Journal of the Scholarship of Teaching and Learning, Vol. 6, No. 1, August 2006, pp. 39 - 55.

Metacognitive Knowledge Monitoring and Self-Regulated Learning: Academic Success and Reflections on Learning

Randy M. Isaacson and Frank Fujita[1]

Abstract: During the past decade the relationship of self-regulated learning (SRL) to academic success has been extensively explored but the impact of metacognition in this process has not been thoroughly examined. This study examined the relationship of metacognitive knowledge monitoring (MKM) to classroom performance. Eighty-four undergraduate students in an introductory educational psychology class completed ten weekly in-class tests in which they were allowed to choose test questions. Students were asked to identify the number of hours they studied, their level of confidence, and to predict their test results after completing the test but before it was graded. High achieving students were: more accurate at predicting their test results; more realistic in their goals; more likely to adjust their confidence in-line with their test results; and more effective in choosing test questions to which they knew the answers. The study supports the relationship of metacognitive knowledge monitoring to self-regulated learning and academic success. *Keywords: self-regulated learning, metacognitive knowledge monitoring*

I. Introduction.

The application of self-regulation to learning is a complicated process involving not only the awareness and application of learning strategies but also extensive reflection and self-awareness. Pintrich (1995) describes self-regulation as the "active, goal-directed self-control of behavior, motivation, and cognition for academic tasks by an individual student." (p. 5) Students who are skillful at academic self-regulation understand their strengths and weaknesses as learners as well as the demands of the specific tasks. They approach learning with an assortment of strategies they might apply to achieve their goals and an understanding of when and how to implement their plan. But students who are expert learners have more than an arsenal of study strategies and the ability to regulate academic resources, they also know when they have mastered, or not mastered, the required academic tasks. That is, expert learners are also skillful at metacognitive knowledge monitoring (MKM).

Self-regulated learners are skillful at monitoring their learning and comprehension which has a direct effect on each step in the self-regulation process. Pintrich et al. (2000) compares monitoring to the thermostat of a furnace. When the temperature falls below a specified level the thermostat tells the furnace to turn on the heat; when a learner is confused or does not comprehend what they are studying the monitor tells the learner to regulate their behavior, cognitive strategies, or motivation and affect to increase learning. To be effective learners, students must adjust their efforts based on their awareness of their own understanding and the level of difficulty of the upcoming task. One of the critical barriers to success for many students may be their inability to objectively assess their mastery of the academic tasks they are facing.

Accurate monitoring of learning can impact self-regulation throughout the learning

---

[1] Respectively, Professor of Education, Indiana University South Bend, risaacso@iusb.edu and Associate Professor of Psychology, Indiana University South Bend, ffujita@iusb.edu.

process. Zimmerman (1998) proposes three phases to self-regulation which incorporate metacognition. The first phase is forethought which includes goal setting, strategic planning, and self-efficacy; students identify their goals, their plans for achieving them, and consider how likely it is they will achieve their goals. The second phase is performance or volitional control which includes attention focusing, self-instruction, and self-monitoring; students attempt the learning tasks and monitor what they are learning. The third phase is self-reflection which focuses on comparing self-monitored information with a standard or goal and reactions to the results. During the reflection stage students assess their success or failure, modify their self-efficacy, make causal attribution, and adapt for future learning.

In all three of these phases students are using academic goals as the yardstick against which they assess their learning; using the thermostat metaphor, goals are the set-temperature the thermostat uses to judge whether to turn-on the furnace. In each phase, self-reflection and self-monitoring are critical to master the skills of self-regulation as well as the content being learned. In each phase students who are skillful at self-regulation are cognizant of their understanding and adjust their goals and self-efficacy based on internal as well as external feedback on their mastery of the tasks.

Students' ability to monitor their learning is one of the key building blocks in self-regulated learning; students who are aware of the level of their mastery of material can adjust their study time and strategies. Over the past decade a number of concepts have been used to describe students' awareness of their learning. These concepts were originally investigated at a micro-level focusing on metamemory using an experimental format.

Ease-of-Learning (EOL) judgements refer to a student's inferences about how easy or difficult a task will be to learn (Nelson & Narens, 1990; Nelson & Leonesio, 1988). EOL judgments occur in advance of actual learning, but clearly are a stumbling block for students who underestimate the difficulty of the material or the level at which it must be mastered. Judgments of learning (JOL) occur during or after learning and are predictions of future test performance (Kelemen, 2000; Koriat, 1997; Nelson & Narens, 1990). Students whose JOL overestimate their actual learning are also likely to terminate their studying prior to mastering the material and fall short of their goals, especially when the demands of the tasks exceed the levels at which they have learned the required material.

These assessments of learning have important implications for self-regulated learning as each influence the thermostat that adjusts how much time and effort a student devotes to studying. But the primary focus of the research on EOL and JOL has been on metamemory with very little attention to the higher level of concepts and problem solving which occurs in classrooms, particularly in postsecondary classrooms. In these studies metamemory is not a measure of understanding or comprehension, but rather the ability to retrieve information from long term memory. While obviously not irrelevant to classroom learning, these measures are primarily applicable to the most basic learning which occurs in school. The connection to the self-regulation of cognitive strategies is probably most suitable to rehearsal or the most rudimentary forms of elaboration. Skillful self-regulators should be able to go beyond the assessment of their ability to recall facts; effective self-regulators should be able to estimate how well they have mastered a body of knowledge and how well they will be able to demonstrate their mastery. Skillful self-regulators should be able to predict how well or poorly they will do, and have done, on a test while naive self-regulators should be less able to estimate their academic success.

A series of studies by Maki and colleagues has shown the significance of metacognition on the comprehension of college students in experimental situations. In a study by Maki and Berry (1984) students who scored above the median on multiple choice items of text showed more accuracy in their test prediction than students who scored below the median. Students who had done well on the test were also better at predicting their test scores: greater metacomprehension could be inferred to influence self-regulation. Students who were asked to process information at high levels were also found to be more accurate in their test predictions (Maki, et.al., 1990) as were students who had increasingly more information about the text (Maki and Serra, 1992). Maki invited the connection between metamemory, metacomprehension, and classroom learning when she found an experimental relationship between higher order thinking questions and prediction accuracy. Studies of metamemory had used simple recall questioning, but Maki (1995) found that there was greater accuracy of predictions when students were asked higher order questions that did not require verbatim recall of information. This was in support of the finding of Weaver (1990) who had proposed that multiple questions had increased the reliability and calibration of comprehension assessment. These finding have important implications for post-secondary education which places greater value on higher level thinking skills. Do successful college students have better metacognitive awareness on tasks that require higher level thinking skills? Can they adjust their study skills in response to the increasing demand of these tasks?

The transition from high school to college puts many demands on young adults. In the classroom the greatest challenge may be the move from the declarative knowledge emphasized in high school to the higher level thinking skills typically required in college. When students do not recognize that these new demands require new approaches to learning and studying they may be unwilling and/or unable to make the necessary changes. At the core of this problem may be that students do not realize that their learning does not match the demands of the task. They assume they have learned the material if they can recall the important terms. They do not recognize that different academic tasks (e.g., different test formats) demand different levels of learning. They overestimate their understanding because they do not recognize the implication of different levels of learning and varying levels of task difficulty. To be successful in college, students may need to have a variety of SRL and metacognitive skills that were not necessarily essential in high school. First, successful college students recognize that professors expect more than the memorization of declarative knowledge. Second, successful college students use accurate MKM while studying to assess their mastery of the required material, particularly in relation to what will be require on the performance task (e.g., test). Third, successful college students have an arsenal of SRL strategies they can choose from to match their level of learning to the demands of the performance task. And finally, successful students are able to self-monitor their understanding and the demands of the performance task during evaluation to adjust their demonstration of the learning (e.g., to choose the right questions to answer during a test.) While it may be very difficult to assess metacognition before or during learning, it is possible to assess MKM during the assessment of their learning. While taking a test are expert students better at identifying what they know and what they do not know?

For the past decade a program of research by Tobias and Everson (2000, 2002) has examined learners' ability to differentiate between what they know and do not know. Their findings indicate learners of all levels of ability and developmental stages are affected by their ability to monitor their learning. In dozens of studies with students of all ages and abilities, Tobias and Everson have found that students who are 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. The studies by Tobias and Everson have focused on the correlation between knowledge monitoring and student's academic performance. But very little evidence exists which explores the relationship between knowledge monitoring and academic choices within classes: How do students who possess effective knowledge monitoring skills use these skills to make decisions which impact their academic success?

These experimental studies of metamemory and metacomprehension are an open invitation to naturalistic studies performed in actual classrooms. Studies on EOL (Nelson & Narens, 1990) suggest that students who underestimate the difficulty of classroom content may abandon their learning efforts before they have mastered the material. Experimental studies on JOL (Kelemen, 2000; Koriat, 1997) indicate that students overestimate their understanding which can lead to discontinuing learning efforts prior to mastery. Schommer & Surber (1986) demonstrated an Illusion of Knowing (IK) when students believe that only shallow processing is necessary when the material is actually difficult, which could lead some students to underestimate the necessity of learning classroom content to the degree which will be required for success. Each of these issues have clear implications for self-regulation in the classroom context. When students in post-secondary education are presented with learning tasks that require higher level thinking they must accurately judge the difficulty of material to be learned, accurately judge the level of their own understanding, and accurately judge the requirement of the performance task (e.g., test) they will be given in class.

While laboratory predictions of test performance are revealing of student's self-monitoring there is little classroom research on this topic. In the first study of classroom confidence, Shaughnessy (1979) reported high achieving students as being better able to distinguish between known and unknown information. In a study of self-efficacy, Sinkavich (1995) reports a significant difference between high and low achieving students on their confidence on individual test items. Hacker et al. (2000) report similar findings; high performing students were accurate in predicting their test results with their accuracy improving over multiple exams, while low performing students were poor at predicting test results.

Hacker et al. (2000) found significant differences between low performing students and high performing students in relation to their ability to predict (before taking a test) and postdict (after taking a test but before receiving their grade) their test results. The implication being that students who were doing poorly in a college course were unable to monitor their knowledge of the course material (i.e., they overestimate their test results in both prediction and postdiction) and therefore were unable to regulate their studying to assure mastery of the course material. The student sample in the Hacker et al. study were college students in an undergraduate educational psychology class in which the students took three tests. On each test the highest achieving students on that test were more accurate in their predictions and postdictions, but the analysis did not focus on the overall achievement of students across the semester. Would the highest achieving students for the semester be more accurate in their postdictions and would they be more likely to accurately adjust their test postdictions on a test-by-test basis? The present study examines students across ten tests during a one semester undergraduate educational psychology course comparing intra-individual differences for low and high preforming students.

All three levels of Zimmerman's (1998) academic self-regulation learning cycle emphasize the importance of goals setting. In the forethought stage students set goals, in the performance and volitional control stage students monitor their learning in relation to their goals, and in the reflection stage students assess their success or failure in relation to their goals. Extensive research has focused on goal setting (Locke & Latham, 1990) but no classroom research to date has examined the impact of goal setting and goal achievement on self-regulation. The present study will explore the relationship among the variables of goal setting, self-monitoring, performance, and self-regulation.

In the performance and volitional control stage of Zimmerman's (1998) learning cycle students are attempting the learning task and monitoring their mastery of the task. In this cycle students who possess good MKM should be able to reflect on the application of their learning to the task and choose appropriate tasks. When given choices of tasks of varying difficulty, such as a selection of diverse test questions, students with good metacognitive awareness should be capable of choosing tasks in which they will succeed and avoid tasks in which they are less likely to succeed. The present study will examine the ability of students to use MKM to choose between tasks of varying difficulty.

The self-reflective stage in Zimmerman's (1998) learning cycle examines the outcomes when students compare self-monitored information to their goals, especially in relation to the impact on self-evaluation, self-efficacy, and adaptation of strategies. The combination of goals with knowledge of performance impacts self-efficacy which heightens motivation (Bandura, 1997). Skillful self-regulators effectively monitor their progress in relation to their goals and then adjust their self-efficacy and future strategies. The present study will begin to explore how the feedback students receive from completing tasks of varying difficulty (e.g., tests) impacts their self-efficacy as they progress through a college course.

This study will explore the metacognitive differences between high achieving students and low achieving students. High achieving students, who have good MKM, should be more reflective and thoughtful about decisions they make in their studying, test taking, and self-efficacy. Are students who excel across the semester more likely to be more accurate in their estimation of their individual test grades and do they accurately adjust their estimations based on their mastery of the material? Do students who excel have goals that are more consistent with their academic performance? Do students who excel have realistic self-efficacy and how does that self-efficacy change over the course of a semester? Are students who excel more likely to make accurate task choices based on their understanding of the required material.

II. Method

A. Participants

The participants were 84 undergraduate college students (59 females, 25 males) enrolled in an introductory educational psychology course on a commuter campus of a mid-western university. All students were university students enrolled as education majors.

B. Procedures

Participants took weekly objective tests (true-false and multiple choice) and completed a questionnaire for each test. Part of the questionnaire was completed before the participants took the test and part immediately after taking the test but before scoring the test. Prior to taking the test each student was asked to report the number of hours they had studied, how many points they would have to achieve to be satisfied with their performance (satisfaction goal), how many points they would have to achieve to be proud of their performance (pride goal), and how confident they were about achieving their satisfaction goal (pre-test self-efficacy).

After completing the test, but before it was graded, each student was asked to identify how many points they believed they would achieve on the test[2] and how confident they now were about their achieving their satisfaction goal (post-test self-efficacy). Then, tests were graded and were returned to the student for review before the student was dismissed from class.

Each of the weekly tests included 40 objective test questions: 18 lower level test questions which emphasized knowledge and comprehension and were worth 1 point each; 18 moderately difficult questions which emphasized application and were worth 2 points each; and 4 difficult test questions which emphasized analysis and synthesis and were worth 3 points each. For each weekly test, students were allowed to answer only 30 of the 40 test questions; their grade being dependant on *both* the accuracy of their answers (number of questions correct) and the type of test questions they chose and answered correctly. To earn an A in the class, students had to choose more difficult test questions (worth 2 or 3 points) and get them correct. To earn a lower grade, students could either take more difficult questions and get a lower percentage correct, or take less difficult test questions (worth 1 or 2 points) and get a higher percentage correct. Therefore, the key to success in the course was not only correctly answering test questions, but also choosing the test questions you could answer correctly. The weekly tests were designed to reveal and substantiate student metacognitive awareness during testing. Ten tests were administered during the semester (approximately one per week).

III. Results

This study examines the differences between high and low achieving students on a number of metacognitive variables: What are the long-term changes between their estimations of test grades early to late in the semester?; How are their satisfaction and pride goals different from their actual performance?; How does their self-efficacy change and how does that compare to their actual test score?; and, Are students able to make academic choices based on their MKM?

A. Test Scores Postdiction

Following the format used by Hacker (2000) we examined the hypothesis that high performing students would be more accurate in predicting their test scores than low achieving students. Using Hacker's terminology, our participants made "postdictions" because they took the test before estimating their score on the test. For each of the ten tests, a correlation was computed between the test score and the squared error of the student's postdiction [(test score – expected test score)^2]: let us call this type of correlation the *matched-score format*. Students who have good MKM should be better at postdicting their test scores which would result in smaller squared error scores and a negative correlation between test scores and squared error scores in the *matched-score format*. The average (median) of these ten correlations was -.26 (all listed correlations are significant at $p < .05$ unless noted otherwise). For the three most difficult tests, the correlations were -.27, -.49, and -.65. Thus, the students who were most accurate in their postdictions (having a low squared error of postdiction) tended to have higher test scores and this was particularly true on the most difficult tests (see Table 1).

From the previous analysis (matched-score format) it is possible that all students are

[2] Hacker (2000) makes the distinction between predictions and postdictions. Predictions are a student's estimates of their test scores *before* they take a test. Postdictions are a student's estimates of their test scores *after* they have taken the test but *before* their tests are graded. In the present study the students estimated their results after having taken the test but before the test is graded. The students were not asked to make predictions before taking the test.

equally good (or poor) at postdicting their test scores, they all postdict that they will do well, and that on any particular test a few students do poorly at random. To insure that it was the students who consistently performed well on their tests who were making the best postdictions, we also computed correlations between the total of the ten test scores (total points) and each of the ten squared error of postdictions. For example, for the first correlation, we took the squared error of postdiction for Test 1, and correlated it with the student's total points for the semester. Let us call this type of correlation the *total score format*. This calculation was done for each of the ten tests. The median of these ten correlations was -0.18. For the three most difficult tests, the correlations were -0.27, -0.24, and -0.53. Thus, the students who have the highest achievement across the semester are better at postdicting test results, and their postdiction accuracy is most pronounced on the most difficult tests. This confirms the findings by Hacker (2000) and others (Maki and Berry, 1984; Maki, et.al., 1990; Maki and Serra, 1992) that demonstrates that high performing students are better at metacognitive awareness, knowledge monitoring, and calibrating how they will do on tests in college. Are there also differences in the goals students set and the changes that occur over time during a semester?

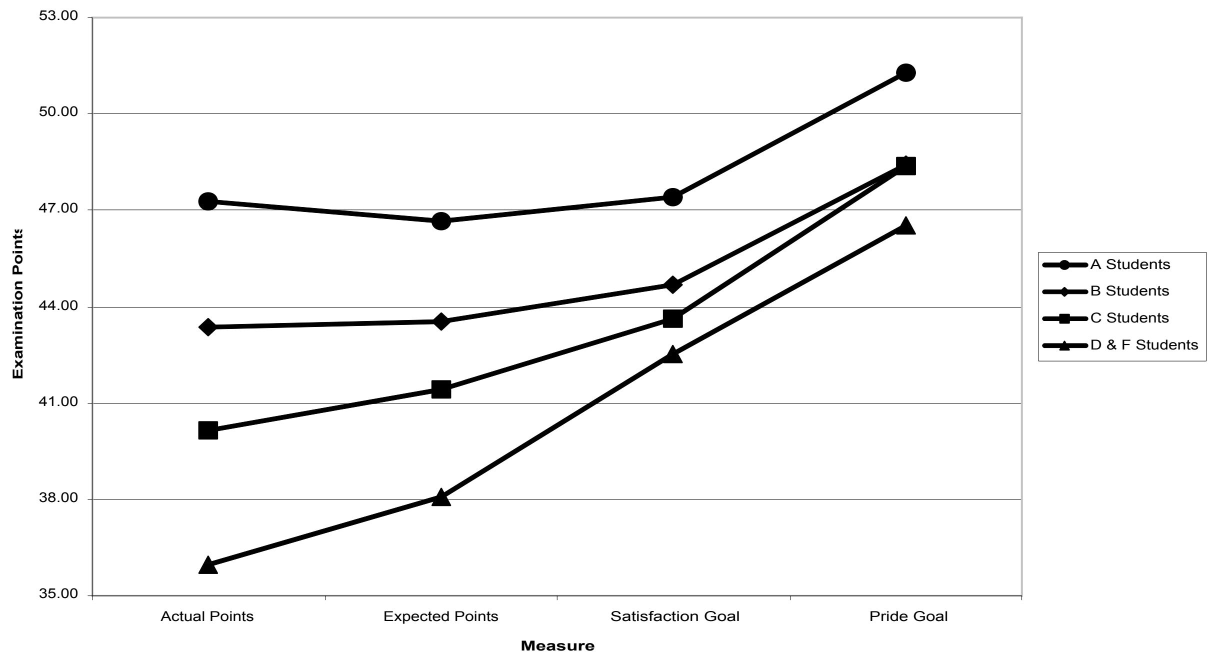
Table 1: Correlations Between Test Score and the Squared Error of Predicted Test Score

			Matched-Score Format			Total-Score Format		
Chapter	Mean	StdDev	r	N	p	r	N	p
2	43.0	5.2	-0.270	81	0.02	-0.228	76	0.05
3	41.2	6.7	-0246	78	0.04	-0.190	77	0.10
4	43.7	5.0	+0.054	81	0.64	-0.014	78	0.91
6	43.5	5.7	-0.392	81	0.00	-0.133	79	0.24
7	38.4	6.8	-0.274	77	0.02	-0.273	75	0.02
8	44.3	6.8	-01.54	81	0.17	-0.180	78	0.11
20	41.7	6.2	-0.163	80	0.15	-0.161	77	0.16
11	43.5	7.0	-0.055	77	0.64	-0.238	76	0.17
14	37.1	6.3	-0.492	80	0.00	-0.527	77	0.04
15	37.1	7.4	-0.651	75	0.00	-0.160	74	0.00
Mean			-0.264			-0.210		
Median			-0.258			-0.185		

B. Group Differences by Goals across Time

Four roughly equally sized groups were created based on the total performance across the ten tests. To examine the differences between high achieving groups and low achieving groups in relation to goal setting, expected performance, and actual performance across the ten tests during the semester, a repeated measures MANOVA was performed. The three factors were performance group (four levels), measure (satisfaction goal points, pride goal points, expected points, actual points), and time (the ten different tests). The important significant difference for this analysis was the group by measure interaction ($F(9,107.2) = 5.76$, $p<.05$, partial $0^2 = .274$). The highest performing group has small differences between goal points, expected points, and actual points. The lowest performing group had very large differences between satisfaction and pride goal points (high), expected points (high), and actual points (low). The two intermediate groups had appropriately intermediate differences between the two extremes. The three-way interaction was not significant. Thus, the lowest performing group of students did not adjust their satisfaction or pride goal points, nor their expected points, to the reality of their actual points. This further supports the hypothesis that low achieving college students are less likely to use metacognitive awareness to make adjustments as they are learning in a college course. The actual test scores of the highest achieving students are very similar to their satisfaction goals and their expected points, with their pride goals

Figure 1: Examination Points as a function of Type of Points and Performance Level of Student



approximately 4 points above their achievement. When the highest achieving students are incorrect in their postdiction they are more likely to *under*-postdict their score which may have led to more extensive studying before the test. As shown in Figure 1, he actual test score of the lowest achieving students are significantly less than their expected points and their goals across the entire semester which may have led to less studying before their test.

Expected points, goals, and self-efficacy can vary dramatically both across individuals and within individuals across the semester. When a high achieving student is able to make accurate postdictions, is that just because the high achieving student lives up to the universal expectation of good performance, or does the high achieving student appropriately lower expectations when they are going to perform at a lower level in a way that the low achieving student does not? That is, in a course which has a weekly test even high achieving students occasionally have a bad day and if the hypothesis about metacognitive awareness is correct these students should make adjustments to their postdictions, self-efficacy, and goals as the semester progresses.

C. Intra-individual Differences across Time

To focus on differences within individual students across the semester, we calculated within-subject correlations of the test score, postdicted test score, hours studied, pre-test self-efficacy, and post-test self-efficacy within each individual student across the ten tests. Each student will have, for example, a correlation between his or her ten test scores and his or her ten postdicted test scores to help explore how changes in test scores compare to changes in postdictions which indicate whether changes in a student's postdictions across the semester reflect changes in their test scores across the semester. The postdictions of students with effective MKM should rise and fall in concert with their test results indicating that they knew when they did well or poorly on test. These individual correlations are characteristics of each student, and as such can be considered a type of individual difference variable. These individual difference variables were themselves correlated across students with each student's final grade. These correlations have a number of important implications for self-regulation.

Intra-individual Post-diction Accuracy across Time. Each student has a within-subject correlation between his or her postdicted points and actual points across the ten tests, which reflects the extent to which each student adjusts their postdictions to match their actual score across the test tests during the semester. Let's call this within-subject correlation "relative postdictive accuracy." The mean relative postdictive accuracy across all students is 0.24. Thus, using the weekly test scores and the student's postdictions for each test, the average student is able to make a somewhat accurate relative postdiction of whether he or she will do better (or worse) on this exam than on the other nine exams. The between-subjects correlation of the relative postdictive accuracy with total points is 0.26 ($p=.056$) meaning that students whose relative postdictive accuracy is greater than the mean tend to score higher across the ten tests during the semester. Students who are more accurate in adjusting their estimation of how well they have done on tests from week to week are more likely to achieve more total points during the semester.

Reliance on Effort for Pre-test Self-Efficacy. Each student has a within-subject correlation between his or her number of hours studied and pre-test self-efficacy. Let's call this within-subject correlation "reliance on effort." The mean reliance on effort is 0.30. The average student is more confident the more hours he or she studied for that test. But the between-subject correlation of reliance on effort with total points is -0.24 ($p=0.12$) meaning that students whose reliance on effort is greater than the mean are less likely to do well in the class. This could be interpreted to mean that the more a student depends on the number of hours they have studied to decide on their confidence for success, rather than on MKM to decide on how confident they are about the test, the less likely they are to do well on a test. It is good to be able to weigh metacognitive feedback during test preparation rather than being forced to depend primarily on the amount of effort expended when making a prediction on how well you'll be able to do on an exam.

Metacognitive Changes in Post-test Self-Efficacy. Each student has a within-subject correlation between his or her pre-test self-efficacy for achieving their satisfaction goal and their post-test self-efficacy for achieving their satisfaction goal across the ten tests. Essentially this measures how much each student is likely to change their self-efficacy for achieving their satisfaction goal from before they take the test to after having taken the test. Let's call this within-subject correlation "self-efficacy constancy." The mean self-efficacy constancy is 0.30. But the between-subjects correlation of self-efficacy constancy with total points is -0.39 ($p<.05$). Thus, students whose self-efficacy constancy is greater than the mean tend to score lower than students whose self-efficacy constancy is less than the mean. Students who use the feedback they receive from taking a test to adjust their self-efficacy are more likely to do better on tests across the semester. This is consistent with what we would expect from students with good metacognitive awareness in that they are aware of how they have done after taking a test but before it is graded.

These three individual difference variable all support the theory that high achieving students are consistently monitoring their understanding of their learning and adjusting their postdictions (relative postdictive accuracy), adjusting the time they spend studying (reliance on effort), and are better judges of how well they have done after completing a test (self-efficacy constancy) than low achieving students. This begs the question: Can metacognitive knowledge monitoring be taught to students and will improvements in MKM lead to improvements in learning?

D. Metacognitive Impact on Choosing Test Questions

Each of the tests in the course used a variability difficulty - variable weight test format where students were given choices about which questions they selected to be graded. This test format allowed students to eliminate questions to which they did not know the answer but at a cost: choosing more difficult questions earns more points but only if you get the correct answer. For each of the 10 exams, students received a number of questions correct out of 30. Each student also had a number of points earned for the exam. Since students are allowed to choose 30 questions from 40 available questions, and since their total points are dependant upon choosing 30 questions that they are likely to get correct, it is critical that students choose the appropriate questions from week to week depending upon their mastery of the material. That is, some weeks students have clearly mastered the material and can choose questions which are more difficult and worth more points. Other weeks when students have not mastered the material as well, the students with good MKM can choose less difficult questions to which they know the answers while students with poor MKM are likely to guess at the answers and guess at which questions to choose to answer. To assess the accuracy of students' choices we correlated the number of points earned with the number of questions answered correctly within student so that each student had a within-subject correlation coefficient that measured the relationship between the number of questions answered correctly and the student’s score on the exam. We will call this within-subject correlation question dependency (QD). If all of the questions on the exam were worth the same number of points, QD would be 1.0 for all students. Because students chose which questions to answer, and different questions were worth differing numbers of points, the average QD was 0.94, the minimum QD was 0.74, and the maximum QD was 0.99. While it is clear that the number of questions answered correctly is the key variable in any student’s score, some students are able to assess their understanding of the material and the test question and choose the appropriate test questions to positively influence their grade.

Students with high metacognitive skill should have a lower QD, because when they know that they do not understand the material well they will choose easier questions to answer, when they understand the material well they will choose more difficult questions which are worth more points, lowering their QD. Students with high metacognitive skills know when they understand the material and are more likely to be able to adjust their choice of test questions based on accurately reflecting upon their learning. Since the test format allows them to not answer some questions they can make their choices based on their assessment of their understanding. Their test points will depend relatively less on the number of questions answered correctly and relatively more on the difficulty of the questions chosen. Students with high metacognitive skill should also, on average, score higher on the exams.

A correlation across students was calculated between each student's average test score and that student's QD. The between-subjects correlation was negative ($r = -0.33$, $p < .01$) as expected. Students with low QD had higher average test scores which is evidence of their metacognitive awareness; they used metacognitive strategies to adjust their test item selection based on their knowledge monitoring. The student whose test scores are more question dependant are most likely to do well based on whether they get questions correct, independent of which questions they choose. Expert college students in the present study adjusted their choice of test questions to enhance their test scores based on knowing-when-they-knew. It can be inferred that low achieving students were less accurate in selecting the questions they knew based on their metacognitive awareness of their understanding of the material and their comprehension of the test questions.

IV. Discussion

The literature on metamemory and metacomprehension has demonstrated that better learners are able to make more accurate judgments about their learning. The classroom application of these findings for self-regulated learning has not been thoroughly explored although the work of Hacker et.al. (2000) begins to shed light on these possibilities. Our findings begin to make a connection between the accuracy of students' metacognitive judgments of their learning and a number of variables related to self-regulation.

The initial step, in Zimmerman's three step academic learning cycle, is forethought which includes goal setting and self-efficacy, the second step is performance and volitional control which is guided by self-monitoring, and the third stage is reflection including assessing success or failure and modifying self-efficacy. To be an effective self-regulated learner a student must use MKM to guide this process and make adjustments in goals, judgments of learning, self-efficacy, and task choice. This study begins to explore these relationships within individual learners.

The first study of metacomprehension within a classroom context (Hacker, et.al., 2000) confirmed the finding that high achieving students are better at predicting and postdicting their learning. Our study supports this finding in relation to student postdictions. Students' postdictions correlated significantly to their test scores. Of special interest was the finding that this correlation was greater for tests that were more difficult which supports the finding of Maki (1995). This relationship has potential implications for metacognition and self-regulation. If difficult tests are better at differentiating high achievers from low achievers the reason may be that difficult tasks require a higher level of metacognition and self-regulation. Given the demands for higher level thinking skills that occur when students transition from high school to college, it would be revealing to explore if effective MKM becomes increasing important in academic tasks that require these thought processes.

The groups in the Hacker et.al. (2000) study were assigned based on their results on each individual test which leaves open the possible interpretation that the reason for the discrepancy between groups could have been a regression to the mean. If the test goals of most students were approximately 80%, and each student were to predict achieving their goal, the group discrepancy between predictions and actual score would be attributable to their initial goals. Since the group assignment changed across each of the three tests it is possible that group membership changed while goals, predictions, and postdictions stayed the same.

In our study the group assignments were made based on the students' accumulated test points across all ten exams. Group assignment based on accumulated points across the entire semester rules out the possible regression to the mean explanation. The twenty students in the top group were the students with the highest scores across all ten tests. The twenty students in the bottom group were the students with the lowest scores. The repeated measures MANOVA demonstrates that the top group of students had the smallest differences between their satisfaction and pride goals, expected points, and actual points across the course of the semester while the lowest group of students had the widest discrepancy between these variables. Future research should examine whether students' MKM changes over time, what factors might impact these metacognitive skills, and whether these skills can be taught to lower achieving students.

The results from Hacker et.al (2000) also suggest that high achieving students are not only more accurate in their judgements but more likely to *under*-estimate their results (leading to under-confidence) while low achieving students were more likely to grossly *over*-estimate their results (resulting in overconfidence.) In these conditions self-regulation theory would predict that high achievers who are under-confident might be defensive-pessimists (Garcia and Pintrich, 1994) which would lead to an increase in their efforts and their success. Students who overestimate their result may be self-handicapping which may result in a decrease in their efforts which would lead to failure. The relationship between calibration, confidence, and self-efficacy will need to be more fully explored since self-protective perceptions may influence the accuracy of calibration (Dembo and Jakubowski, 2003). What is the impact when students underestimate or overestimate their mastery of the course material? How does this impact studying before a test? What impact does this underestimation and overestimation have on motivation to study?

According to Zimmerman (1998), adjusting goals, expectations, and self-efficacy over time is a critical skill in self-regulated learning. The weekly class test format in our study allowed students to set and readjust goals and expectations before each test during the course of the semester. The pattern of pride and satisfaction goals, expected points, and achievement for the high achieving group is consistent with self-regulation, while the pattern for the low achieving students does not reflect the reality of their actual test scores. Low achieving students maintain the same general expectations which may lead them to learned helplessness. Future studies should explore the affective reactions of students who do not adjust their goals, their expectations in the face of continual failure, and their attributions for success and failure.

One of the advantages of the present study is the longitudinal nature of the data set. Since students are given weekly tests and the same data is collected for each test, we can examine the intra-individual differences for a number of variables. Self-regulation involves small changes over time within, as opposed to across, individuals. To explore changes in self-regulation it is important to examine how the changes of one student week-to-week compare to the changes within another students. These intra-individual relationships can be examined as individual difference which can then be compared to variables such as final course grade. Our study explored three such relationships.

The first intra-individual difference we are calling "relative postdiction accuracy." This is a measure of whether a student is able to predict whether they will do better on *this test* relative to other tests. The average student is somewhat accurate in estimating whether they will do better on this test compared to other tests, but the highest achieving students are better at judging when they will do better or worse on a test compared to their own performance across the semester. While teachers at all levels may view this as "common sense", the ability of students to judge how they have performed on this test compared to other tests is indicative of the MKM that is critical to academic success in college. High achieving students are not only better at estimating their score on an individual test, they are also sensitive to whether they will do better or worse than they usually perform on the class tests. Whether the reason is that they were not able to study as much as usual or they estimate that the material will be more difficult, high achieving students are aware that on this test they are likely to perform better or worse than they usually perform. Self-regulated learning in Zimmerman's performance stage (1998) is dependent upon on-line monitoring of performance and when students are not able to assess whether they are doing better or worse than they normally do they are less likely to adjust their behavior.

The second intra-individual difference is "self-efficacy constancy." Students were asked to estimate their confidence for achieving their satisfaction goal before they took the test, and after they took the test but before the test was graded. Students who are metacognitive about the feedback they receive from taking a test adjust their self-efficacy based on this feedback. Students who lack metacognitive awareness do not change their self-efficacy. High achieving students are more likely to modify their confidence for an individual test after having taken the test; "That test was harder/easier than I thought it would be." This is also evidence of the metacognitive awareness necessary for self-regulation. When students do not, or cannot, adjust their self-efficacy after taking a test, it is likely they are not metacognitively aware of how they have done.

The third intra-individual difference is "reliance on effort." Many students link their test confidence (i.e., pre-test self-efficacy) to the number of hours they study for a test. When a student has studied four hours for a test they are usually more confident than when they have only studied one hour. But students who are truly self-regulating do not tie their confidence solely to the number of hours they have studied. Self-regulating students make judgments of when to stop studying based on how well they know the material and whether they believe they will be able to achieve their goals. Many students seem to have a pre-designated number of hours they plan on studying and if they complete those hours they believe they are sufficiently ready for the test. When students rely primarily on time to regulate their learning they may be less likely to succeed. The correlation of "reliance on effort" with total points indicates that the more a student relies on time to determine their pretest self-efficacy the lower their total points in the class. The relationship of metacognitive awareness, pre-test confidence, and self-regulation is an important variable to explore. Many students depend on time as the leading indicator of learning - - - and the result is often failure. How students decide they have spent enough time learning is an important question to be explored.

The process of self-regulation depends on intra-individual changes in expectations, goals, and self-efficacy. Students who are skillful at self-regulation modify what they expect based on the feedback they receive from self-monitoring and external input (e.g., tests). They also adjust their goals and self-efficacy based on these results. The process of learning self-regulation requires extensive time and feedback. The present study begins to follow the development of this process in a college classroom with frequent feedback over an entire semester.

The variable weight - variable difficulty test format in this study places a strong emphasis on student's metacognitive awareness. Given this test format it is possible for a student to get all 30 of the questions they choose correct and still only earn a C, *if* they choose the easiest test questions. Students are told, and quickly learn, that their grade is dependant not only on the number of questions they get correct, but also choosing the correct test questions. This test format encourages students to take questions that are worth more points, which require higher level thinking skills, with the expectation that they will eventually learn to regulate their study time and strategies to go beyond the simple memorization of facts to the application, analysis, and synthesis of information. But learning at this higher level is not the only skill necessary to succeed with this test format. It is also important for students to know-when-they-know. Students who used elaboration and organizational strategies in learning would be more likely to do well if all students were required to take all the questions of highest value. If all students took all the higher level questions, the student who got the most questions correct would have the highest score. But since students can choose which questions to take their score is also dependant upon their ability to choose the question they will get correct.

The question dependency (QD) variable helps to reveal a student's ability, across the semester, to choose questions that will enhance their test score based on their MKM. Students who lack MKM are less likely to be able to assess whether they understand, and can answer, each individual test item. For these students their test score is entirely dependent upon the number of questions they get correct because they do not make metacognitive choices on individual items on the test. For students with high MKM their test score is influenced by the questions they choose as well as the number of items they get correct. For example, a student with high MKM who, in a given week, is not able to devote as much time as usual to studying for the higher level thinking test questions is likely to choose lower level difficulty test questions on that test. These students are aware that they do not know the answer to the more difficult test questions and, instead of guessing, choose the easier questions for which they know the answer. Their MKM influences their total score which makes them less question dependent. But students with poor MKM cannot make these choices because they do not know-when-they-know and therefore are guessing: guessing not only on the answer to the question, but also on whether they should choose the question.

This finding invites further exploration of the relationship of metacognitive awareness and learning particularly when higher level thinking is required. Many undergraduate students have difficulty when they first enter college because they are not familiar with the academic demands of higher level thinking. If they were effective in memorizing information in high school they were likely to have received good grades and also were likely to have believed their were good at learning. This would lead them to believe they would be successful in college if they were to use the same learning strategies. When they enter college and fail to meet their own expectation, they are then more likely to externalize the blame for their failure to the teacher or an unfair task. This may keep them from engaging in the reflection which is essential to changing their study behavior and improving their metacognitive awareness. Are high achieving college students better able to predict their scores on difficult tests because the questions are more difficult (Maki, 1995) or because the test requires higher level thinking skills?

This study also raises the question of whether metacognitive awareness can be improved over time. It is clear that the highest achieving students in this study were better able to predict their test scores and also choose the right test questions. Are the metacognitive skills that allow these students to choose the right test questions a stable characteristic, or can these skills be improved over time? If these skills can be improved, what instructional approaches would facilitate an improvement in their metacognitive awareness that would impact their self-regulated learning and success in school? This study and others (e.g., Hacker et.al., 2000; Tobias and Everson, 2000, 2002; Maki, 1995) are demonstrating a strong relationship between MKM and academic performance but the causal relationship is unclear. Can an improvement in MKM lead to academic improvement or is improved MKM a result of improved academic performance?

This study demonstrates that expert students are effective at estimating their understanding (postdicting their test scores) and they are more inclined to vary their goals and self-efficacy based on past results and the feedback they receive from taking a test. Expert students in this study were also more likely to make choices which demonstrate mastery and non-mastery of tasks of varying levels of difficulty (choosing the appropriate test questions.) This has important implications for the teaching-learning process in higher education and secondary education. As students are required to take on academic tasks of increasing difficulty it is critical that they have the metacognitive skills to assess their mastery of the material on a variety of levels. This metacognitive self-assessment is essential to the application of self-regulated learning.

This study also invites a thorough examination of the relationship of MKM to learning. It seems clear that expert students are skilled in reflecting on their own learning but the origin and nature of these skills is not clear. Does MKM facilitate student learning or does the mastery of a body of knowledge assist students in their judgement of their mastery of the material? Does MKM change over time and can it be taught? If MKM is a skill that can be learned, what pedagogical changes in classroom practice are most likely to encourage students to be more reflective of their own learning? What types of assessment practices are most likely to encourage students to be more metacognitively aware? Are there procedures that can be implemented in most postsecondary classrooms (e.g., frequent evaluation and immediate feedback) that would encourage students to be more metacognitively aware? Can technology be used outside the classroom to assist students to improve their MKM? The author is presently implementing pedagogical approaches that are designed to make students aware of the impact of metacognitive knowledge monitoring and assist them in improving their self-awareness of their learning.

This study also raises questions about how to assess metacognitive knowledge monitoring. This study demonstrates the challenges inherent in assessing the metacognitive awareness of students while they are taking a test which reveals the even more difficult task of assessing metacognitive decision making while students are studying - which clearly is the more crucial connection between metacognitive knowledge monitoring and self-regulated learning. This study clearly demonstrates that low achieving students frequently over-estimate how well they know a body of information which leads to disengagement early during studying. This process typically results in failing the test which often leads them to blame the teacher/test for their failure instead of examining their own learning. A thorough examination of the impact of metacognition, and pedagogical approaches that might increase metacognition, are an important issue that should be addressed in the scholarship of teaching and learning.

References

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

Dembo, M., T.G. Jakubowski (2003, April). "The influence of self-protective perceptions on the accuracy of test predictions." Presented at the Annual conference of the American Educational Research Association, Chicago, IL.

Garcia, T., P. R. Pintrich (1994). "Regulating motivation and cognition in the classroom: The role of self-schemas and self-regulatory strategies." In D.H. Schunk & B.J. Zimmerman (Eds.), ***Self-regulation of learning and performance: Issues and educational application.*** (pp. 127-153). Hillsdale, NJ: Erlbaum

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.

Kelemen, W. L. (2000). “Metamemory cues and monitoring accuracy: Judging what you know and what you will know”. ***Journal of Educational Psychology**, 92*(4), 800-810.

Koriat, A. (1997). “Monitoring one's own knowledge during study: A cue-utilization approach to judgements of learning.” ***Journal of Educational Psychology**, 126*(4), 349-370.

Locke, E. A., & Latham, G. P. (1990). ***A theory of goal setting & task performance**.* Englewood Cliffs, NJ: Prentice-Hall.

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

Maki, R. H., & Berry, S. (1984). Metacomprehension of text material.” ***Journal of Experimental Psychology: Learning, Memory, and Cognition**, 10*(4), 663-679.

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

Maki, R. H., & M. Serra. (1992). “The basis of test prediction for text material.” ***Journal of experimental psychology: Learning, memory, and cognition**., 18*(1), 116 - 126.

Maki, R. H., & Serra, M. (1992). The basis of test prediction for text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(1), 116-126.

Nelson, T., & Narens, L. (1990). “Metamemory: A theoretical framework and some new findings.” In G. Bower (Ed.), ***The psychology of learning and motivation***. San Diego, CA: Academic Press.

Nelson, T. O., & Leonesio, R. J. (1988). “Allocation of self-paced study time and the "labor-in-vain" effect.” ***Journal of Experimental Psychology**, 14*(4), 676-686.

Pintrich, P. R., Walters, C., & Baxter, G. P. (2000). “Assessing metacognition and self-regulated learning.” In G. Schraw and J. C. Impara (Ed.), ***Issues in the measurement of metacognition*** (pp. 43 - 97). Lincoln, NE: Buros Institute of Mental Measurement.

Pintrich, P., R. (1995). “Understanding self-regulated learning”. In P. R. Pintrich (Ed.), ***Understanding self-regulated learning*** (pp. 3 - 12). San Francisco, CA: Jossey-Bass.

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

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

Sinkavich, F. J. (1995). "Performance and metamemory: Do students know what they don't know?" ***Journal of Instructional psychology, 22***(1), 77.

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.

Weaver, C. A., III. (1990). "Constraining factors in calibration of comprehension." ***Journal of Experimental Psychology: Learning, Memory, and Cognition,*** *16*(2), 214-221.

Zimmerman, B. J. (1998). "Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models." In D. Schunk & B. Zimmerman (Eds.), ***Self-regulated learning: From teaching to self-reflective practice*** (pp. 1 - 19). New York, NY: Guilford.