasoning in middle school. *Dissertation Abstracts International: Section A: Humanities and Social Sciences, 59*(5-A), 1459.

Reitman, W. R. (1965). *Heuristic decision procedures, open constraints, and the structure of ill-defined problems. Cognition and Thought*. New York: Wiley.

Rotondo, D. M., & Perrewe, P. L. (2000). Coping with a career plateau: An empirical examination of what works and what doesn't. *Journal of Applied Social Psychology, 30*(12), 2622–2646.

Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence, 4*, 181–201.

Tekian, A., Jalovecky, M., & Hruska, L. (2001). The impact of mentoring and advising at-risk underrepresented minority students on medical school performance. *Academic Medicine, 76*(12), 1264.

Weinstein, M. J. (1999). What makes a good mentor. The role of mentor flexibility in mentor functions and protégé outcomes. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 59*(12-B), 6508.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*, 806–820.

Wheelmen, T. (2003). *2000 uses for WD-40*. Retrieved on April 15, 2003 from http://www.twbc.org/wd40.htm.