Isaacson, R. M., & Was, C. A. (2010). Believing you're correct vs. knowing you're correct: A significant difference? *The Researcher, 23*(1), 1-12.

The 2009 NRMERA Distinguished Paper

Believing You're Correct vs. Knowing You're Correct: A Significant Difference?

Randy M. Isaacson
Indiana University South Bend

Christopher A. Was
Kent State University

Metacognition is often described as thinking about one's thinking. A more precise definition is knowledge and control over one's cognitive processes. Poor metacognition often causes students to overestimate their understanding which leads to ineffective learning. This overestimation is specifically caused by poor metacognitive knowledge monitoring. The current study examines college students' ability to monitor and accurately assess their prior knowledge. We found that undergraduates who are more accurate at assessing their existing knowledge perform better on course exams than students who are less accurate at assessing their knowledge. There is hope however, that metacognitive knowledge monitoring is a skill that can be taught and learned. The findings have important implications for the college and k-12 instruction.

Keywords: *metacognition; knowledge monitoring; knowledge assessment*

Every college teacher has the experience of working with a student who claims they have mastered the course content but did poorly on the exam. Upon further questioning it becomes clear the student does not understand the material despite their continuing accusation that their failure was due to an unfair test. Our first reaction may be that the student is lying, but further inquiry often leads us to recognize the student honestly believes they understand the concepts. Students often believe memorization is understanding, which can explain their perspective, but a deeper exploration demonstrates that a student's poor monitoring of their understanding not only leads them to argue with you, but also causes problems with learning. Successful students have developed strategies to monitor their comprehension as they are learning; they are effective at metacognitive knowledge monitoring.

Metacognition is often defined as knowledge and control over one's cognitive processes. For example, Shimamura (2000) refers to metacognition as the evaluation and control of one's cognitive processes. Shraw and Moshman (1995) described the difference between metacognitive knowledge and metacognitive control processes. The basic distinction is what one knows about cognition and the control and regulation one exerts over their own cognition. Pintrich, Wolters, and Baxter (2000; as cited in Tobias & Everson, 2002, 2009) described three distinct components of metacogntion: knowledge regarding metacognition, monitoring of one's learning processes and the control over the learning processes. In the current study, we focus on

Contact information: Christopher A. Was, Kent State University, Lifespan Development and Educational Sciences, 405 White Hall, Kent, OH 44242, Tel: 330-672-2929, Fax: 330-672-2512, E-mail: cwas@kent.edu

a specific form of regulation: monitoring. Monitoring as described by Shraw and Mosham is the online awareness of understanding and performance. This ability develops gradually and has been demonstrated to be poor in both children and adults (Pressley & Ghatala, 1990). Pressley and Ghatala present a review of their work and work conducted with colleagues. It is clear from the review of their recent research in monitoring that both adults and children are poor at on-line evaluation of knowledge and comprehension.

The ability to monitor prior knowledge has been argued to be necessary for new learning from instruction (Tobias & Everson, 2009). However, it is not clear how the monitoring of prior knowledge or previous learning is used to learn from instruction? In order to answer this question, it is first necessary to determine if individuals are effective monitors of their own knowledge. Flavell (1979) stated that metacognitive experiences, such as the feeling that you understand something, are likely to occur in situations or tasks in which you are required to utilize that type of thinking. Judgments of learning (JOLs) could be considered one such task. Son and Metcalfe (2000) reviewed the literature regarding metacognitive JOLs when participants learn a list of word pair associates and then predict the likelihood they will be able to recall the target word from a pair when presented with the cue word. Son and Metcalfe found that people tend to dedicate more time to items judged to be difficult to learn than to those judged to be easy or less difficult to learn. This tendency to spend more time on items judged as difficult leads to a negative correlation between JOLs and study time.

At least two explanations the use of metacognition to control study during learning have been proposed. The region-of-proximal learning explanation, proposed by Metcalfe and Kornell (2005), describes the negative correlation often found between high JOL and study time. Metcalfe and Kornell proposed that individuals choose to eliminate to-be-studied items they believe they already know or have mastered. When people eliminate the high JOL items they tend to choose to study the easiest items next rather than the hardest items. The more difficult items are only studied once the next set of items is mastered. Metcalfe (2009) compares the region-of-proximal-learning framework to Vygotky's zone of proximal development and other theories in that it is a region in which one's study efforts are most likely to be efficient and effective.

Dunloksy and Hertzog (1998) proposed an alternative model to describe the negative correlations between JOL and study time. The discrepancy-reduction model states that individuals devote more time to items judged to be farther from an internal learning criterion. That is, items judged to have not been learned to an internal standard are studied for longer periods to reduce the discrepancy between the criterion and the JOL. In the context of both the region-of-proximal-learning and the discrepancy-reduction model, it is clear that if students cannot to distinguish between what they know and they do not know, they are unlikely to engage in efficient study practices. When a student mistakenly overestimates their previous learning (an inaccurate high JOL) they are likely to disengage from learning prematurely. When a student underestimates their previous learning (inaccurate low JOL) they are likely to spend time more time than necessary to achieve understanding. It is clear, as Tobias and Everson (2009) stated, without a foundation of good knowledge monitoring students are unlikely to effectively engage in higher-order metacognitive processes such as planning, strategy use and evaluation of new learning.

Knowledge monitoring has also been measured using calibration techniques. Calibration of knowledge monitoring is operationalized as the difference between predicted performance and actual performance. Classroom research in college settings (e.g., Hacker, Bol, Hogan & Rakow, 2000; Isaacson & Fujita, 2001) has illustrated a striking difference between high and low performing students in their ability to predict their test scores (calibration): successful students accurately estimate their grades before taking a test, while students on the verge of failure often over-estimate their future performance. Hacker et al. gave undergraduates three multiple-choice exams over the course of a semester. The course in which the study was conducted had a focus of self-evaluation and the relationship to performance. Results indicated that the highest performing students were accurate in the predictions of exam scores as well as the post-diction (estimates of performance following the exam) of performance. In turn, the lowest performing students were the least accurate in prediction and post-diction of exam scores and tended to greatly over estimate their performance. Isaacson and Fujita had undergraduate students complete 10 weekly examinations over the course of a semester. Prior to each exam students were asked to identify the number of hours studied and their level of confidence in achieving their goals on the test. As in the Hacker et al. study, following the exam but before having it graded, students were asked to post-dict their exam score. As in the Hacker et al. study high achieving students were better at post-dicting their test scores than the lower achieving students.

Other studies have also demonstrated a relationship between college student performance and metacognitive knowledge monitoring. Vadhan and Stander (1994) found that the difference between the post-diction of exam scores and actual scores was negatively correlated with the actual scores. In other words, the better the calibration, the better the exam score. The empirical evidence suggests that good students can predict when they are going to be successful in a class or on a test, but how does metacognition impact their learning?

The body of literature investigating the relationship between reading comprehension and metacognition may be informative. A number of studies have demonstrated the importance of metacognition in the reading comprehension of college students. Maki and Berry (1984) found that following the reading of texts, students who scored above the mean on a subsequent multiple-choice test covering the text were significantly better at predicting their scores on the tests than those below the mean. Maki, Foley, Kajer, and Thompson (1990) varied the amount of processing needed to comprehend a text. Deleting letters from words within the text accomplished increasing the processing requirements. Increased processing requirements led to better calibration (the relationship between prediction of comprehension performance and actual performance). These and other findings by Maki (e.g., Maki, 1995; Maki & Serra, 1992) support the conclusion that successful college students possess the levels of metacomprehension required to successfully complete tasks requiring higher-order thinking skills.

The knowledge monitoring assessment (KMA) developed by Tobias and Everson (2000, 2002) demonstrates that students able to differentiate between when they know and when they do not know are more likely to excel than students who are not able to distinguish their level of comprehension or learning. Tobias and Everson (2009) provide a summary of the research using the KMA as a measure on the ability to monitor knowledge and the relationship to academic performance. Germane to the current study, is research from Everson and Tobias (1998) that in two studies demonstrated a correlation between the KMA and college student grade point average. From this research, it is easy to ascertain that metacognitive knowledge monitoring is an important component to academic performance. The issue central to educators and researchers alike is whether knowledge monitoring is a skill that can be taught or innate ability.

Paris and Winograd (1990) argued that instructional techniques can promote and enhance student metacognition and in turn learning. However, the evidence regarding whether one can be taught to better monitor their knowledge is, inconclusive. Studies conducted with elementary and middle school students appear to have somewhat consistent conclusions regarding students increasing their monitoring. Delcos and Harrington (1991) demonstrated that 5th and 6th grade students trained in problem solving and self-monitoring solved more difficult problems and in less time than students trained only in problem-solving and a control group. Ghatala, Levin, Pressley and Lodico (1985) trained 2nd graders in the utility of monitoring strategies, the affective consequences of strategy use or no strategy training. Students trained in the utility of monitoring strategies most quickly abandoned in effective strategies for the memory task. Thiede and Anderson (2003) demonstrated an increase in judgment accuracy when students summarized texts, improving the relationship between confidence and performance. In two studies, Theide and Anderson had college students read texts and then summarize each text either immediately following reading or after a delay. Students were better at meta-comprehension accuracy when summarizing came after a delay rather than immediately following reading. The measure of accuracy was the calibration for self-rating of comprehension and a 6 item multiple-choice test for each text. Theide and Anderson proposed that summarizing the text after a delay improved meta-comprehension accuracy. They describe their findings from a discrepancy-reduction model of self-regulated learning and the findings support the hypothesis that knowledge monitoring or at least metacomprehension is a skill that can be learned. Son (2004) found that distributed practice also increased relative judgment accuracy. However, studies by Bol, Hacker, O'Shea and Allen (2005) and Nietfeld, Cao and Osbourne (2005) found no improvement in monitoring accuracy even after a semester of monitoring practice. A more recent study conducted by the same group of researchers (Neitfeld, Cao and Osbourne, 2006) found that an intervention of monitoring exercises and feedback had a significant impact of students' calibration and test performance. These inconsistent findings beg the question, if knowledge monitoring skills improve over time in classroom situations and these skills are critical to academic success, what instructional approaches can be used to facilitate these improvements?

In other related research, metacognitive knowledge monitoring has been shown to impact self-regulated learning. In an experimental study (Thiede, Anderson, & Therriault, 2003) college students who were instructed to reflect on their comprehension of reading text were found to regulate their study behavior more effectively and perform better on subsequent learning tasks than students who were not encouraged to monitor their learning. It was hypothesized that when students are encouraged to focus on the discrepancy between their current state of learning and their desired learning state they are more likely use self-regulated learning to remove this discrepancy to achieve their goals. The authors propose that metacognitive monitoring plays a critical role in the connection between academic performance and self-regulated learning in a laboratory environment where it is encouraged but, "...it seems likely that left to their own devices people will not accurately monitor comprehension." (p. 71) If students who are instructed to reflect on their understanding in a laboratory environment are more effective in regulating their learning, what can be done in a college course to encourage students to be more reflective, and does this metacognitive monitoring improve their academic performance?

The goal of the current study was to answer a number of questions important to education. Knowledge monitoring is clearly related to student learning but how does it impact learning? Do some errors in knowledge monitoring have a more negative impact than others? Is knowledge monitoring a disposition or skill some students possess and use in their classes, or is this skill content specific? Can knowledge monitoring be measured to the extent that the disposition/skill at the start of a class can predict student learning at the end of a class? Although the goals are quite lofty, we feel the current study can shed some light on these issues.

METHOD

PARTICIPANTS

One hundred twenty nine undergraduates enrolled in an educational psychology course participated for course credit. Not all of the students completed all measures and therefore there is missing data. All analyses were completed using list-wise deletion and no analysis was completed with less than 106 participants. All participants were of sophomore or junior standing. Females represented 74% of the sample.

MATERIALS

KMA. The KMA utilized for this study is an evaluation designed to measure metacognition by measuring the difference between students' estimates of how well they are to perform and their actual performance (Tobias & Everson, 2000). The current task was adopted from Tobias and Everson (2000; for a review of the literature regarding this assessment see Tobias & Everson, 2009). The KMA designed for the current study required participants to state whether they knew or did not know the meaning of 50 English words, and then respond to a multiple choice vocabulary test of the same words. Using terminology common to signal detection theory, the procedure generates the following four scores with students assessing the word was: 1) known and correctly responded to the item on the vocabulary test [hits: + +]; 2) known but responded to incorrectly on the test [false alarms: + -]; 3) unknown but the correct response was given on the test [misses: - +]; and 4) unknown and responded to incorrectly on the test [correct rejections: - -] (see Table 1).

Table 1. *Accuracy Assessments Generated by KMA*

	Response Accuracy	
Feeling of Knowing	Correct Response	Incorrect Response
Know	Hits [+ +]	False Alarms [+ -]
Don't Know	Misses [- +]	Correct Rejections [- -]

Final Exam. The final exam was a multiple-choice exam with one hundred items; half of the questions were from prior tests (i.e., old questions) and half of the questions were new questions. Students were asked to use metacognition to distinguish their level of confidence for each answer using an incentive point system for questions at three levels of difficulty. Students were required to place answers for which they were confident "on the left" of their answer sheet (74 questions for increased points) and the answers for which they were less confident "on the right" (26 questions for one point each). Each student has a total score (points of a possible 298) and an absolute score (number of items correct of a possible 100).

DESIGNS AND PROCEDURES

Participants completed the KMA twice during the semester in their educational psychology course. The first administration (PreKMA) took place within the first two weeks of the semester and the second administration during the last two weeks (PostKMA) of the semester. Both administrations were completed online. During both administrations of the KMA, participants were presented with 50 vocabulary words one at a time. Thirty-three of the words represented vocabulary items derived from the text used for the educational psychology course and 17 represented general vocabulary items. When each of the 50 items was presented, participants were required to signify whether they knew or did not know the meaning of the word. After all 50 items were presented, participants were presented with the vocabulary items and five possible synonyms. Participants responded by indicating which of the five alternatives they believed to be the synonym.

During the semester, students were presented with a number of tasks that invited or demanded that they examine their metacognitive knowledge monitoring and other tasks where they were asked to reflect on their ability to monitor their understanding of concepts in the course. Each week students were administered a test based on variability difficulty - variable weight and metacognition. Each test had three levels of difficulty based on Bloom's taxonomy (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956), each assigned a different weight. Fifteen Level I questions, worth 2 points each, assessed knowledge and comprehension; Fifteen Level II questions, worth 5 points each, assessed application; and 5 Level III questions, worth 6 points each, assessed analysis and synthesis. Students were also required to choose the questions they were least confident about and "put them on the right" of the answer sheet where these questions were only worth one point (5 of the 15 Level I and II questions, and 2 of the 5 Level III questions). The total score possible for each test was 100 points. The test format creates an incentive for students to reflect on and improve their metacognition. Students are asked to complete a pre-post test questionnaire which asks them to predict and postdict their test scores before the test is graded and also asks them to reflect on the test after it is graded.

Throughout the semester long course, students were also presented with a number of resources in the curriculum to improve metacognition. Students were encouraged to take on-line practice tests each week that have a format similar to the KMA where students are asked about their confidence of each answer before the practice test is graded on-line. The course had a website (Isaacson, 2009; http://mypage.iusb.edu/~risaacso/ed_psych/) with a variety of resources that have been developed to improve metacognition (e.g., Ed Sykes and Pick 'Em). The course had small discussion classes led by peer mentors where students take a quiz with a format similar to the test each week and other resources designed to focus on metacognition. Students also submitted a journal to their peer mentor each week that focuses on self-regulated learning and metacognition. The class had two lectures each week and students were presented with a Question of the Day twice during class with their answer to these questions recorded using a student response system (i.e., "clickers") that allows students to indicate whether they are Absolutely Sure, Fairly Sure, or Just Guessing at the answers: the students response system has an incentive for knowing-when-you-know.

RESULTS

A discussion of the KMA outcome measure used in the current study is warranted before discussing the results. Schraw (2009) describes the differences between outcomes measures used to capture monitoring accuracy. The decision to use a specific outcome measure should be determined by the construct one wishes to measure. Measures of relative accuracy focus on the reliability of a set of judgments relative to a set of outcomes rather than the precision of item specific judgments (Shraw, 2009). For our purposes in the current study a measure of relative accuracy is the appropriate outcome measure. Our goal was to capture the students' consistency between their judgments-of-knowing and their actual performance on the vocabulary test, not the degree of accuracy on judgments on an item-by-item basis. Nelson (1984) recommended using the Goodman-Kruskal gamma correlation (Goodman & Kruskal, 1954). Schraw (1995) also stated that when the researcher is interested in the association between performance and judgments-of-knowing, the gamma correlation is the appropriate statistic. For these reasons, we chose to use the gamma correlation for the measure of knowledge monitoring. The mean of the final exam total score was 174.75 out of a possible 300 with a standard deviation of 31.84. The mean of the final exam absolute score was 70.51 out of a possible 100 with a standard deviation of 13.58. The correlation between final exam total score and absolute score was r = .98, therefore we choose to focus the remaining analyses using absolute score or (percent correct). Table 2 presents the means and standard deviations of the four scores generated on both the PreKMA and PostKMA.

Table 2. *PreKMA and PostKMA Means and Standard Deviations by Accuracy Assessment.*

Variable	PreKMA Mean (SD)	PostKMA Mean (SD)
Hits [+ +]	16.99(8.77)	21.53(11.97)
Correct Rejections [- -]	14.94(7.19)	8.25(5.95)
False Alarms [+ -]	6.95(5.04)	7.33(6.54)
Misses [- +]	9.63(5.18)	5.33(3.90)

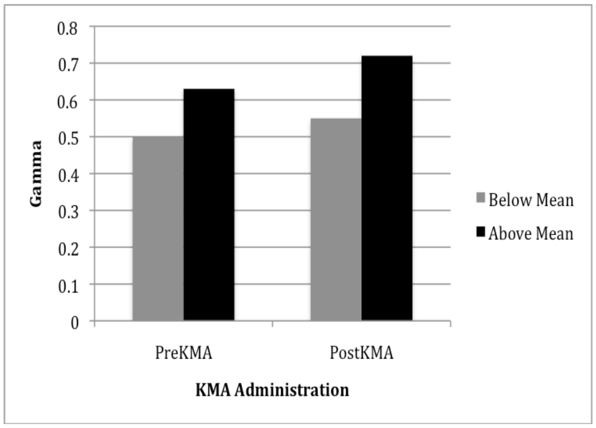
Table 3. *Correlations Between Final Exam and Pre and PostKMA Gamma Coefficients.*

Variable	PreKMA	PostKMA
Correct Hits [+ +]	.41**	.38**
Correct Rejections [- -]	-.20*	-.20*
False Alarms [+ -]	-.39**	-.36**
Misses [- +]	.29**	.26**

In order to examine the overall relationship between the final exam in the educational psychology course and metacognition as measured by the KMA, a Gamma coefficient (□; Goodman & Kraskul, 1954) was calculated for each participant for both the PreKMA and PostKMA. As described above, □ is a measure of relative accuracy. The correlation between PreKMA gamma and PostKMA gamma coefficients was r =.28, p =.003. The correlations between both PreKMA and PostKMA □ coefficients and final exam absolute score was significant, r =.42, p < .001, and r =.37, p < .001 respectively (see Table 3).[1] These results indicated that knowledge assessment accuracy in the first two weeks and last two weeks of the semester are related to performance on the final exam at the conclusion of the semester.

Figure 1. *Gamma Coefficients for Students Scoring Above and Below the Mean on the Final Exam.*

[1]The correlations between both PreKMA and PostKMA γ coefficients and final exam total score r =.43, p < .001, and r =.38, p < .001.



To further highlight this relationship, two groups were created using a mean split of absolute final exam scores (see Figure 1). Students scoring above the mean had significantly higher PreKMA scores (Mean □ =.55) than those below the mean (Mean □ =.50), $t(113) = -2.71$, $p = .008$, $d = -0.13$, CI [-0.22, -0.03] and significantly higher PostKMA scores (Mean □ =.72) than those below the mean (Mean □ =.63), $t(104) = -4.79$, $p < .001$, $d = -0.18$, CI [-0.25, -0.10].

To further explore the relationship between knowledge monitoring accuracy and performance on the final exam, a final analysis was conducted. PreKMA and PostKMA □ coefficients were correlated with final exam hits, correct rejections, false alarms and misses (see Table 3). Using final exam assessment accuracy scores based on the left/right design of the final exam, we calculated four scores similar to the scores generated by the KMA. For example, a response to an item on the final exam placed on the left and answered correctly was considered a hit. As is evident in Table 4, final exam performance is related to both accurate and inaccurate assessment. A repeated measures ANOVA found a significant effect of group [above vs. below the mean; $F(1, 99) = 8.05$, $p < .05$], indicating a difference between mean gamma for pre and post KMA's, but no interaction effect ($F < 1$). This indicated that the magnitude of increase in monitoring accuracy was equal for the groups, but that the gamma correlations were different.

Table 4. *Final Exam Knowledge Assessment Scores.*

Knowledge Assessment Accuracy	Above Mean on Final Exam (*n* = 53) Mean (SD)	Below Mean on Final Exam (*n* = 76) Mean (SD)
Correct Hits	64.43(3.93)	48.91 (7.32)
Correct Rejections	10.38 (3.59)	15.49(3.59)
False Alarms	8.62(3.91)	23.66(6.73)
Misses	15.57 (3.63)	9.74(2.77)

Table 4 presents the final exam assessment accuracy means and standard deviations separated by mean split for final exam total score. Employing a two group ANOVA it was found that those scoring above the mean made significantly more correct hits than those scoring below the mean, $F(1, 125) = 237.49$, $p < .001$, □$_p^2$ = .66. Those scoring above the mean also had fewer false alarms, $F(1, 125) = 251.94$, $p < .001$, □$_p^2$ = .67. Both of these results indicated that students scoring above the mean on the final were more accurate in their assessment of items they knew.

Finally, we tested the hypothesis that the self-regulation and metacognition training provided in the course would increase students' knowledge monitoring accuracy. A paired samples t-test found that PostKMA gamma was significantly higher than PreKMA gamma, $t(104) = -3.217$, $d = -0.09$, CI [-0.14, -0.03] $p = .002$.

Discussion

The correlations between the final exam and the PreKMA and PostKMA demonstrate a strong relationship between students' metacognitive ability and their learning, even when the measure of their metacognitive ability occurs two months prior to the demonstration of their learning as was the case with the PreKMA. As utilized in the current study, the KMA represents a measure of relative accuracy of the students' metacognitive knowledge monitoring. In other words, the KMA represents a measure of each student's ability to accurately monitor his/her existing knowledge not simply a measure a of item-by-item feeling of knowing. Not surprisingly, this ability to monitor one's knowledge was found to have a relationship with the students' final exam as demonstrated by the correlation analysis.

Further evidence of this relationship was provided by the final exam mean split analysis. Although the correlations between the PreKMA and PostKMA gammas and the final exam are informative, the mean split analysis provides more information regarding the relationship between the KMA gamma scores and performance on the final exam. It is quite clear from this analysis that students performing poorly on the final exam are inaccurate in the KMAs. Based on previous research and theory, one can assume that these poor performing students do not possess the basic metacognitive skills, such as monitoring, to support efficient higher level metacognitive skills such as evaluation and planning. However, the data indicated that the mean increase in knowledge monitoring accuracy from the PreKMA to the PostKMA did not differ between those scoring above the mean and those below the mean of the final exam (see Figure 1).

When comparing the PostKMA to the PreKMA (see Table 2) students had more hits, fewer correct rejections, and fewer misses – all of which can be attributed to classroom learning and increased knowledge monitoring skill. A surprising result was the significant increase in the number of false alarms from the PreKMA to the PostKMA[2]. As proposed in the opening paragraph of this manuscript, this may indicate that some students are judging "knowing" based on recognition or memorization rather than true understanding. Schomer and Surber (1986) describe an illusion of knowing (IK) when students believe that shallow processing is all that is necessary to learn material that is most likely difficult to learn. Several other researchers have proposed metacognitive errors related to IK (e.g., Dunloksy, Rawson, & Middleton, 2005; Glenberg, Wilkinson, & Epstein, 1982; Nelson & Dunlosky, 1991). As described, the KMA contain 33 items that were drawn from the vocabulary and content of the course in which the students were enrolled. It is highly likely that on the second administration of the KMA, many students had a strong IK based on familiarity with the vocabulary, but not a true understanding of the terms. This false sense of understanding may explain the student who argues with the teacher that he understood the material but failed the test – looking over your notes and recognizing all the terms does not mean you understand all the information required to succeed on the test.

There is some empirical evidence (and extensive anecdotal evidence) that effective students improve their metacognitive ability when given an incentive to reflect on their learning and the appropriate resources to improve learning. Given this association it is important for educators to examine this relationship and explore what classroom practice might be used to improve the metacognition of students. The current study presented one professor's attempts to increase knowledge monitoring among a class of undergraduate students. The results support the conclusion that the metacognitive and self-regulation training may indeed increase at least the most fundamental component of metacognition.

As educators we must ask ourselves, "If students' learning is so closely associated with metacognition, can student

[2]An independent samples t-test created by splitting students into two groups: those above the mean on the PostKMA and those below the mean indicated that those below the mean had significantly more false alarms on the PostKMA, $t(106) = 4.67$, $p < .001$, $d = 2.72$, CI [1.46, 3.98].

metacognition be improved over time and what classroom instruction will support the development of effective metacognitive knowledge monitoring?" We feel the current analysis comparing PreKMA and PostKMA gamma coefficients is a strong indicator that students can be trained to be better knowledge monitors. This basic metacognitive skill was demonstrated to increase on average over the course of the semester. Although this data is encouraging, there are a number of questions yet to be answered.

The current study was conducted within a college classroom setting and there were a number of interventions employed to increase students ability to accurately assess their knowledge and understanding. It is likely that the weekly test format was responsible, at least in part, for the increase in monitoring accuracy as measured by the KMA. However, current design did not allow for directly testing this hypothesis, due to the number of other activities in this classroom that might increase one's knowledge monitoring accuracy (e.g., the practice quizzes and self-regulation journals). Future research should include an experimental design in which two sections of this course are taught in the same manner and including the same materials. An experimental group would complete the exams in the same manner described in this study, while a control group completed the same exams (the same items) but without the weighting and left/right options. If the experimental group is found to have greater increases in knowledge monitoring accuracy that would provide evidence that the test format itself is an effective treatment to increase monitoring accuracy.

Finally, any number of individual difference variables might interact with or mediate the relationship between monitoring accuracy and academic achievement. Personality variables, motivational variables, and even interest in the topic may impact the relationship between monitoring accuracy and academic achievement.

REFERENCES

Bol, L., Hacker, D., 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*(4), 269-290.

Delcos, V. R., & Harrington, C. (1991). Effects of strategy monitoring and proactive instruction in children's problem-solving performance. *Journal of Educational Psychology, 83*(1), 35-42.

Dulosky, J. & Herzog, C. (1998). Training programs to improve learning later in adulthood: Helping older adults educate themselves. In D.J. Hacker, J.D. Dunlosky, & A. C. Grasser (Eds.), *Metacognition in educational theory and practice* (pp. 249-276). Mahwah, NJ: Earlbaum.

Dunloksy, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate monitoring and accessibility hypothesis. *Journal of Memory and Language, 52,* 551-565.

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

Flavell, J. H. (1979). Metacognitive and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-911.

Ghatala, E. S., Levin, J. R., Pressley, M., & Lodico, M. G. (1985). Training cognitive strategy-monitoring in children. *American Educational Research Journal, 22*(2), 199-215.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition, 10*, 597-602.

Goodman, L. A. & Kruskal, W. H. (1954). Measures of association for cross classification. *Journal of the American Statistical Association, 49*, 732-764.

Hacker, D., Bol, L., Horgan, D., & Rakow, E. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(4), 160-170.

Isaacson, R. (2009). Retrieved from: http://mypage.iusb.edu /~risaacso/ed_psych/index.htm

Isaacson, R., & Fujita, F. (2001, April). *The effects of goals, expectations, and self-efficacy on self-regulation and performance in college students.* Presented at the Annual conference of the American Educational Research Association, Seattle, Washington.

Maki, R. (1995). Accuracy of meta-comprehension judgments for questions of varying importance levels. *The American Journal of Psychology, 108*(3), 327-344.

Maki, R. H., & Berry B. S. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 663-67.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 609-616.

Maki, R. & Serra, M. (1992) Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology, 84*(2), 200-210.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159-163.

Metcalfe, & Kornell (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*, 463-477.

Nelson, T. O. (1984). A comparison of the current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*(1), 109-133.

Nelson, T., & Dunlosky, J. (1991). When People's Judgments of Learning (JOLs) Are Extremely Accurate at Predicting Subsequent Recall: The" Delayed-JOL Effect". *Psychological Science, 2*(4), 267-270.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy and self-efficacy. *Metacognition and Learning,* 159-179.

Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. Dimensions of thinking and cognitive instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of Thinking and Cognitive Instruction.* (pp. 15-52). Hillsdale, NJ: Lawrence Erlbaum.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp 43-98). Lincoln, NE: Buros.

Pressley, M., & Ghatala, E. S., (1990).Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25*, 19–34.

Schomer, M. & Surber, J. R. (1986). Comprehension monitoring failure in skilled adult readers. *Journal of Educational Psychology, 78*(5), 353-357.

Schraw, G., (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology, 9*, 321-332.

Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education.* (pp. 415-429). New York, NY: Routledge.

Schraw, G., & Moshman, D. (1995) Metacognitive theories. *Educational Psychology Review 7*(4), 351-371.

Shimamura, A.P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition, 9*, 313–323.

Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*(3), 601-604.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*(1), 204-221.

Thiede, K., & Anderson, M. (2003) Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology* (28), 129-160.

Thiede, K., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66-73.

Tobias, S., & Everson, H. (2000). Assessing metacognitive knowledge monitoring. In G. Schraw & J. Impara (Eds.), *Issues in measurement of metacognition* (pp. 147-222). Lincoln, NE: Buros Institute of Mental Measurement.

Tobias, S., & Everson, H. (2002). *Knowing what you know and what you don't: Further research on metacognitive knowledge monitoring.* College Board Report No. 2002-3. College Board, NY.

Tobias, S., & Everson, H. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education.* (pp. 107-128). New York, NY: Routledge.

Vadhan, V., & Stander, P. (1994). Metacognitive ability and test performance among college students. *The Journal of Psychology, 128*(3), 307-309.