Journal of Experimental Psychology: Learning, Memory, and Cognition

Unskilled but Aware: Reinterpreting Overconfidence in Low-Performing Students

Tyler M. Miller, and Lisa Geraci

Online First Publication, January 24, 2011. doi: 10.1037/a0021802

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/11/$12.00 DOI: 10.1037/a0021802

Unskilled but Aware: Reinterpreting Overconfidence in Low-Performing Students

Tyler M. Miller and Lisa Geraci
Texas A&M University

People are generally overconfident in their self-assessments and this overconfidence effect is greatest for people of poorer abilities. For example, poor students predict that they will perform much better on exams than they do. One explanation for this result is that poor performers in general are doubly cursed: They lack knowledge of the material, and they lack awareness of the knowledge that they do and do not possess. The current studies examined whether poor performers in the classroom are truly unaware of their deficits by examining the relationship between students' exam predictions and their confidence in these predictions. Relative to high-performing students, the poorer students showed a greater overconfidence effect (i.e., their predictions were greater than their performance), but they also reported lower confidence in these predictions. Together, these results suggest that poor students are indeed unskilled but that they may have some awareness of their lack of metacognitive knowledge.

Keywords: metacognition, overconfidence, prediction

When people make self-assessment errors, and they often do, they are usually in the direction of overconfidence. For example, people overestimate their reasoning ability, their ability to recognize humor, and their knowledge of grammar (Kruger & Dunning, 1999), and they underestimate the time they need to complete tasks (Buehler, Griffin, & Ross, 1994). In addition, most people believe they have better than average leadership skills (Dunning, Heath, & Suls, 2004), driving skills (Knouse, Bagwell, Barkley, & Murphy, 2005; A. F. Williams, 2003), and dating popularity (Preuss & Alicke, 2009). Similarly, people are overconfident in their academic abilities. In the classroom, undergraduate students tend to overestimate their performance on upcoming exams (cf. Hacker, Bol, Horgan, & Rakow, 2000; Miller & Geraci, 2010). For example, in Hacker et al. (2000), many students predicted that they would earn scores more than 30% higher than their actual scores.

This overconfidence effect is greatest for people who score below average compared with those who attain above average scores (e.g., Bol, Hacker, O'Shea, & Allen, 2005; Burson, Larrick, & Klayman, 2006; Hacker et al., 2000; Kelemen, Winningham, & Weaver, 2007; Krueger & Mueller, 2002; Kruger & Dunning, 1999; Miller & Geraci, 2010; Nietfeld, Cao, & Osborne, 2005). When undergraduate students are asked to predict their exam scores, students with the higher scores have more accurate predictions than do students with lower scores (Hacker et al., 2000). Similarly, in the laboratory, students with high SAT scores are more accurate than students with low SAT scores when asked about their mastery of a set of Swahili–English word pairs (Kelemen et al., 2007).

Several studies have confirmed that low performers are more overconfident than high performers, but the reason for this greater metacognitive inaccuracy is debated. Some have suggested that the exaggerated overconfidence effect in low performers is the result of a measurement artifact whereby low performers have room to make predictions that are much higher than their level of performance, whereas high performers do not (Krueger & Mueller, 2002). But perhaps the leading interpretation is that low performers are overconfident because they have a general deficit of metacognitive insight. The double-curse account suggests that in addition to lacking knowledge of the material, poor students also lack awareness of the knowledge that they do and do not possess. According to recent characterizations, low-performing students are "blissfully incompetent" (W. M. Williams, 2004) and "unskilled" and "unaware" (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008). Evidence for this characterization of low performers as unskilled and unaware comes from studies showing that low performers in particular predict that they will perform much better than they actually do (the overconfidence effect) even in the face of counterinformation (Kruger & Dunning, 1999).

It follows from the double-curse account that if low performers are blissfully incompetent, then, in addition to making inaccurate performance predictions, they would also be unduly confident in these predictions. Indeed, Dunning (2005) likened low performers' inflated self-assessments to brain damage (i.e., anosognosia) and suggested that "people performing poorly cannot be expected to recognize their ineptitude" and that "the ability to recognize the depth of their inadequacies is beyond them" (p. 15).

The current study examines whether this characterization of low performers is correct. Are low performers entirely unaware of their deficits? To answer this question, we asked low- and high-performing students to rate their confidence in their prior grade predictions. If low performers are unaware of their metacognitive deficits, then they should be at least as confident in their performance predictions as high performers are in theirs. However, if

Tyler M. Miller and Lisa Geraci, Department of Psychology, Texas A&M University.

Correspondence concerning this article should be addressed to Tyler M. Miller, Department of Psychology, Texas A&M University, College Station, TX 77843. E-mail: milltyl@tamu.edu

low performers have some awareness of their errors in judgment, then they should be less confident than high performers in their performance predictions.

To examine the nature of the overconfidence effect in low and high performers, we highlight the distinction between two forms of confidence. Herein, we refer to errors of overestimating one's ability (predicting that one will perform better than one does) as *functional overconfidence* and errors of overcertainty (being overly certain of one's predictions) as *subjective overconfidence*. As far as we are aware, only one study has examined subjective confidence associated with predictions of performance (Dunlosky, Serra, Matvey, & Rawson, 2005). In this study, participants made judgments of learning (JOLs) to indicate the likelihood that they would remember unrelated noun pairs. For each JOL, participants made a second-order judgment (SOJ) indicating their confidence in the JOL. Results showed that JOLs and SOJs were functionally distinct from each other, displaying a U-shaped curvilinear relationship with higher SOJs at extreme JOLs. In addition, the curve was asymmetrical, showing that SOJs associated with high JOLs were much greater than SOJs associated with low JOLs.

In the current study, we used a methodology similar to that used by Dunlosky et al. (2005) to examine the nature of the overconfidence effect in low-performing students. College students in an upper level cognitive psychology course were asked to predict their exam scores and rate their confidence in their predictions. In Study 1, students made letter grade predictions for the first exam immediately before the exam, and in Study 2, they made percentage predictions immediately before the first and final exams. In both studies, students rated their confidence in their predictions. We predicted that low-performing students would show a greater functional overconfidence effect than would high-performing students in that low-performing students would predict that they would receive disproportionately higher grades than they did. Of interest was whether the low-performing students would also be subjectively overconfident. According to the double-curse explanation, subjective confidence should be just as high, if not higher, for the low-performing students compared with the high-performing students because the low-performing students are assumed to lack awareness of their metacognitive deficit. If low-performing students are less confident than high-performing students in their grade predictions, then this would suggest that low-performing students may not be unaware of their metacognitive difficulties, as previously suggested.

Study 1

Method

Participants. Ninety-one students from a cognitive psychology course at Texas A&M University participated in the study. The students were largely junior and senior psychology majors. Seventy-four percent of psychology majors at Texas A&M University are female, 72% are European American, 5% are African American, 17% are Latino or Hispanic American, 5% are Asian American, and 1% were in an "other" category.

Design and procedure. Immediately before the first exam, participants recorded a letter grade prediction on the exam cover sheet. These letter grades were converted to numeric values for the analyses using the standard grading scale (i.e., scores between 100% and 90% received an A, between 89% and 80% received a B, between 79% and 70% received a C, between 69% and 60% received a D; and 59% or less received an F. The same grading scale was used for both studies). For example, if a student predicted a B+, that score was converted to a numeric value of 88% on the basis of the midpoint of the B+ range. Students also rated their confidence that their exam score prediction was correct using a Likert-type scale ranging from 1 = *low confidence* to 5 = *high confidence*. Students were given an incentive to be as accurate as possible in these predictions: If their exam score prediction was within the same letter grade equivalent as their actual performance, they received an additional 2 percentage points on their exam score.

Results

We divided students into quartiles on the basis of their exam performance. We created difference scores for each student by subtracting the actual score from the predicted score. As such, positive numbers indicated overconfidence and negative numbers indicated underconfidence (see Table 1). As expected, the omnibus F test indicated significant differences in calibration by quartile, $F(3, 86) = 25.23$, $MSE = 30.92$, $p < .001$, $\eta_p^2 = .47$. Tukey post hoc tests indicated that students in the top two quartiles, Quartiles 3 and 4, were significantly more calibrated (less overconfident) than were students in the bottom two quartiles, Quartiles 1 and 2 (ps $< .05$). Students in Quartile 2 were significantly more calibrated than were students in Quartile 1 ($p < .05$). Thus, results replicated previous findings showing the largest errors in calibration, with guesses being higher than performance, in poor-performing students.

When asked to rate their confidence in their exam prediction, however, low-performing students were less certain than high-performing students that their predictions were correct, $F(3, 86) =$

Table 1
Mean Prediction, Grade, Difference, and Confidence Scores for Studies 1 and 2 by Quartile

Quartile	Prediction	Grade	Difference	Confidence
		Study 1 Exam 1		
1	79.60 (1.31)	69.09 (0.88)	10.50 (1.45)	3.23 (0.11)
2	83.57 (1.27)	78.26 (0.54)	5.30 (1.06)	3.65 (0.18)
3	85.78 (1.06)	84.87 (0.41)	0.91 (1.00)	3.74 (0.11)
4	87.36 (0.86)	90.73 (0.50)	−3.36 (1.14)	3.77 (0.11)
		Study 2 Exam 1		
1	79.10 (1.44)	66.83 (1.33)	12.28 (1.98)	3.24 (0.16)
2	84.45 (0.95)	80.14 (0.71)	4.31 (0.97)	3.55 (0.13)
3	86.83 (0.93)	88.62 (0.27)	−1.79 (0.94)	3.66 (0.13)
4	89.09 (0.80)	96.00 (0.58)	−6.91 (1.08)	3.75 (0.12)
		Study 2 Final Exam		
1	78.10 (1.74)	60.38 (0.90)	17.71 (2.12)	64.76 (4.48)
2	79.55 (1.72)	69.30 (0.47)	10.25 (1.71)	72.70 (4.51)
3	79.64 (1.59)	76.00 (0.46)	3.64 (1.51)	80.95 (2.17)
4	84.57 (1.58)	85.62 (0.84)	−1.05 (1.51)	72.38 (3.96)

Note. Students in Studies 1 and 2 reported confidence on a scale of 1–5 for Exam 1 and students in Study 2 reported confidence on a scale of 0–100 for the final exam. Standard errors are shown in parentheses.

3.42, $MSE = 0.41$, $p = .012$, $\eta_p^2 = .11$. Tukey post hoc tests indicated that students in the top two quartiles, Quartiles 3 and 4, were significantly more confident in their predictions than were the students in the bottom quartile ($p < .05$). Thus, the results demonstrate a dissociation between functional and subjective confidence, showing that although low-performing students may show a greater functional overconfidence effect than do high-performing students, they are subjectively less confident in these judgments compared with high-performing students (see Figure 1).

In Study 1, students predicted their grades by providing a letter grade. In Study 2, students were asked to make more fine-grained predictions. They predicted the specific number grade they believed they would receive on the exam. For example, students were allowed to predict they would receive an 86 on the exam rather simply indicating they would receive a B. This change allowed us to improve measurement precision by allowing students to make more specific grade predictions, and it also removed the need to transform students' guesses into average number grades. Thus, the purpose of Study 2 was to replicate the findings from Study 1 using number predictions as opposed to letter grade predictions. Additionally, in Study 2, there were no incentives for accuracy to mimic standard classroom conditions. Would low-performing students continue to be less confident than high-performing students in their exam predictions?

Study 2 was also designed to examine if subjective or functional overconfidence might change over time and with course experience. In addition to the first exam, students were asked to make grade predictions and to rate their confidence in these predictions on a cumulative course final. Students also rated their confidence in their exam grade prediction a second time, after they completed the final exam. Previous research suggests that predictions are more accurate when they are made after the material is retrieved (Pierce & Smith, 2001) or after a delay when retrieval processes are engaged (Nelson & Dunlosky, 1991). One might predict that low performers would show a greater reduction in subjective confidence than high performers because low performers would use the exam information to adjust (lower) their subjective confidence.

Study 2

Method

Participants. One hundred thirteen students from a cognitive psychology course at Texas A&M University participated in the study. Note that the students in Study 2 were from a different cognitive psychology course than the students in Study 1. The students were largely junior and senior psychology majors. Seventy-four percent of psychology majors at Texas A&M University are female, 72% are European American, 5% are African American, 17% are Latino or Hispanic American, 5% are Asian American, and 1% were in an "other" category.

Students made grade predictions and confidence ratings on both the first exam and the final exam. After the students completed the final exam, they again rated their confidence in their original exam prediction. Only a subset of these 113 students took the final exam because of a course policy that allowed students to opt out of taking the final if they had high scores on the previous exams. This left 86 students that completed the final exam. Because the best students did not take the final exam, one might expect that the final exam data would not be representative of high-performing students. However, to anticipate, the data from the final replicate the results from Exam 1 of the current study and Exam 1 in Study 1.

Design and procedure. The design and procedure were similar to those of the previous study with the exception that participants in Study 2 were also asked to make an exam prediction and a confidence rating prior to the final exam and a second confidence rating after the final exam.

Results

Exam 1. As in Study 1, we divided students into quartiles on the basis of their exam performance Again, high-performing stu-

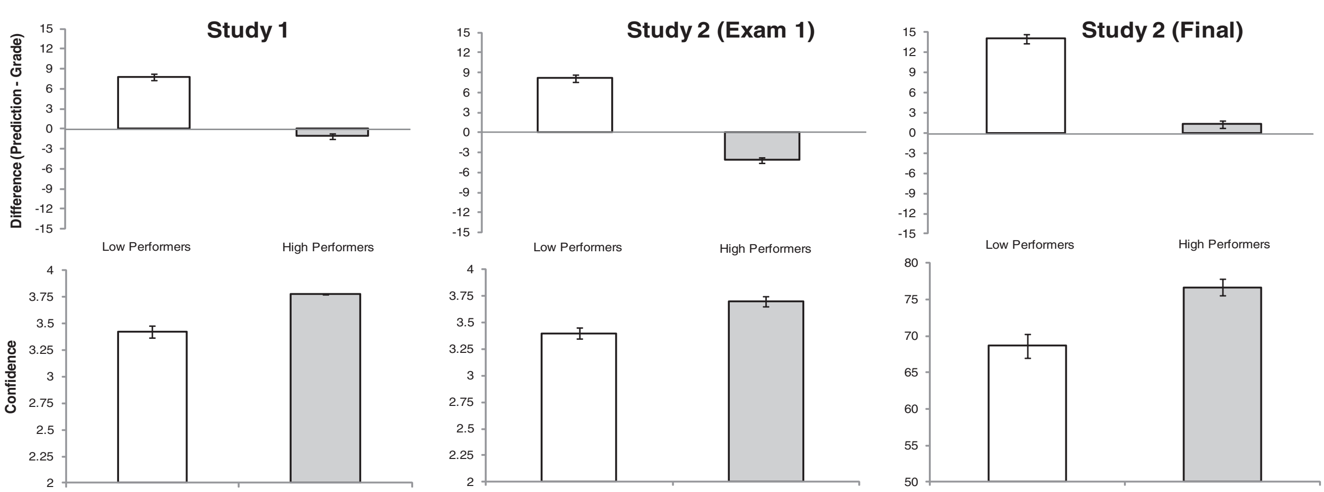

Figure 1. Functional and subjective confidence for Studies 1 and 2 for the bottom and top quartile participants. Note that participants rated subjective confidence on a scale of 1–5 on Exam 1 for both studies and on a scale of 1–100 on the final exam in Study 2. Error bars depict standard errors.

dents predicted their exam scores more accurately than did low-performing students, $F(3, 111) = 39.04$, $MSE = 50.00$, $p < .001$, $\eta_p^2 = .51$. More important, though, results again showed that under different exam prediction instructions, students' subjective confidence ratings differed by quartile (see Table 1), $F(3, 111) = 2.58$, $MSE = 0.54$, $p = .057$, $\eta_p^2 = .07$, showing that confidence was highest in the high performers and lowest in the low performers (see Figure 1). Thus, the current results using number grade estimations replicated those from Study 1 using letter grade estimations.

Final exam. We also examined grade predictions and confidence on the final exam for this study. As expected, higher performing students were significantly more calibrated than lower performing students in their exam grade predictions, $F(3, 79) = 22.53$, $MSE = 61.97$, $p < .001$, $\eta_p^2 = .46$. In line with Exam 1 findings from Studies 1 and 2, we again showed a dissociation between functional and subjective confidence: Low-performing students were less confident than high-performing students in their exam predictions, $F(3, 79) = 2.94$, $MSE = 312.68$, $p = .038$, $\eta_p^2 = .10$ (see Table 1 and Figure 1). Note that although the pattern of data is less regular for the final exam compared with the previous exam, which was likely due to the fact that only a subset of participants took that final exam, the overall results are the same. Thus, despite the fact that students had much experience with the course material and exams, they continued to show functional overconfidence; that effect continued to be larger in low-performing students than in high-performing students. In addition, although low-performing students continued to be functionally overconfident, they also continued to be less subjectively confident in their predictions than high-performing students.

After the students completed the final exam, they again rated their confidence in their original exam prediction. Postexam confidence ratings were lower than preexam confidence ratings, $t(77) = 4.66$, $p < .001$, but each quartile similarly reduced their confidence after the exam as indicated by the lack of a significant F test, $F(3, 74) = 1.43$, $MSE = 741.69$, $p = .242$, $\eta_p^2 = .05$ (see Table 2). Subjective confidence remained lower for low-performing students than for high-performing students after the final exam was completed.

Discussion

Both Study 1 and Study 2 showed a consistent dissociation between functional and subjective overconfidence, such that low-performing students were more functionally overconfident but not more subjectively overconfident than high-performing students. This pattern of results occurred regardless of whether participants predicted their scores as a letter grade or as a percentage. The pattern also held regardless of whether participants could earn incentives for accuracy and regardless of whether their predictions were for the first exam or the final exam in the course.

The fact that low-performing students were less subjectively confident in their predictions than were high-performing students supports the notion that low-performing students may have some awareness of their ineptitude. That low-performing students have any metacognitive awareness provides evidence against the strongest version of the double-curse account, which suggests that low performers overestimate performance because they are unaware of their lack of metacognitive knowledge (see Kruger & Dunning, 1999).

We have offered the straightforward interpretation that low performers are less confident in their predictions than are high performers because they simply have less confidence that these predictions are correct. However, it could be that low-performing students were less confident in their predictions than were high-performing students because they thought that their guesses were too low and that they might perform better on the exam than predicted. Alternatively, the majority of low performers could have lacked confidence, compared with high performers, because they thought that their guesses were too high. Of course, it is always possible that there may be individual differences in how people make confidence judgments (or any judgments). For these potential differences in the reasons for lowered confidence to explain the pattern of data, they would have to be systematic (with the vast majority of low performers making their confidence judgments because they thought that their grade prediction was either too low or too high). We think this is unlikely, but future research could examine this issue by asking students to indicate whether they believe that they guessed too high or too low, or whether they simply do not know, after they make their subjective confidence rating.

If, as the current data suggest, low-performing students have some awareness of their metacognitive failing, then why do they consistently predict that they will perform so much better than they do? Perhaps this finding results from a measurement artifact, as previously suggested (Krueger & Mueller, 2002), whereby low-performing students have more room to predict higher than they perform than do high-performing students. Alternatively, we suggest that low-performing students may not know how they will perform so they simply make a reasonable guess—they guess that they will receive a low B or high C (arguably a fairly average grade in many college courses) but they actually receive a low C. In fact, the current data, in conjunction with previous findings (Hacker, Bol, & Bahbahani, 2008; Hacker et al., 2000), show that low-performing students predict, on average, that they will receive approximately an 80% on exams, but they actually perform much worse. Regardless of the exact reason for the greater functional overconfidence effect for low performers, the present data indicate that low performers are not as confident in these predictions as high performers are, implying that low-performing students are not entirely unaware of their deficits. Further, the current data demonstrate a dissociation between metacognitive ability and awareness of this ability. Functional overconfidence was greater for low

Table 2
Study 2: Mean Preexam and Postexam Confidence for Exam 5 by Performance Group

	Preexam		Postexam	
Quartile	Confidence	*SE*	Confidence	*SE*
1	64.76	4.48	59.94	6.36
2	72.70	4.51	49.37	5.75
3	80.95	2.17	65.29	4.87
4	72.38	3.96	63.14	5.13

Note. Students reported confidence on a Likert-type scale (1–100).

performers than for high performers, whereas subjective confidence was lower for low performers than for high performers.

Other research shows that one can parse overconfidence into different types (Moore & Healy, 2008). These researchers identified three types of confidence: (a) *Overestimation* occurs when people predict that they will perform better than they do, (b) *overplacement* occurs when people believe that they are better than others, and (c) *overprecision* occurs when people are unduly certain that their prediction is correct. To our knowledge, Dunlosky et al. (2005) and the current study are the only empirical dissociations between two of these types of overconfidence (i.e., overestimation and overprecision), and the current study is the first to examine these types of overconfidence in the classroom. Future studies could examine whether low and high performers differ on other forms of confidence.

Given the current results showing consistent differences between functional and subjective confidence, future research might also consider the contribution of both types of confidence to performance. For example, classroom studies show that low performers' functional overconfidence is fairly resistant to intervention (see Miller & Geraci, 2010; Nietfeld et al., 2005). Future research might take advantage of the fact that subjective confidence judgments appear to more accurately reflect actual performance than functional confidence. For example, if one asked student participants in the classroom to make functional and subjective confidence judgments and then asked participants to focus on the subjective confidence to inform their studying behavior, even low-performing students may be able to improve their performance.

References

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *The Journal of Experimental Education, 73,* 269–290. doi:10.3200/JEXE.73.4.269-290

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67,* 366–381. doi:10.1037/0022-3514.67.3.366

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90,* 60–77. doi:10.1037/0022-3514.90.1.60

Dunlosky, J., Serra, M. J., Matvey, G., & Rawson, K. A. (2005). Second-order judgments about judgments of learning. *The Journal of General Psychology, 132,* 335–346. doi:10.3200/GENP.132.4.335-346

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself.* New York, NY: Taylor & Francis. doi:10.4324/9780203337998

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5,* 69–106. doi:10.1111/j.1529-1006.2004.00018.x

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105,* 98–121. doi:10.1016/j.obhdp.2007.05.002

Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning, 3,* 101–121. doi:10.1007/s11409-008-9021-5

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92,* 160–170. doi:10.1037/0022-0663.92.1.160

Kelemen, W. L., Winningham, R. G., & Weaver, C. A., III. (2007). Repeated testing sessions and scholastic aptitude in college students' metacognitive accuracy. *European Journal of Cognitive Psychology, 19,* 689–717. doi:10.1080/09541440701326170

Knouse, L. E., Bagwell, C. L., Barkley, R. A., & Murphy, K. R. (2005). Accuracy of self-evaluation in adults with ADHD: Evidence from a driving study. *Journal of Attention Disorders, 8,* 221–234. doi:10.1177/1087054705280159

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82,* 180–188. doi:10.1037/0022-3514.82.2.180

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,* 1121–1134. doi:10.1037/0022-3514.77.6.1121

Miller, T. M., & Geraci, L. (2010). *Training metacognition in the classroom: The influence of incentives and feedback on exam predictions.* Manuscript submitted for publication.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115,* 502–517. doi:10.1037/0033-295X.115.2.502

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270. doi:10.1111/j.1467-9280.1991.tb00147.x

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in a postsecondary classroom. *The Journal of Experimental Education, 74,* 7–28.

Pierce, B. H., & Smith, S. M. (2001). The postdiction superiority effect in metacomprehension of text. *Memory & Cognition, 29,* 62–67.

Preuss, G. S., & Alicke, M. D. (2009). Everybody loves me: Self-evaluations and metaperceptions of dating popularity. *Personality and Social Psychology Bulletin, 35,* 937–950. doi:10.1177/0146167209335298

Williams, A. F. (2003). Views of U.S. drivers about driving safety. *Journal of Safety Research, 34,* 491–494. doi:10.1016/j.jsr.2003.05.002

Williams, W. M. (2004). Blissfully incompetent [Editorial]. *Psychological Science in the Public Interest, 5,* i–ii. doi:10.1111/j.1529-1006.2004.00017.x

Received April 7, 2010
Revision received September 14, 2010
Accepted September 18, 2010 ■