

Metacognitive awareness and academic achievement in college students

Andria Young and Jane D. Fry¹

Abstract: The researchers examined the Metacognitive Awareness Inventory (MAI) (Schraw and Dennison, 1994) to determine how it relates to broad and single measures of academic achievement in college students. Correlations were found between the MAI and cumulative GPA as well as end of course grades. Scores on the MAI significantly differ between graduate and undergraduate students. Professors' use of the MAI as a potential screening tool to identify students requiring metacognitive strategy intervention is discussed as well as implications for future research.

Keywords: metacognition, metacognitive awareness, GPA, grade.

College professors today are faced with classrooms full of students who come to them with varying levels of knowledge about how they learn. Some students are active, self directed learners who know how they learn and are able to apply what they know to various learning situations. Others may be average students who work hard and who have awareness of their learning strengths and weaknesses, but who may not adequately regulate their learning. Still others may be passive learners who have little awareness of how they learn and how to regulate their learning. In essence, professors are faced with classrooms full of students who come to them with various levels of metacognitive skills.

Metacognition is generally defined as the activity of monitoring and controlling one's cognition. It can further be defined as what we know about our cognitive processes and how we use these processes in order to learn and remember (Ormrod, 2004). Researchers further conceptualize metacognition by breaking down metacognition into two subcomponents, metacognitive knowledge and metacognitive regulation. These two subcomponents have been theorized to be related to one another (Brown, 1987; Flavell, 1987; Schraw and Dennison, 1994).

Metacognitive knowledge can be described as what we know about our own cognitive processes. Declarative, procedural and conditional knowledge may all be considered subcomponents of metacognitive knowledge (Schraw and Moshman, 1995). Declarative knowledge involves what we know about how we learn and what influences how we learn. Procedural knowledge is our knowledge about different learning and memory strategies/procedures that work best for us. Conditional knowledge is the knowledge we have about the conditions under which we can implement various cognitive strategies. As a whole, our knowledge of cognition refers to what we know about how we learn; what we know about the procedures and strategies that are the most effective for us; and, what we know about the conditions under which various cognitive activities are most effective (Schraw and Moshman, 1995).

¹ School of Education and Human Development, University of Houston –Victoria, 3700 N. Ben Wilson St., Victoria, TX 77901, younga@uhv.edu.

Metacognitive regulation in contrast to metacognitive knowledge may be thought of as the actual activities in which we engage in order to facilitate learning and memory (Schraw and Moshman, 1995). Metacognitive regulation can be broken down into three component activities. These include planning, monitoring and evaluating. Planning involves just that, planning out a cognitive task by selecting appropriate strategies and cognitive resources. Monitoring involves the awareness of our progress through a cognitive task and our ability to determine our performance. Finally, evaluating involves taking a look at the outcome and determining if the learning outcome matches our learning goals and if the regulation processes we used were effective (Schraw and Moshman, 1995).

It stands to reason that if students have well developed metacognitive knowledge and metacognitive regulatory skills and they use their metacognition they will excel academically. Consequently, it is important to be able to assess metacognition of college students to determine if this knowledge and skills are related to academic achievement. If we can say that metacognitive knowledge and skills are related to measures of academic success then professors can use various techniques to assess their students' metacognition and develop means by which to improve students' metacognition when necessary.

I. Metacognitive assessment and academic achievement.

Researchers have examined metacognition and how it relates to measures of academic achievement. In these studies metacognitive skills are measured in terms of metacognitive regulation, metacognitive knowledge or both of these components. However, these components are measured differently within the literature. Some researchers use self report inventories to assess metacognitive skills and relate them to achievement measures (Schraw and Dennison, 1994; Sperling et al., 2004). Other researchers examine metacognitive judgments in the form of monitoring accuracy as a measure of metacognitive regulation on various tests (Everson and Tobias, 1998; Nietfeld et al., 2005; Schraw, 1994). Monitoring accuracy is measured in terms of what is considered calibration of performance. Calibration of performance judgments are made at the local and global levels. Local judgments are made after each item on a test. Local monitoring accuracy is determined to be the average difference between the actual answer of each test question and the students' judgment of how well they answered each question. Global judgments are made after the entire test is completed. Students are to judge how well they think they did on the test as a whole. Global monitoring accuracy is determined to be the difference between the overall test score and the students' judgment of how they did on the test. Local monitoring accuracy is thought to be a measure of ongoing metacognitive regulation during testing and global monitoring accuracy is thought to be a measure of cumulative metacognitive regulation (Nietfeld, et al 2005). The following is a brief review of studies utilizing both survey and measures of monitoring accuracy to assess metacognitive knowledge and/or metacognitive regulation.

Everson and Tobias (1998) were interested in knowledge monitoring accuracy. This skill is thought to be involved in metacognitive regulation. They developed a means to assess students' knowledge monitoring ability (KMA) by examining the difference between students' estimates of their knowledge in the verbal domain and their actual knowledge as determined by performance on a standardized verbal test. They found the greatest relationship to be between the KMA and students' end of course grade in English, then the humanities and the students' overall

GPA. They also found that this measure of metacognitive regulation, the KMA, was related to academic achievement in college and it was a good predictor for success in college.

Schraw (1994) was interested in the relationship between metacognitive knowledge and metacognitive regulation. He measured metacognitive knowledge by asking students to rate how well they thought they could monitor their accuracy on a series of multiple choice reading tests. He measured metacognitive regulation at both the local and global levels by having students rate accuracy for each question then rate their accuracy after completing the tests. Based on the results of his study, Schraw suggested that adult students may differ not so much in their metacognitive knowledge skills but in their metacognitive regulation skills. He further suggested that metacognitive knowledge may develop independently of metacognitive regulation. Finally, Schraw found that actual test performance was significantly correlated with judgments of test performance made before testing, a measure of metacognitive knowledge. Test performance was also correlated with metacognitive regulation in that he found correlations between performance and local and global judgments.

Nietfeld et al (2005) examined metacognitive regulation by measuring monitoring accuracy at the local and global level on a series of multiple choice tests given as a part of a semester long course. They found that monitoring accuracy remained stable across tests throughout the semester. They also found that students were more accurate in their global predictions than their local predictions. They found that student performance on the tests was related to local monitoring accuracy.

Schraw and Dennison (1994) developed the Metacognitive Awareness Inventory (MAI) to assess metacognitive knowledge and metacognitive regulation which they referred to as the knowledge of cognition factor and the regulation of cognition factor. The MAI consists of 52 questions tapping into these two components of metacognition. They found that there was strong support for the knowledge of cognition and regulation of cognition components and that these two components were related as had been suggested in the research (Brown, 1987).

Schraw and Dennison (1994) also tested the convergent validity of the MAI by comparing MAI scores with other measures thought to be related to metacognition such as pretest monitoring ability, actual test performance and the ability to accurately monitor test performance. They did not find a significant relationship with regard to monitoring accuracy and the MAI or between pretest judgments and monitoring accuracy. They found the knowledge of cognition factor of the MAI was related to higher test performance and the regulation of cognition factor of the MAI was not. They also found that knowledge of cognition as measured by pretest judgments was related to the MAI. Pretest judgments were also related positively to test performance.

Sperling et al (2004) utilizing the MAI to determine college student metacognitive awareness, found a significant correlation between the knowledge of cognition factor and the regulation of cognition factor. They also were interested in whether the MAI would be correlated with other measures of academic achievement such as SAT scores and high school average. They found no relation between scores on the MAI and measures of academic achievement. They were surprised to find a negative correlation between SAT math scores and the MAI scores.

Overall, the findings in the research reviewed above regarding the correlation of metacognition with academic and achievement measures indicate that when regulation of cognition is measured by having students estimate their performance on either a local or global

level, regulation of cognition is related to test performance, domain specific GPA scores and overall GPA scores (Everson and Tobias, 1998; Nietfeld et al, 2005; Schraw, 1994).

It appears that when metacognition is assessed through calibration of performance measures there is support for the relationship between metacognitive skills and measures of academic achievement. Unfortunately, determining monitoring ability and monitoring accuracy at the local and global level to assess metacognitive knowledge and regulation skills is a labor intensive endeavor. This situation is especially true for students who are assessed in their actual college classes and not a laboratory or contrived setting. Students monitoring their accuracy on a local and global level must take the time to answer the test questions and then respond to how confident they were about their performance on each question. This process can be a time consuming and possibly stressful task for students while taking tests that will count toward their end of course grades (Nietfeld, 2005). It is important to assess students in a less intrusive manner in order to ascertain their metacognitive awareness and skill level. A less intrusive assessment such as a questionnaire, will allow instructors to quickly identify struggling students early on and assist them in developing effective metacognitive skills.

In a departure from utilizing metacognitive judgments as a method to determine metacognitive skills Schraw and Dennison (1994) developed the MAI as a quick and easy means to assess metacognitive awareness. As reported above they found the MAI correlated with reading comprehension test performance, a measure of academic achievement, only on the knowledge of cognition factor. Sperling et al (2004) did not find a correlation with more comprehensive measures of academic achievement such as SAT scores or high school GPA. Obviously the results of the studies in which the MAI was used to assess metacognition are mixed.

The MAI, needs to be examined further and in a broader context. Instruments used to assess metacognition must be sensitive to comprehensive measures of academic achievement that require a variety of cognitive skills in addition to general verbal ability. Assessments must be easy to administer and score so professors can use the information to help students over the course of a semester. Metacognitive assessments must also be comprehensive assessments of the theorized components of metacognition, namely metacognitive knowledge and metacognitive regulation.

The purpose of the present study was to further examine the relation between metacognition and broad based measures of academic achievement within a natural classroom setting. The MAI was chosen because it is an easy to administer survey for adults, which can be delivered in both face to face and online classes. Additionally, the MAI taps into the two component model of metacognition, metacognitive knowledge and metacognitive regulation cited in the research (Brown, 1987; Schraw and Dennison, 1994). Furthermore, with the MAI researchers can analyze relationships between metacognitive skills and specific academic skills such as scores on classroom tests, reading comprehension tests etc. Researchers can use the MAI to analyze for relationships between metacognitive skills and broader measures of academic achievement such as cumulative GPA, SAT scores and other standardized scores. The final purpose of the study is to add to the body of knowledge regarding the validity of the MAI in terms of the statistical relationship between metacognitive knowledge and regulation; and in terms of the convergent validity of the MAI with measures of academic achievement.

Given the findings thus far regarding the MAI and academic achievement measures the primary goal of this study was exploratory in nature. The researchers were interested in correlations between the MAI and end of course grades; the MAI and cumulative GPA; and the

MAI and single tests within a semester long course. Furthermore, the researchers were interested in whether scores on the MAI would distinguish between experienced and less experienced students as measured by class standing as graduate or undergraduate.

II. Method.

A. Participants.

Undergraduate and graduate education students at a small upper level (junior, senior and graduate level) institution located in Southeast Texas were invited to take part in the study. Junior and senior level students in undergraduate teacher education classes in Reading and Human Learning were asked to voluntarily complete the Metacognitive Awareness Inventory (MAI) during summer and fall semesters. Additionally, graduate students in master's education programs taking core classes were asked to voluntarily complete the MAI during summer and fall semesters. The MAI was offered in a total of 15 classes. Two classes were delivered face to face, while the remaining classes were delivered online. The requirements of the fifteen classes from which students were drawn included multiple choice tests, online discussions with specific content criteria, projects and lesson plans.

Students in both face to face and online classes had access to the MAI online through WebCT. They were told they could take the MAI at any point during the semester in which they were enrolled in the participating class. One hundred and seventy eight students completed the MAI. Forty five or 25.3% were graduate students and 133 or 74.7% were undergraduate students. One hundred and fifty eight or 88.8% of the respondents were enrolled in online classes. The remainder was enrolled in face to face classes with access to an online component of the class. See Table 1 for student characteristics.

Table 1. Student Characteristics.

Gender	Age	Level of College Education	Credit Hours Taken of Semester Participation	GPA
Male 3.9%	20-25 yrs 32.8 %	<2yrs. 1.7%	3-6 35%	4.0-3.5 46.9%
Female 96.1%	26-30 yrs 22.6%	2yrs. 17.4%	7-12 36.3%	3.49-3.0 32.8%
	31-35 yrs 15.3%	3yrs. 20.8%	13-18 17.5%	2.99-2.5 15.3%
	36-40 yrs 5.6%	>3yrs. 34.3%	19-21 2.1%	2.49-2.0 5.1%
	41-45 yrs 14.7%	Bachelors 22.5%	>21 .9%	
	Over 45 9.0%	Masters 3.4%		

B. Materials.

The MAI (Schraw and Dennison, 1994) with permission of the first author was used to measure students' metacognitive awareness. The MAI consists of 52 statements which students rate as being false or true on a five point likert scale. The two components of metacognition discussed above are represented within the scale, metacognitive knowledge and metacognitive regulation. Within the MAI these are referred to as the knowledge of cognition factor and the regulation of cognition factor. Within the inventory there are 17 questions related to the knowledge of cognition factor for a possible point total of 85. There are 35 questions related to the regulation of cognition factor for a possible point total of 175. The factor scores are calculated by adding the scores on questions related to each of the factors. Higher scores correspond to greater metacognitive knowledge and greater metacognitive regulation. In addition to the knowledge of cognition score and the regulation of cognition score a MAI total score is derived by summing responses to all 52 questions. The instrument was designed for use on adult populations. The MAI was transformed into a web format so it could be completed by students online.

C. Procedure.

In each of the fifteen classes a link to the MAI was set up on WebCT for students to access. Students in face to face classes had an online component to their classes so they too had access to the MAI via WebCT. The letter within the link explained the MAI and the purpose of the study. Students were asked to consent to complete the MAI and to provide their names on the MAI so their end of course grades could be associated with their score on the MAI. Students were not provided incentive in the form of additional points to complete the MAI as this extra credit would skew their end of course grades and confound the results of the study. Consequently, each class had approximately a 50% response rate. Students were told they could complete the MAI at anytime during the semester. The experimenters downloaded the MAI responses only after the end of course grades were submitted to the registrar in order to avoid bias in assigning end of course grades.

III. Results.

A. Correlations between MAI and measures of academic achievement.

For the 178 respondents the mean MAI score was 206.85. The mean score for the knowledge of cognition factor and regulation of cognition factor was 68.69 and 138.16 respectively. See Table 2 for means and standard deviations of the MAI.

Table 2. Means and Standard Deviations of the MAI.

	M	SD
MAI Total	206.85	20.99
Knowledge of Cognition Factor	68.69	7.28
Regulation of Cognition Factor	138.16	14.94

(n=178)

In order to determine if there were relationships between the knowledge of cognition and the regulation of cognition factors, as well as correlations between scores on the MAI and achievement measures of GPA and end of course grades Spearman's Rho, nonparametric correlation analysis was completed. There was a significant correlation between the knowledge of cognition factor and the regulation of cognition factor $r = .73, p < 0.01$. See Table 3.

Table 3. Correlations between MAI scores and broad measures of Achievement.

	Course Grade	GPA	MAI Total	Regulation Factor	Knowledge Factor
Grade	1.00	0.36**	0.19*	0.19*	0.20**
GPA	0.36**	1.00	0.23**	0.20*	0.26*
MAI Total	0.19*	0.23**	1.00	0.97**	0.87**
Knowledge Factor	0.20**	0.26**	0.86**	0.73**	1.00
Regulation Factor	0.19*	0.20**	0.97**	1.00	0.73**

** Correlation is significant at the 0.01 level

*Correlation is significant at the 0.05 level

There was a correlation between the total score of the MAI and end of course grades. Breaking this down into the two factors of knowledge of cognition and regulation of cognition a correlation was found between each of these factors and end of course grades. There was a correlation between GPA and the knowledge of cognition factor and also between GPA and regulation of cognition factor. See Table 3. Albeit relatively modest correlations, these results show a relation between both the knowledge of cognition and regulation of cognition factors of the MAI and broad measures of academic achievement.

A subset of 65 students from the current sample was drawn in order to examine whether the MAI would correlate to single tests within semester long courses. Subjects from four undergraduate sections of an online course titled, "Human Learning and its Application to Education" were drawn for this purpose. This sample was chosen from the larger sample as it contained the largest number of students within the sample taking the same course; each section had the same tests; and course delivery was consistent across sections. Three multiple choice tests were given throughout the semester. Multiple choice questions covered course content and required students to know facts and be able to synthesize and apply information. Spearman's Rho nonparametric correlation analysis was completed on the data. There were no significant correlations between test 1 and scores on the MAI; nor were there significant correlations between test 2 and scores on the MAI. There was a correlation between test 3 and the knowledge factor $r = 0.26, p < 0.05$ and the regulation factor $r = .27, p < 0.05$ of the MAI.

B. MAI scores and individual differences.

The researchers were also interested in whether there were differences in scores on the MAI between more experienced graduate student learners and less experienced undergraduate student learners. Consequently analysis of variance (ANOVA) was performed. The independent variable was whether a student was a graduate or undergraduate student and the dependent variables were the regulation of cognition and knowledge of cognition factor scores. There was not a significant difference between the two groups on the knowledge of cognition factor. There

was a difference between graduates and undergraduates with regard to the regulation of cognition factor $f(1,177) = 4.13, p < 0.05$. The mean score on the regulation of cognition factor for graduates was 142.04 and 136.85 for undergraduates.

IV. Discussion.

The purpose of the present study was to further explore the MAI and its relationship to broad and single measures of academic achievement. As was expected and found in previous research (Schraw and Dennison, 1994; Sperling et al, 2004) there was a significant correlation between the regulation of cognition factor and the knowledge of cognition factor. Significant correlations were found between the MAI and broad measures of academic achievement. The knowledge of cognition factor of the MAI was correlated with GPA and end of course grades. The same is true for the regulation of cognition factor. These results also provide support for the validity of the MAI as it relates to academic measures.

Within this study significant differences were found between graduate and undergraduate students with regard to their scores on the regulation of cognition factor of the MAI but not the knowledge of cognition factor. This supports the authors' contention that if the MAI is a good measure of academic achievement then it should yield scores that distinguish between more and less experienced students. Graduate and undergraduate students do not differ in relation to knowledge of cognition, they do differ in terms of their regulatory skills. This finding supports that of Schraw (1994) who found that adult learners tend to differ with regard to the use of metacognitive regulatory skills and not so with regard to metacognitive knowledge skills.

The results of the correlations between the MAI and single test scores within a course were unexpected. In the current study, the MAI is better correlated to broad measures of academic achievement such as GPA and end of course grades rather than single measures. It seems there may be other factors that confound the relation between the MAI and single test performance. Single test performance grades may be impacted by many variables other than one's utilization of metacognitive regulation and knowledge skills. These confounding factors may be physical illness, variations in personal motivation, and, potential problems with the technology required for the online class. Broad measures such as GPA and end of course grades which are measures of academic performance over time are much less sensitive to these vagaries of everyday life. This possibility is one that warrants further research to determine how factors other than an individual's metacognitive abilities temper learning as measured by single test scores.

The results of this study are promising. Given the positive correlations between the MAI and end of course grades as well as GPA it can be a tool for professors to use to screen students in need of direct instruction related to metacognition. This may become especially important in large classes as well as online classes where professors have little opportunity to get to know their students on an individual basis. Professors can flag students who obtain low scores on the MAI and then use the MAI as a means to determine what type of metacognitive knowledge and regulatory skills the student reportedly utilizes while learning.

The MAI is set up so professors can complete an item analysis for low scoring students. Each of the 52 items within the MAI is a statement about one's knowledge of learning or the activities one must undertake to regulate learning. For example, "I understand my intellectual strengths and weaknesses." and "I have control over how well I learn." (Schraw and Dennison, 1994, p. 473) are examples of questions related to metacognitive knowledge. "I pace myself

while learning in order to have enough time.” and “I set specific goals before I begin a task.” (Schraw and Dennison, 1994, p. 473) are examples of questions related to metacognitive regulation. Professors can examine responses to statements like these and specifically pinpoint areas students are reporting weaknesses. Professors can then tailor instructional intervention related to metacognitive knowledge and regulation to meet the needs of individual students. The use of the MAI as a screening tool and a tool to identify specific metacognitive weaknesses merits further research.

V. Future Research.

In the future the goal will be to further examine the relation between the MAI and measures of academic achievement with larger, random samples of students. The intent behind this is to determine if more robust correlations can be obtained when sampling is random and sample sizes are larger. In addition, future research will focus on using the MAI in the applied setting of a classroom to identify and assist students in developing their metacognitive skills. The MAI will be administered to two sections of the same class, one class of students will serve as the control and the other class of students who score low on the MAI will be identified for metacognitive instruction. Analysis will be completed to determine if students with similar MAI scores in the two classes differ on course related achievement measures when the experimental class members with low scores receives metacognitive instruction and the control class members with low scores receives no metacognitive instruction.

VI. Limitations.

Students participating in the study may not be representative of all adult learners as they were primarily education majors. Additionally, students self selected for the study by volunteering to participate. Thus, higher performing students may have been overrepresented in the sample. Finally, associations between various measures within this study may be confounded by additional variables that were not measured such as motivation, students’ personal time constraints etc.

References

- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Weinert and R. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.
- Everson, H. T. and Tobias, S. (1998). The ability to estimate knowledge and performance in college: A metacognitive analysis. *Instructional Science*, 26, 65-79.
- Flavell, J.H. (1987). Speculations about the nature and development of metacognition. In F. Weinert and R. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 21-29). Hillsdale, NJ: Erlbaum.

Hacker, D. J., Bol, L. Horgan, D.D. and Rakow, E.A. (2000). Test prediction and performance in classroom context. *Journal of Educational Psychology*, 92(1), 160-170.

Nietfeld, J. L., Cao, L. and Osborne J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental Education*, 74(1), 7-28.

Ormrod, J.E. (2004). *Human Learning*. Upper Saddle River, NJ: Pearson Prentice Hall.

Schraw, G. (1994). The effect of knowledge on local and global monitoring. *Contemporary Educational Psychology*, 19, 143-154.

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

Schraw, G. and Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.

Sperling, R. A., Howard, B. C., Staley, R. and DuBois, N. (2004). *Educational Research and Evaluation*, 10(2), 117-139.

Questions first: Introducing critical thinking using the Text Analysis Matrix (TAM)

J. Gregory Keller¹

Abstract: Critical thinking skills are crucial for both academic and everyday life. This paper presents the author's Text Analysis Matrix (TAM), a model for developing skills for the critical examination of texts. The TAM guidelines involve finding and clarifying the main claims of a text, discovering and assessing arguments, uncovering the implications for thinking and for action of the claims made in a text, and, finally, looking at the social and political impact of a text in the form of a critique. Results of the application of the TAM with an introductory Ethics course indicated more sophisticated arguments and more awareness of the broad implications of one's beliefs, as well as a greater awareness of social and political critique, at the end of the semester than at the beginning.

Keywords: Critical thinking, critique, critical reading, philosophy, teaching practice

The following ideas arose in response to an apparently simple question: How does one think philosophically? If, as I have been doing for a number of years, I set out to teach philosophy, especially to those with no prior acquaintance with the subject, it seems reasonable to ask what exactly I am trying to accomplish. When I first addressed the question, however, I found it remarkably intractable. What am I trying to teach? I often, and with some justification, have claimed that I cannot specifically put into words what I am trying to teach because it involves not only a body of knowledge and sets of skills but also a certain level of intuitive or artistic expertise that can be coached more easily than described. One can fall back on the following analogy: if someone wants to learn to play the piano or to play tennis, the most competent teachers will not make a list of items to memorize or even of skill sets only but will guide the student through a series of actions that, based upon the coach's refined mentoring skills, will provide on-going feedback until, seemingly miraculously, one day the diligent student will be performing actions that she could not have dreamed of mastering when she began.

I will admit to finding this analogy comforting when I could not at first define what I want students to learn. The more I thought about the situation, however, the less convinced I was that (1) the good coach cannot describe the results she seeks, even when such a description cannot itself produce those results, and (2) I was such a good coach that I could merely point to my consistent results without further contemplation of my aims. Furthermore, as a philosopher I can hardly claim that thinking about my pedagogic aims is either useless or too taxing. So I began to wonder and work, slowly at first, at clarifying what exactly I want to accomplish in teaching philosophy. In the end I produced a Text Analysis Matrix of sets of questions under four categories that provides me with a starting point for describing to myself and my students what we are undertaking together (see Table 1).

¹Department of Philosophy, Indiana University-Purdue University Indianapolis, 425 University Blvd., CA 331A Indianapolis, IN 46202 jgkeller@iupui.edu.

This undertaking, I came to realize, is not applicable only to philosophy. What I describe falls in the general category of critical, analytical, or one might say merely competent thinking. As discussed below, my recent research with an introductory Ethics class indicates that the TAM leads to more sophisticated arguments and more awareness of the broad implications of one's beliefs, as well as a greater awareness of social and political critique, among introductory level students. The mastering of text analysis lies at the heart of what I intend to teach in addressing with students the various texts of philosophy. Each semester I point out to students the important place in their philosophical education of the particular texts we read and of the interrelated sets of concepts we address. Yet since I hope that students will take classroom lessons both into the larger world of philosophic inquiry and into that other large world of everyday life, the critical thinking skills addressed by the TAM offer one kind of expertise that provides crucial competence in both these realms.

I. Literature Review.

One could describe the questions I raise and the approach I take as part of the long-standing tradition of critical thinking. A voluminous literature concerning this topic already exists. For example, Wolcott and Lynch (2001) provide "Steps to Better Thinking."² Based upon a foundation of "Knowledge and Skills" the steps are as follows:

- Step 1—identify the problem, relevant information, and uncertainties;
- Step 2—explore interpretations and connections;
- Step 3—prioritize alternatives and communicate conclusions; and
- Step 4—integrate, monitor, and refine strategies for re-addressing the problem. (p. 2)

Each of these steps then is divided into sub-steps to be used in the classroom.

In *Learning to Think: Disciplinary Perspectives* (2002), Janet Donald presents a variety of approaches to thinking based on the work of different academic disciplines. She provides a "working model of thinking processes in higher education" (pp. 26-27) in which she describes a perspective on different "methods of inquiry." The working model offers an over-arching set of procedures followed by most disciplines that includes: description, selection, representation, inference, synthesis, and verification. Under each topic on this inventory, then, she lists sub-topics indicating their relation to the diverse inquiry methods she previously connected to academic areas. In a later summary of the "most important thinking processes used generally across disciplines," she lists separately "Identify the context" and "State assumptions," then reduces her previous set of common procedures to those of selection, representation, and synthesis (p. 284).

In Maclellan and Soden (2001), as summarized in Heron (2006) we see the following set of critical thinking skills:

1. unpacking concepts—ability to unpack or break down ideas, concepts or theories;
2. recognizing contradictions—differentiating between viewpoints and counterarguments;
3. development—explaining a phenomenon, joining ideas together to form lines of arguments;
4. providing evidence—supporting or justifying assertions;

²A version of these steps and much additional information can be found at <http://www.WolcottLynch.com>.

5. examining implications of evidence—generating hypotheses about consequences or examining the relationships between key factors;
6. alternative interpretation—questioning or challenging an interpretation of the evidence and offering an alternative. (pp. 212-213)

Heron goes on to say that “Adopting this type of framework risks simplifying a complex area. Moreover, it is important to recognise that critical thinking cannot easily be separated from other forms of thinking” (p. 213).

Pithers and Soden (2000) mention various educational “myths” that limit the learning of critical thinking, and say, “By rejecting the myths, it is possible to consider notions such as looking for novel approaches, the notion that the ‘truth’ may be fluid and context dependent and that the learner needs to develop more control and independence over their own learning” (p. 243).

Christine Loh (2003) suggests a method for improving dialogue that consists of the following practices:

- Listen to understand, find meaning and reach agreement
- Identify and re-evaluate assumptions
- Re-examine all positions raised
- Admit that others’ thinking can improve one’s own
- Search for strengths and value in other positions
- Discover new options, not closure. (p. 13)

We need further to note a philosophical argument by Michael Huemer (2005) that in general one is better served by believing experts or, when experts disagree substantially, remaining skeptical than by engaging in critical thinking about public issues. The reason for this stance has to do with the danger that one might substitute conclusions based upon one’s own ill-informed critical thinking for the informed consensus arrived at by those who have better information and a history of critical analysis in a particular field of study. Of course, Huemer goes on to say, one should use critical thinking to assess who counts as an expert and in personal cases in which it is the case both that one has access to the relevant information and that no expert opinion applies.

In summary, many approaches are available that offer methods for thinking more critically. Each of the suggestions surveyed above points to the importance of a structured system for developing student analytical skills. Significant cautions about the use of critical thinking methods are also mentioned by Heron, Pithers and Soden, and Huemer. While each of the approaches described briefly above offers unique and helpful ideas and techniques, there are some significant differences between these ways of looking at the topic and the Text Analysis Matrix I have developed. The questions raised in the Text Analysis Matrix extend far beyond the critical thinking ideas mentioned here. Not only do the specific questions raised in the TAM offer a sharpened approach to texts and ideas, but the last two categories of the TAM go well beyond what is usually addressed in the literature on critical thinking. The study both of the implications of a claim, in thought and in action, and of possible social and political critique lends a larger and extremely valuable perspective to student work. It opens the way for the broadening of individual thinking and for the preparation of a concerned and capable citizenry. In the next section, I have identified in footnotes specific connections between the approaches to critical thinking discussed above and the use of the TAM.

Table 1. Text Analysis Matrix (TAM).

TOPIC	QUESTION 1	QUESTION 2	QUESTION 3	QUESTION 4
<i>Claims</i>	What's the main idea, thesis, or <i>conclusion</i> ? (Begin by listing important statements in the text, and then ask what idea is supported by most of the rest of the text.)	Is the main idea stated in one T/F sentence? (If not, state it in one T/F sentence. If this seems difficult, perhaps what has been found is a theme or issue instead.)	Is the main idea clear or can I make it clear? (Try to make it clear – define terms as specifically as possible and seek a coherent understanding of its meaning.)	(1) What <i>interpretations</i> of the claim (implications/critique) seem possible or reasonable to me? (2) What interpretations seem unlikely, impossible, or unreasonable? What are the differences between 1 and 2?
<i>Arguments</i>	What claims support the main claim? (This process can continue – the argument might contain claims that support its supporting statements.)	How do supporting claims (<i>premises</i>) work together to provide an 'argument' for the thesis (<i>conclusion</i>)?	Does the argument make the conclusion necessarily true, likely (or even highly likely), or does the argument really not support the conclusion at all?	Can the argument be made stronger or clearer? On the other hand, what objections or counterarguments might be raised?
<i>Implications</i>	What ideas or actions might follow from the main idea?	What argument might be given that these implications truly follow from the main claim?	What objections might be raised to the implications of the main idea? How might the claim be defended against these objections?	What implications might follow from the claim's being false? Assess these implications.
<i>Critique</i>	Whose perspective, voice, or feelings are upheld, strengthened, or made dominant by this claim? Who wins or loses if the claim is true? (Who or what is "governed" or "policed" by this claim?)	What "obvious" truths are taken for granted in this claim? What "common sense" but otherwise hidden assumptions are required to make this claim true and its argument work?	What ideas or sense of self is upheld, strengthened, or weakened by this claim and its argument or implications? What sense of community or society is strengthened or weakened?	What happens when we apply the claim and its "logic" (interpretations, arguments, counterarguments, and implications) to itself? (Look for what the claim says about philosophy, truth, society, persons, and so on.)

II. The Text Analysis Matrix (TAM).

My Text Analysis Matrix offers a methodical approach to assisting students in the development of an important set of thinking skills. Four categories of questions are addressed in the TAM. First, we ask about the claims found in a text. Second, we ask about evidence in the form of an argument or set of arguments. Third, the issue of implications comes to our attention. Finally, we look in the direction of critique, asking questions about the social and political situation being fostered or undermined by the claims, arguments, and, especially, implications of a reading. (For a summary of the questions that follow, please see Table 1, “Text Analysis Matrix”).

A. *Claims.*

I begin, then, with an attempt to isolate and examine the claims found in a particular text and the claims one makes in addressing that text.³ The first question to be put to a text might well be simply, “What’s the main idea?” Or, “What idea is supported by most of the rest of the text?” Teacher and learner can enter into an initial investigation of a philosophic work by wondering together about its main idea. I warn students that a given text might have multiple ideas that appear to be central or that appear to be supported by most of the rest of what is said. Further, we might legitimately disagree about the main idea, although presumably we can usually come to a consensus about what statements are central, even if we fail to agree about one single main idea. At the very least, we might agree that two or three statements have a legitimate claim to centrality in a given text and that certain other statements remain under dispute. The task of ascertaining what claims are central and what claims we wish to place on the list of possibilities for centrality can take considerable time and effort, but allows us a beginning point on the mutual journey into a critical assessment of a text.

Once we settle on a candidate or some candidates for main claim, we turn to questions of clarification, such as whether this claim, as we have understood it, is stated in a sentence that is either true or false.⁴ Frequently, students point to a question or a phrase as a main claim. When, however, we try to state the main claim as a sentence that is either true or false, we can immediately see that phrases and questions do not satisfy this criterion. It might be the case, of course, that a phrase that represents a major theme of the text can be transformed into a sentence that functions as the main claim. A question might also turn out to represent in a rhetorical way the main claim and might easily be changed into the correct form. The instructor nevertheless needs to remind students to be wary of taking a theme or a major question *as* the main claim, and students can quickly begin to see the difference, once they are asked whether they have found a statement or only a phrase or question. It is important, of course, not to denigrate the asking of questions or the forming or discovering of themes in a text. What we want to do instead is focus our critical energies on the statements in a text for which evidence will be given and that will present us with implications to consider. Questions and themes often organize our thinking in

³ Compare to Wolcott and Lynch (2001), Step 1 (“identify the problem”), Donald’s (2002) “description” phase, Maclellan and Soden’s (2001) “unpacking concepts,” and Loh’s (2003) “listen to understand” and “identify ... assumptions.”

⁴ The clarification steps that follow are generally comparable to Wolcott and Lynch’s (2001) identifying the problem, Donald’s (2002) steps of description and selection, and Maclellan and Soden’s (2001) “unpacking concepts.”

helpful ways and turn us to important new ideas. A claim, however, opens our thinking specifically to arguments, implications, and the possibility of a critique and helps us place the ideas to be addressed in clarifying order.

The next step involves assessing the clarity of claims. In thinking critically, one must ask oneself if the main idea, as currently understood, is clear.⁵ At this point, the instructor can easily direct students to the “principle of charity” (See Davidson, 2001, and Quine, 1960).⁶ Student assessments of claims frequently begin with exaggerating a minor flaw in someone’s claims or with interpreting claims or arguments in ways one can only assume the author would not have intended. Making claims clear then involves not only seeking to put claims in specific, plain terms but also seeking to place claims in their best light.

As a further step toward clarity, I often recommend to students that they seek to get a “picture” of the meaning of a claim or to find out if the claim “sounds right” or to grasp the main “feeling” of the claim – it is crucial that we recognize in working with ideas that different individuals represent ideas differently and that we seek to accommodate these differences. The use of the above mentioned different sensory approaches can help with at least some of those differences. Sometimes, of course, it will be difficult to clarify a claim.⁷ That might mean that the text itself leaves the claim essentially unclear or that we have failed fully to grasp the claim given in the text. Sometimes the lack of clarity can be attributed to the fact that the reader has misunderstood some larger portion of the text or its context in the life and culture of its author. Often lack of clarity in representing claims provides the instructor with an opportunity to work on students’ critical thinking skills and on fallacious ideas they hold about the topic as a whole or even about the single sentence being discussed.

Following the question of clarity one must address the question of interpretation.⁸ At this stage, I ask students whether a given interpretation of a claim seems reasonable to them or not. Questions concerning the implications and critique of claims will be discussed later. The pedagogical issue at this point lies with the importance of helping students to see that one person’s reasonable interpretation might be another’s unreasonable one. That is, the instructor can raise here the issue of the variety of interpretations and the variety of reasons that might be given for those interpretations.⁹ At the same time, one can use this opportunity to point out that not all interpretations are equal. If I say, “My house is on fire,” an interpretation that suggests that I am discussing how much I enjoy sailing the Mediterranean slips off the edge of the reasonable. On the other hand, varieties of interpretation can remind us of the multi-valence of language and of all there is to learn from those who see things differently.

⁵ The approaches to critical thinking discussed earlier suggest the importance of this step as well, for example clarification can play a role in Wolcott and Lynch’s (2001) step one, “identify the problem,” and in Donald’s (2002) procedure of “representation.”

⁶ Compare, for instance, Loh’s (2003) idea that one must “listen to understand” and “Search for strengths and value in other positions.”

⁷ Although none of the sources I mention above explicitly addresses the difficulty of one’s first step, it clearly must be taken into account and any approach to critical thinking does so, if only implicitly.

⁸ See Wolcott and Lynch (2001), Step 2; also Donald (2002), “State assumptions”; and Loh (2003), “Identify and re-evaluate assumptions.”

⁹ See Maclellen and Soden’s (2001) “alternative interpretation.”

B. Arguments.

I next turn to the issue of the support offered for main claims.¹⁰ Having identified a claim, we need to see whether the claim will stand up to scrutiny. One might, at this point, present arguments and counter-arguments of one's own, but the standard approach is to turn to the text for an understanding of the author's evidence for the claim. One must look first, then, for claims that seem intended to support the main claim (intended by the author, one might say – this issue of author intentions can, quite naturally, provide another fruitful line of questioning). There are numerous ways to go about this process, and the reader can adopt any of the standard procedures or can take a more novel approach. The main point here lies in looking for the evidence supporting a main claim in the form of other claims that appear as that evidence. The individual or class might, of course, disagree with the argument given and might even present a stronger, self-constructed argument. With beginning students, however, the instructor can focus on unearthing the author's (apparent) argument so that it can be assessed by the class. In everyday life the student will be faced with many claims as well as arguments allegedly supporting them. If she can develop some reasonable facility at noticing arguments, that alone will put her at an advantage in responding reasonably to the many claims of politicians, advertisers, religious advocates, and, for that matter, family and friends.

One must not only find an argument, however, but discover ways to assess it. For beginning students, my own hope lies in the broad sweep of more or less intuitive assessments of whether supporting claims work properly together and whether they make the conclusion necessarily or merely probably true (if they themselves are true). There are highly structured technical means for making these assessments, but in the midst of an introductory class I find an intuitive approach with some serious nods at structured technique to be all that's possible and more likely to remain with the student in everyday application. So I aim to enhance the students' intuitive skills at argument assessment rather than to produce proficient technicians of logic. This approach can be an aid for students in making sense of everyday claims and, in the best case, to be a method that can be generalized beyond the classroom and even disseminated to a student's circle of friends. Enhancing intuitive argument skills can be approached in a variety of ways. The main approach I take is to address a large number of arguments throughout the semester, thinking aloud with my students about whether a given argument works and why. Sometimes I simply describe the argument and its main strengths or weaknesses. Sometimes I present the argument and ask students questions to draw out of them an assessment of the argument. At other times I divide the class into smaller groups and ask the groups to work on various aspects of an argument that they then present to the class for our general assessment.

A further and important question concerns whether the given argument can legitimately be strengthened, enhancing the argument in the text by providing other arguments or by supplying missing or weak claims. The reader also needs to become aware of possible objections or counterarguments that have not been discussed by the text itself. In both cases, students need especially to address those points of view that they themselves find problematic. If a student, for instance, finds an argument for the existence of God appealing, it is important for that student to look seriously at objections or counterarguments. On the other hand, if a student finds an argument objectionable, say an argument concerning protecting the environment or feeding the hungry, that student needs to spend more time looking for ways to strengthen the argument. In

¹⁰ This point potentially includes Donald's (2002) "inference, synthesis, and verification" and Maclellan and Soden's (2001) "development" step.

both cases, however, the aim is to increase an awareness of the need to compensate for one's initial leanings.

C. Implications.

I want now to turn to an especially important issue for the critical understanding of a text – discovering the implications of key claims.¹¹ At this point one might either look for the implications of the main ideas as presented by the text itself or for implications seen by the reader that have not been explicitly mentioned in the text. Addressing such implications involves, first of all, recognizing what actions or ideas might follow from the main claim or claims of a text. To return to examples mentioned in the previous paragraph, we might ask: if “God exists” were true, what other claims about the universe and human life might also be true? If “protect the environment” is a legitimate normative claim, what actions might follow? For each claim we examine, we need to ask both what other claims or ideas seem linked to it and what actions or human practices does it imply. I sometimes ask students to consider what implications arise when a politician tells us that in order to save our freedom the government must take away our freedom (of course hardly any politician states such a claim this directly, but students usually understand the way hidden claims often lie behind political statements). To recognize implications often requires that we examine, as in this case, how words have been defined as well as what complex set of ideas and actions surrounds the given idea.

We must of course assess both arguments for and objections to associating certain implications with a given claim, recognizing that any statement about implications is itself a claim to be assessed alongside the original claim. Students might also find it useful to compare the possible implications of a claim with the possible implications of that claim's being false.

D. Critique.

The final aspect of thinking critically that I ask students to consider has to do with a critique of the claims, arguments, and implications already addressed.¹² Critique in this sense turns the reader toward the social and political aspects of claims. It is easy for academicians to think of our work as somehow outside the realm of social and political life, as though we spend our time amidst pure ideas, untouched by the winds of prejudice and privilege that affect everyday life. Yet especially if we are trying to teach students to think in ways that might affect them and their world we must recognize the ways our classroom remains part of that world. Ideas are living things; they resonate with personal, social, and political meaning. To claim that God exists or even that the sky is blue or that $a=a$ means something not just at a “purely philosophical” level, whatever that might be, but also in terms of how people think, live, relate to others, and decide about the way a community or state is going to function.

¹¹ At this point we move beyond the views of critical thinking mentioned in the Literature Review above. Although Maclellan and Soden (2001) mention “examining implications of evidence” they do not address under that topic the issues I focus on in this aspect of my critical thinking matrix. This is not to say that critical thinking approaches other than mine entirely fail to look at this issue only that it is not usually taken up as crucial to critical thinking.

¹² This aspect of my critical thinking matrix, one I find especially important to moving critical thinking from academic exercise to real-world endeavor, is not, to my knowledge, part of any other approach to critical thinking, though, of course, it is addressed by many thinkers in attempts to study and expose the play of politics and power in academe and in society.

So when one looks at claims, arguments, and implications, if the instructor wishes to be honest with her students and herself, she must also attend to the privileged perspectives, the apparently obvious truths, the sense of self and of community appealed to or assumed, and the internal logic contained within the texts addressed. There is a powerful tradition in academe of the neutrality of such endeavors, including the notion that if one approaches issues of identity, community, or state, except in a purely ‘academic’ fashion, the instructor or research is somehow sullying himself and his work. Yet the work of education itself lies squarely within processes of identity formation (“I am a philosopher”; “I am an educated person”; “I know x”; “I attended Harvard”, “...State U”, “... community college”, or “I only finished high school”, and so on), community solidarity (forms of knowledge, use of grammar, and peculiarities of accent as formalized or shunned in an educational setting mark the boundaries of a community as much as does geography or history), and national, racial or ethnic feelings and aims (we can think, e.g., of how history “changes” over time, not just because there’s more of it but because perspective and power shifts, thus shifting who or what is heard, seen, or valorized).

I turn then, in attempting to think critically, to questions concerning what perspectives, voices, or feelings have the upper hand in the claims found in a text. One can ask whether there are winners and losers if the claim is true, and who or what is governed or policed by the claim and its implications. The student and instructor do this in order to understand both the possible bias of the text itself and any personal need to reject or modify a claim that might otherwise be accepted due to its social or political implications. The idea here may be illustrated in the case of “God exists” or even “a=a”. If God does exist, then anyone who denies that it is so is wrong and, further, given how important the claim seems, might (it seems) deserve to be ignored at best or shunned, tortured, or killed at worst. Further, if God exists, then God’s chosen spokespersons (or, usually, *spokesmen*) must be given a key role in society. It is also worth noting that, if God exists and is male (whatever it might mean for God to have a gender or sexual role), then males are to be taken more seriously than females because they are more like God. I could continue, but this might serve as a sufficient beginning in understanding that claims such as “God exists” play a role that is far from neutral. In the case of “a=a” one can see that such a claim focuses on the identity of objects, presumably over time, since to say that something is identical to itself at one time seems silly at best. When taken out of the realm of abstract logic, which must be done if one intends to “use” the claim, we need to ask: What does such identity mean? Does it imply that anyone who learns anything is automatically “not herself” now? The implications of identity statements range over a wide territory. Politicians will frequently seek to convince potential voters that they have not changed their minds or “flip-flopped” on an issue. This implies, of course, that they are “themselves” over time, I suppose. It also valorizes ignorance over learning and emphasizes the significance of lying to the public as a political necessity. I once heard a philosopher talk about a famous contemporary philosopher in terms of two time periods, the later being “after x went crazy”. How did he go crazy? He changed his mind. Of course, in all fairness the change of mind resulted in a later position that the critic herself held in very low regard; the point, nevertheless, stands that one cannot very well learn anything without “changing one’s mind” and that changing one’s mind cannot therefore be held in low esteem if we hope for better ideas in philosophy, in politics, or in society in general.

Educators often seek to “govern” or “police” their students’ thinking. This activity is undertaken, usually at least, with the best of motives. One hopes to govern grammar and style, claims and arguments; and class climate and participation. All of these acts work, at best, to further pedagogical aims. But educators are not the only ones governing what may be thought or

spoken or what feelings and voices may be heard. The texts chosen and the standard interpretations play a role in what one can think or utter. If instructors fail to turn toward this issue, they ignore the role of education in identity, community, and state and the role of each of these in education. What gets ignored plays a role in the level of ignorance instructors and students embrace. Critical thinking skills cannot end with one's being able to assess arguments; one must also be able to think about the issues that lie hidden within the accepted range of arguments and ideas.

So I claim that critical thinking needs, in the end, not only to assess claims, arguments, and implications of claims, it needs further to take at least a preliminary look at the "obvious" and at "common sense" in order to encourage and enhance a critical consciousness in instructors and in students. As a final step in this process, then, one needs to apply the "logic" of a claim to itself. A famous philosophic example can be found in the basic claim of logical positivism. In brief summary, the claim is that any statement is meaningless (is not truly a statement but is instead merely an expression of one's feelings) if it is not either a tautology (a logical truth such as "it is either raining right now here or it is not") or an empirically verifiable statement ("the speed of light in a vacuum is x " – a verifiable statement need not be true it only needs to be able to be verified by experiment or observation). The immediate problem with this claim lies in the fact that when its logic is turned upon it, it fails its own test. That is, this claim is itself neither a tautology nor empirically verifiable.

"Only white males speak the truth" may pass the test of its own logic so long as it is stated by a white male, recognizing that such claims are seldom in our time period made as explicit as this. But suppose a significant part of this claim's target audience, if I may so speak, a woman for example, were to say this. Her words obviously cannot be trusted. So although she can acquiesce when a white male says these words, even her demure agreement is suspect, to the extent that the original statement is true. One could, of course, say that given the original claim it might be the case that those who are not white males can truthfully give signs of agreement that are not spoken. The point to be noticed here is the importance of looking and listening for ways that the logic of a claim says more than we first notice.

III. Application of the TAM.

I now offer some examples of ways the use of the TAM works for students. To do so I present here changes in student comments from the first day in a recent introductory Ethics course, before introduction of the TAM, to its end, after regular reference to it.¹³ In both cases, I asked students to address, in an in-class essay, the issue of lying, presenting them with the following set of alternatives: Lying is 1) always wrong, 2) often wrong, 3) sometimes wrong, or 4) never wrong. To judge the learning process of students, I compared answers given by students at the beginning and the end of the semester. There were 43 participants in the study. In the students' excerpts below, I use exact student wording except in a few cases of editing for spelling or for grammatical clarity. Pseudonyms are used in all cases.

A. Findings and Analysis.

The most striking difference overall in student essays about lying from the beginning to the end of the semester appears in relation to two aspects of the Text Analysis Matrix (see Table

¹³ The Ethics syllabus includes the TAM.

1); more sophisticated arguments and more awareness of the broad implications of one's beliefs about this practical issue. A number of students provided brief arguments in their first essays and mentioned some implications of their claims, but the development displayed in both areas is significant. Few students engaged in an obvious level of critique; I attribute the limited amount of stated critique to the specificity of the question I asked. On the surface, discussions of lying and truth-telling lend themselves less to questions of critique than some other areas of ethical interest because telling or withholding truth might appear to be a very personal decision and to relate primarily to direct personal relationships. Upon further reflection, however, I saw that questions of "who wins and loses," of the work of "common sense," and of one's "sense of community or society" (see Table 1, "Critique", Questions 1-3) definitely arose in student comments.

I found considerable critique demonstrated best in the final paper written by Steve, a nontraditional age male student. He addressed two major ethical theories in relation to lying and ended that discussion with the following:

What both of the positions lack is a grounding in real life. ... [It] seems to me that in this issue of moral principle, theory must intersect with real life or experience. Not to mention the fact that there is an emotional component to moral principle.

He went on to define "lying as ... 'knowingly telling an untruth with intent to harm.'" "In that case," he continued, "lying would always be wrong." Here we find the aspect of the TAM that I at first thought to be missing in these student essays; the question, "Who loses?" is addressed explicitly. Who loses? – The person lied to, and, most students agree in their final essays, the person doing the lying as well. What passes for common sense seemed to give students, as revealed by their first essays, the impression that lying was to be avoided due to its being "sinful" or to its being condemned by society but in their final essays they voiced a more "internal" critique of lying and of the alleged "common sense" that makes lying a purely subjective or cultural issue. Most of them, in their final essays, saw lying as a way of upholding the perspective of the stronger (or simply of the speaker) that cannot be morally or even logically maintained on that basis. The stronger (or the speaker) holds no special moral rights that allow her to get away with what is denied to others. The sense of community reflected in their concluding essays pushes the students beyond simple individualism to a consideration of a strong relationship between self and society and thus to issues of critique.

I turn now to a few additional excerpts from student papers that illustrate changing student responses from the beginning to the end of the semester. Matt, a nontraditional age male student, wrote at the beginning of the semester:

A Lie is merely a statement, it does not have a right or a wrong a correct or incorrect value to it until it is assessed by an individual or group using their perception. Values, such as Right and Wrong, are established by Society, Culture, Religion, and Individual weights of what is and is not seen as acceptable. If an individual lived in a society that had no set standards of acceptability, and that individual had no personal, cultural, or religious bias, then lying would just be a statement, and have no value of either right or wrong.

At the end, he said:

In order to determine the right or wrong ... of lying, one must first look upon the expectations of society and culture. ... [A]n individual within a culture must be able to communicate within that culture. It is expected that if one is speaking to his neighbor, her is speaking truth; if this was not the case ... society and culture would break down very

quickly. Under this assumption, this universally expected behavior, one can say that to tell the truth is what you should do, what is good to do, and what is right.

However, even taking into account the universal thought discussed above ... it is difficult to establish a complete answer. Some would argue that the result of the lie is what determines its value. ... [The] necessity to account for the individual makes it impossible to establish a complete and absolute answer.

We see, then, at the end of the semester a greater sophistication in the answer, even though his choice from the options provided is the same in both cases. More significantly, one can trace in the final essay themes related to the TAM. Although Matt provided a *claim* in each case, he backed up his final claim with stronger and more precise *arguments*. Of even greater importance, Matt's final essay took into account a level of analysis frequently not found in initial student essays by setting forth some of the *implications* of his claim. He did not venture noticeably into the area of *critique* in his essay, but in referring to the "necessity to account for the individual" he made a gesture in that direction.

Another student, Marti, a traditional age female student, changed her answer from saying at the beginning of the semester that lying is always wrong to concluding at the end that it is often wrong. In her initial essay, she said, "I can not think of any instance where I think it would be completely "right" or accepted by society to tell a lie" At the end of the semester, she offered as a reason for her change of mind the idea that "... in order to 'practice' becoming virtuous, we will make ethical mistakes." She went on to say, "As long as a person is working toward a virtuous life [basing her idea here on Aristotle] and using their mistakes to build their 'skill' of virtue, a lie can be forgiven and ... may not always be wrong." She concluded her essay by saying, "My theory may be criticized by those who have a different definition [of lying] than I." There is considerable sophistication in this concluding student essay. As in the previous student's essays, it is important to recognize that she offered not only a claim and arguments to back up that claim, but explicitly recognized the implications of her claim and noted that the obvious (to her) truth of her particular definition of lying need not carry over into everyone's assessment of the idea (see Table 1, "*Critique*", Question 2).

Alicia, a traditional age female student, said at the beginning of the semester, "Lying is always wrong because it's a sin." In her final essay, however, she claimed that her parents told her that "in order to be a good kid and get to go to heaven I had to tell the truth always no matter what." She said that she "often felt like that was stupid because if you could spare someone's feelings without them finding out, then why tell them the truth and hurt them?" She concluded that "After taking this class I believe my parents were right." She followed this statement with a long argument for the value of telling the truth. In this case the student originally had a sense of the negative implications for oneself of lying; presumably as a sin, lying is subject to severe penalties. In the end, however, she provided a lengthy discussion of exactly why lying is wrong and how to respond in the most desperate situations, citing a specific example that is parallel to a commonly cited ethical dilemma; if a girl hides in one's house for fear of her boyfriend should one lie to the boyfriend? The implications she mentioned in her final essay reflect an interest in the implications for others of one's actions and for a more sophisticated assessment of one's moral situation than those offered at the beginning.

B. Conclusion concerning Application of TAM.

Of the four aspects of the TAM (*claims, arguments, implications, and critique*), claims were most obviously presented both at the beginning and at the end in all student papers, arguments were found more frequently in final papers than in initial ones, implications appeared either in a more sophisticated form or solely in end-of-semester work, and critique appeared almost exclusively at the end, when most students saw beyond the personal issues of getting caught, having to pay, or either following or denying societal claims to the larger issues of one's intimate and undeniable engagement with one's community and world. The distinction between societal claims that one unthinkingly either takes up or denies and a clear sense of community that one embraces thoughtfully lies at the heart any of genuine critique as applied to texts, ideas, or one's own writing and thinking. Most students come into any class having been sufficiently imbued with the claims of their society and culture. A crucial part of critical thinking, as I view it, is to move students beyond the simple "common sense" they share with peers, family, and friends to new ways of critically assessing the ideas offered them.

IV. Summary.

In the end, then, I am far from thinking that I have now stated for once and all what it means to think critically or to approach texts philosophically, but the TAM can aid instructors as teachers/learners and students as learners/teachers to approach academic texts and the texts of everyday life with more clarity and insight as we work together to hone our natural, useful, and too often underdeveloped critical skills for the academic and life tasks that confront us.

Acknowledgements

Thanks to Deborah Biss Keller, IUPUI School of Education, for comments, questions, and clarifications. Special thanks are due as well to the IUPUI Center for Teaching and Learning, the Gateway Scholars program, and especially to Dakin Burdick formerly of the Center for Teaching and Learning for encouraging me to clarify many of the ideas contained in the TAM. Last, but far from least, thanks to numerous students over many years whose wisdom, questioning, and constant desire to read, write, and think more clearly has shaped and continues to shape my own thinking about these issues.

References

- Davidson, Donald. (2001). Radical interpretation. In *Inquiries into truth and interpretation*, second edition. (Oxford: Clarendon Press). Originally published in *Dialectica: International Journal of Philosophy of Knowledge*, 27(1973), 314-328.
- Donald, Janet. (2002). *Learning to think: Disciplinary perspectives*. San Francisco, CA: Jossey-Bass.
- Heron, Gavin (2006). Critical thinking in social care and social work: Searching student assignments for the evidence. *Social Work Education*, 25(3), 209-224.

Huemer, Michael (2005). Is critical thinking epistemically responsible? *Metaphilosophy*, 36(4), 522-531.

Loh, Christine (2003). Promoting sustainability tools: Connecting thinking and dialogue skills. *Civic Exchange*, January 2003.

Lynch, Cindy L. and Wolcott, Susan K. (2001). Helping your students develop critical thinking skills. IDEA Paper #37. Manhattan, KS: The IDEA Center.

MacLellan, E. and Soden, R. (2001) Making explicit tutors' judgements about quality in students' academic writing in the Faculty of Education. Report of an Initiative for the Teaching and Learning Quality Improvement Committee, University of Strathclyde.

Pithers, R.T. and Soden, Rebecca (2000). Critical thinking in education: a review. *Educational Research*, 42(3), 237-249.

Quine, Willard Van Orman. (1960). *Word and object*. Cambridge, MA: MIT Press.

Effectiveness of Calibrated Peer ReviewTM for improving writing and critical thinking skills in biology undergraduate students

Adalet Baris Gunersel, Nancy J. Simpson, Karl J. Aufderheide, and Li Wang¹

Abstract: This study focuses on student development with Calibrated Peer Review (CPR)TM, a web-based tool created to promote writing and critical thinking skills. Research questions focus on whether or not students showed improvement in writing and reviewing competency with repeated use of CPR in a senior-level biology course and whether the difference between higher performing and lower performing students decreased over time. Four repeated measures analyses were conducted with different sets of students. Repeated measures analyses indicate that students showed improvement in writing skills and reviewer competency with repeated use of CPR. The difference between higher and lower performing students decreased over time in both writing skills and reviewer competency.

Keywords: science education, critical thinking, innovative teaching tools, writing skills, peer review, undergraduate education.

Calibrated Peer Review (CPR)TM is a web-based tool for authoring and managing student writing assignments (for more information, see <http://cpr.molsci.ucla.edu/>). CPR assignments engage students in writing and in reviewing their peers' work, and include a calibration phase during which students practice reviewing according to an instructor-designed rubric. While there are a few published studies that provide evidence of the value of CPR as a tool for improving students' conceptual learning, as well as their writing and critical thinking skills (e.g., Furman and Robinson, 2003; Gerdeman, Russell, and Worden, 2007; Margerum, Gulsrud, Manlapez, Rebong, and Love, 2007; McCarty et al., 2005) more research is needed. What are the characteristics of effective CPR assignments? Is CPR effective for all students? What strategies for implementation lead to success? Questions such as these intrigued a biology instructor and two faculty developers, all of whom had been working on an NSF-funded project focusing on CPR. This joint curiosity led to the current study investigating student outcomes in three semesters (Spring 2005, Spring 2006, and Spring 2007) of a senior-level biology course by using repeated measures analyses of CPR-generated data. The focus is the effectiveness of repeated use of CPR for improving student writing and reviewing competency in biology with CPR. The course's instructor, who was a part of the research team, provides information regarding the course to give a context for the study. This study adds to the literature about the effectiveness of CPR by investigating the development of student writing in biology and reviewing competency in a senior-level biology class. Specifically, there are three research questions:

1. Did repeated use of CPR in a senior-level biology course result in improvement in writing and reviewing skills of initially lower performing students?

¹ Center for Teaching Excellence, 533 Blocker, MS 4246, Texas A&M University, College Station, TX 77843, bgunersel@tamu.edu; n-simpson@tamu.edu; kauf@mail.bio.tamu.edu; wangli_99@yahoo.com

2. Did the difference between higher performing and lower performing students decrease with repeated use of CPR?
3. Did repeated use of CPR in a senior-level biology course result in improvement in writing and reviewing skills in general?

I. Calibrated Peer Review (CPR)TM.

In order to provide a greater understanding of the study, further explanation of CPR and research related to CPR are presented in this section. Developed at UCLA for the Molecular Science Project, one of the NSF-supported Chemistry Systemic Reform Initiatives, CPR was designed to give students practice in writing and peer review, since both are expected competencies in scientific fields (Russell, 2001).

One of CPR's aims is to develop students' skills of discipline-specific writing, which has become prominent in education (Emerson, MacKay, MacKay, and Funnell, 2006; Lea and Street, 1998). The underlying the pedagogy of CPR is reinforced by numerous studies supporting the educational value of both writing (Holliday, Yore, and Alvermann, 1994; Klein, 1999; Kovac and Sherwood, 1999; Lowman, 1996; Rivard, Stanley, and Straw, 2000) and peer review (Falchikov, 1995; Orsmond, Perry, and Callaghan, 2004; Searby and Ewers, 1997; Sluijsmans, Brand-Gruwel, and Van Merriënboer, 2002; Sluijsmans, Docky, and Moerkerke, 1999; Topping, 1998). Although peer review may be a source for hesitation among students, several studies suggest that peer review can be as reliable as faculty assessment (Falchikov, 1995; Freeman, 1995; Saavedra and Kwun, 1993; Sluijsmans et al., 1999; Stefani, 1994; Topping, 1998).

In addition to having students write and review peers' work, CPR has the students practice reviewing in the "calibration phase." In order to create a CPR assignment, instructors produce the following components:

Instructions for writing. Instructions include suggested resources, questions to guide student thinking, and a "writing prompt" that tells students such things as the topic, format and audience for their writing.

Calibration questions. A set of questions that direct students attention to content and style characteristics of a completed assignment and form the basis for assigning a text rating.

Three sample essays. High, average, and low quality essays that are the responses to the assignment. (Instructors review and rate these essays using the calibration questions.)

Student work on a CPR assignment occurs in three phases:

Text entry phase. Students read instructions, access suggested resources, and write and submit their essays.

Calibration phase. Students are presented with the three sample essays, along with the calibration questions. For each essay, students answer the calibration questions and assign a rating. CPR assigns a reviewer competency index based on a comparison of the student review to the instructor review of each essay;

Review phase. Students are presented first with three classmates' essays (randomly assigned and anonymous) and then with their own essay, all of which they review and rate using the same set of calibration questions.

Instructor-reported experiences and a limited number of studies have suggested that CPR is a tool that can help students master content, improve writing skills, and become more competent reviewers (Furman and Robinson, 2003; McCarty et al., 2005; Russell, 2001). Gerdeman, Russell, and Worden (2007) examined the development of 1330 students' writing and

reviewing skills in an introductory biology course and found that students showed improvement in writing and reviewing over three CPR assignments. Margerum et al.'s (2007) survey with first-semester general chemistry students suggested that students felt they were becoming "better technical reviewers" with CPR assignments (p. 294). They also found that students mastered the class content through both the calibration phase and the review phase. Pelaez (2002) compared the learning outcomes of undergraduate nonscience majors taught with lectures and taught with CPR™ in an introductory physiology course. The results suggested that the performance of students who had completed problem-based learning assignments in CPR was better than or equal to the performance of students who had received "traditional instruction" (p. 181). Pelaez (2002) noted:

The favorable results may be a product of the work students complete when writing about their thinking, or perhaps students did better because PW-PR (problem-based writing with peer review) made it possible for them to confront and resolve difficulties they encountered relating concepts. (p. 181)

II. The Context of the Study.

Data from students in three semesters (Spring 2005, Spring 2006, and Spring 2007) of a senior-level biology course were used. Each semester, students completed the same four CPR assignments and three highest scores counted for the final grade. The assignments were ordered in increasing difficulty: "Why Do We Use The SI System Of Measurement In Science?" "Mitosis Through the Microscope: Advances in Seeing Inside Live Dividing Cells," "Microtubules and Motor Proteins," and "Cajal Bodies and Coilin—Moving Towards Function." While the first one was an example assignment from the CPR assignment library, the other three were created by the instructor.

This ordering made the assignments get "more focused on a specific area of cell biology and much more detailed in the kinds of information a student would have to collect and condense into a series of paragraphs." The instructor used the assignments for a dual purpose: They were related to lecture topics and there was "a sequence of increasing complexity and specific focus as to the nature of the information that they're going to have to deal with."

The second assignment (in 2005 and 2006) ("Mitosis Through the Microscope: Advances in Seeing Inside Live Dividing Cells") was a historical overview of how a specific microscope has been used in cell biology. The article that the students had to work with to complete the assignment was a "general article, an overview" that was "roughly coordinated to some of the classes they did in the beginning of the semester." The third assignment ("Microtubules and Motor Proteins") was "much more detailed about a specific set of cellular structure and motor proteins that interact with them." The structure of this assignment was slightly different: Students didn't have one article as a source, rather they were linked to a series of research websites. This assignment was more difficult than the second, since students were "doing much more of a diffuse search to several sources of information." The fourth assignment ("Cajal Bodies and Coilin—Moving Towards Function") was the most difficult of all, partly because of the topic of the assignment, but also but also because the source material (the review article) they read was not very well written and was poorly organized. Therefore to answer the guiding questions, they had to read the entire essay and select information from paragraph to paragraph to construct a comprehensive narrative. So, in fact, students had to write an essay that was better organized than the original source material.

III. Methods.

Specifically, the study addresses three questions:

1. Did repeated use of CPR in a senior-level biology course result in improvement in writing and reviewing skills of initially lower performing students?
2. Did the difference between higher performing and lower performing students decrease with repeated use of CPR?
3. Did repeated use of CPR in a senior-level biology course result in improvement in writing and reviewing skills in general?

Data from students in three semesters (Spring 2005, Spring 2006, and Spring 2007) of a senior-level biology course were used. For the analyses two CPR-generated scores, reviewer competency index (RCI) and text rating (TR), were included as dependent variables. The reviewer competency index (RCI) is computed (by the CPR program) following student review of three instructor-provided essays. RCI computation uses a comparison of student and instructor responses to instructor-generated calibration questions, as well as of student and instructor global rating of the essays. Text rating (TR), on the other hand, is a weighted average of scores given by three peer reviewers. Weighting is based on reviewing competency (RCI) of the peer. Peer reviewers are instructed to base the score on analysis guided by the calibration questions. Since the calibration questions include both content-related questions and writing-related questions, TR can reflect both content understanding and writing competence. In summary, TR is used as a measure of writing quality and content understanding, while RCI is used as a measure of students' ability to review. For each CPR assignment students receive a TR ranging from 1 to 10 and a RCI ranging from 1 to 6. Students who had completed fewer than three of the four assignments were eliminated from the analysis.

Students were categorized into two groups according to their TR and RCI scores from the first assignment: higher performing (third quartile; highest 25%) and lower performing (first quartile; lowest 25%) (see Table 1). The second quartile was eliminated in order to focus on the development of higher and lower performing students. Thus, TR scores of 47 students (18 from Spring 2005, 15 from Spring 2006, and 14 from Spring 2007) were included and RCI scores of 83 students (27 from Spring 2005, 26 from Spring 2006, and 30 from Spring 2007) were included (see Table 1). This discrepancy between the numbers occurred since the second quartile was larger for TR scores than RCI scores.

In addition to this, all students regardless of performance level were included in separate repeated measure analyses. Table 4 presents the number of students for each assignment.

Table 1. Higher and lower performing groups, 2005-2006-2007.

	Lower Performing	Number	Higher Performing	Number	All students
TR1	≤ 6.4925	23	≥ 8.9600	24	47
RCI1	≤ 2.000	45	≥ 6.000	38	83

IV. Data Analysis.

Four repeated measures analyses were conducted, two of which focused on students at initial performance levels and two of which included all students, regardless of performance level. The first analysis included the TR scores of a total of 47 students. The second analysis

included the RCI scores of a total of 83 students (27 from Spring 2005, 26 from Spring 2006, and 30 from Spring 2007) from groups of higher and lower performance. Both analyses included lower performance and higher performance as the grouping variable and the number of assignments (4) as the within-subjects factor.

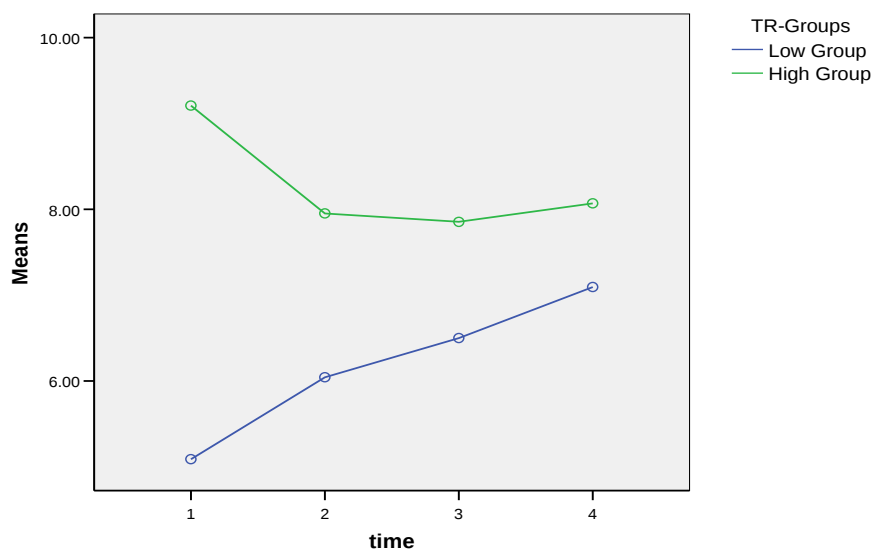
The other two repeated measures analyses included all students regardless of performance level. Students' TR and RCI scores were used as dependent variables, the semester as the grouping variable, and the number of assignments (4) as the within-subjects factor.

V. Results.

When considering initially higher and lower performing students in TR, although there was no overall statistically significant change over four assignments ($df= 3$, $F= 1.813$, $p< 0.149$), the change of means of higher and lower performing students was statistically significant over four assignments at alpha level .01 ($df= 3$, $F=14.370$, $p< 0.000$). The mean for the lower performing group increased steadily throughout the semester, while the higher performing group's mean decreased (see Table 2 and Graph 1). Also, the difference between the groups decreased throughout the semester.

Table 2. TR Descriptive Statistics for Spring 2005-2006-2007.

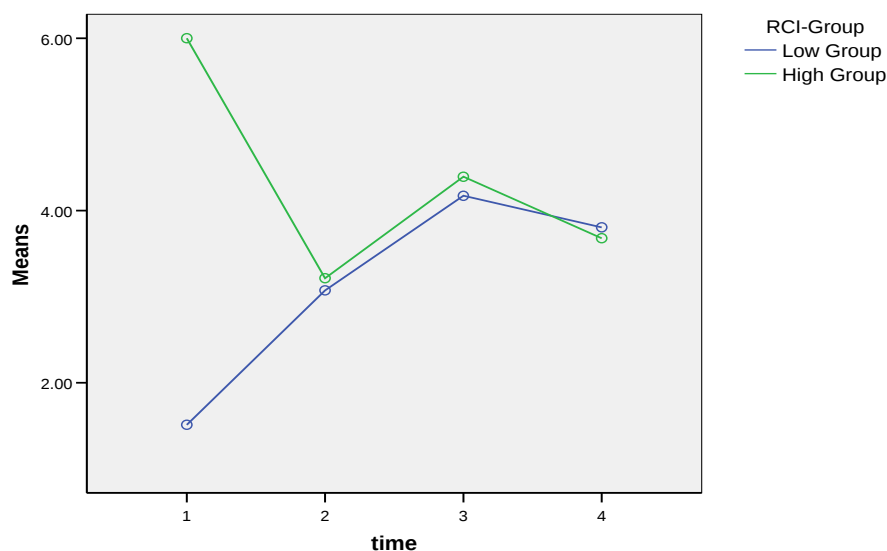
		<i>M</i>	<i>SD</i>
TR1	Higher performing	9.2079	.30570
	Lower performing	5.0876	1.04912
	All students	7.0448	2.22479
TR2	Higher performing	7.9521	1.45644
	Lower performing	6.0419	1.65149
	All students	6.9493	1.81961
TR3	Higher performing	7.8542	1.03237
	Lower performing	6.4995	1.56828
	All students	7.1430	1.49083
TR4	Higher performing	8.0689	.96489
	Lower performing	7.0938	1.02561
	All students	7.5570	1.10106

Graph 1. Spring 2005-2006-2007 means plot of TR Groups.

When considering initially higher and lower performing students in RCI, there was overall statistically significant change four assignments at alpha level 0.01 ($df=3$, $F=8.479$, $p<0.000$) and significant change in higher and lower performing students at alpha level 0.01 ($df=3$, $F=47.829$, $p<0.000$). The lower performing students' group improved throughout the semester, while the higher performing group fluctuated (see Table 3 and Graph 2). However, the means of both groups decreased in the fourth, most difficult, assignment. All students together showed a statistically significant increase from the first assignment to the last (means of 3.3333 to 3.7536) and the difference between the groups had almost disappeared by the second assignment (means of 3.6786 and 3.8049).

Table 3. RCI Descriptive Statistics for Spring 2005-2006-2007.

		<i>M</i>	<i>SD</i>
RCI1	Higher performing	6.000	0.0000
	Lower performing	1.5122	0.77852
	All students	3.3333	2.29876
RCI2	Higher performing	3.2143	1.54817
	Lower performing	3.0732	1.43858
	All students	3.1304	1.47442
RCI3	Higher performing	4.3929	1.59488
	Lower performing	4.1707	1.56369
	All students	4.2609	1.56855
RCI4	Higher performing	3.6786	1.46701
	Lower performing	3.8049	1.70616
	All students	3.7536	1.60336

Graph 2. Spring 2005-2006-2007 means plot of RCI Groups.

When the TRs of all students were included, there was a statistically significant change over four assignments at alpha level 0.05 ($df= 3$, $F= 2.814$, $p< 0.041$). However, there was no statistically significant change when separated according to different semesters ($df= 6$, $F= 0.888$, $p< 0.0506$). While the mean of the TR scores initially decreased from the first assignment to the second, they increased steadily from the second to the fourth assignment (see Table 4).

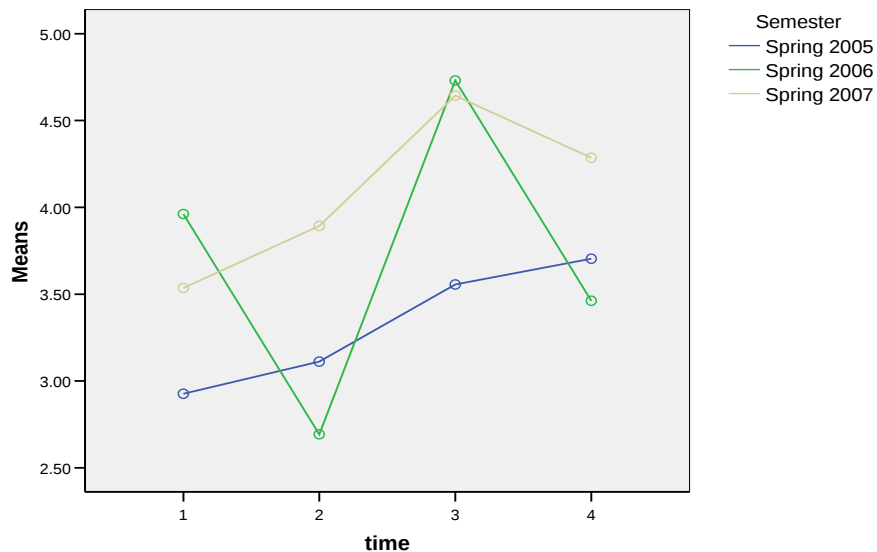
Table 4. TR Descriptive Statistics (semesters combined).

Assignment	# of Students	Minimum	Maximum	<i>M</i>	<i>SD</i>
1	94	2.89	10.00	7.5089	1.62404
2	93	3.12	11.98	7.2424	1.51903
3	87	3.44	10.00	7.3366	1.39329
4	78	4.06	9.52	7.6663	1.09819

When the RCIs of all students were included, there was statistically significant change over four assignments at alpha level 0.01 ($df= 3$, $F= 6.709$, $\eta^2=0.088$, $p< 0.000$). There was also statistically significant change when separated according to different semesters at alpha level 0.05 ($df= 6$, $F= 5.871$, $p< 0.042$). While RCIs slightly declined from the first assignment to the second, they increased from the second to the third, but then decreased again on the fourth, most difficult assignment (Table 5). While students in Spring 2005 showed constant improvement in their RCIs, students' means in Spring 2006 fluctuated (see Table 5 and Graph 3). Students in Spring 2007 showed improvement during the first three assignments, and then a decline. However, the mean on the final assignment was higher than the mean on the first one.

Table 5. RCI Descriptive Statistics.

		<i>M</i>	<i>SD</i>
RCI1	Spring 2005	2.9259	2.12903
	Spring 2006	3.9615	2.48967
	Spring 2007	3.5357	1.75293
	Combined	3.7238	2.06386
RCI2	Spring 2005	3.1111	1.18754
	Spring 2006	2.6923	1.43581
	Spring 2007	3.8929	1.87260
	Combined	3.6058	1.75940
RCI3	Spring 2005	3.5556	1.57708
	Spring 2006	4.7308	1.56353
	Spring 2007	4.6429	1.39348
	Combined	4.4796	1.58751
RCI4	Spring 2005	3.7037	1.83586
	Spring 2006	3.4615	1.44861
	Spring 2007	4.2857	1.54750
	Combined	3.8876	1.51836

Graph 3. Means plot for RCI.

VI. Conclusions.

Our study suggests that repeated practice with CPR is an effective way to help students develop writing and reviewing skills in biology, supporting other studies that have found CPR's usefulness with writing and reviewing skills (e.g., Furman and Robinson, 2003; Gerdeman, Russell, and Worden, 2007; Margerum et al., 2007; McCarty et al., 2005; Pelaez,

2002). Since the instructor-created rubric included both content-related criteria and writing-related criteria, improved TRs indicate increased student ability both to understand the content focus of the CPR assignments and to write about this content in a coherent manner. In addition to this, results showed that the difference between the higher performing and lower performing students decreased in both TRs and RCIs.

Repeated use of CPR appears to be particularly beneficial for initially lower performing students. In our study, the students who did poorly on the first assignment exhibited statistically significant improvement with repeated use of CPR in both TR and RCI. While the difficulty of the final assignment did not impact TRs, it did impact RCIs, as student scores showed a decrease. Still, the RCIs of the lower performing students on the final assignment were higher than those on the first assignment. This improvement occurred despite the fact that grading rubric was different for each assignment. The students appeared to have become more adept at, as the instructor put it, “internalizing” a set of criteria for evaluation.

Initially higher performing students showed a slight, but significant, decrease in their TRs (Graph 1) which could have two reasons: Since three of the highest scores counted, students’ efforts may have decreased, or it may have been the “regression to the mean” which suggests that students’ initially high scores would be more likely to decrease. These students fluctuated in their RCIs (Graph 2). Just like the lower performing students, their RCIs were impacted by the increased difficulty of the fourth assignment, while their TRs were not. A possible reason for this finding may be that the difficulty of the assignment impacted students’ ability to match the evaluations of the instructor, which is what the RCI is based on. Another possible reason is the following: The difficulty of the fourth assignment rose from its text which was not well written. Thus, it may have been more difficult for the students to examine the details of the text and rate their peers in a way that is similar to the instructor’s rating, which impacts their RCIs. On the other hand, the difficulty of the text may not have impacted students’ ability to get enough information to write a medium- to high-quality essay, which leads to TRs remaining unaffected by the difficulty.

The results for the initially higher and lower performing student groups are consistent with the findings of Gerdeman, Russell, and Worden (2007): Students with the lowest initial levels of performance gained the most over time, while students with the highest levels of performance slightly declined.

Analysis including all students regardless of performance level also bore interesting results. There was a significant change in TRs over four assignments with a steady increase after the second assignment. There was no significance when students were separated according to semesters possibly due to the small sample size in each semester (22 from Spring 2005, 20 from Spring 2006, and 28 from Spring 2007).

Changes in RCIs in each semester varied and were statistically significant. While in Spring 2005, students’ RCIs showed a statistically significant increase, in Spring 2006 scores fluctuated, with means decreasing from the first to the second assignment, then increasing from the second to the third, and decreasing again from the third to the fourth. Students’ scores in 2007 showed a statistically significant increase over three assignments, and then a decline. This decline is not unexpected as the fourth assignment was the most difficult. In 2005 and 2007, RCIs on the final assignment were higher than the initial, which was not the case in 2006.

In addition to this, CPR brought students with initially different levels of performance closer together in their scores: for both variables, TR and RCI, the difference between lower performing and higher performing students decreased over four assignments, which can be

observed in Graph 1 and 2. The initial differences in performance could be attributed to levels of preparation and ability which the students brought to class. CPR seems to be a useful tool that helps students overcome initial shortcomings and brings students together in skills.

It should also be noted that both the advancement of students who were at a lower performing level in the first assignment and the decline of the differences between students at different levels of abilities took place without a significant amount of feedback from the instructor outside of the CPR program. This is an important aspect of CPR as it which is frees time for other instructional tasks and gives instructors with large classes the opportunity to use writing.

In today's learning environment where it is important to be able to critique and to probe, CPR shows promise as a learning tool that gives students the opportunity to exercise their writing and critical thinking skills and opens new avenues to learning. Our interview with the instructor revealed that these aspects of CPR were his reasons to continue using this educational tool, although it was unfamiliar for both him and the students. He believed that college students needed further experience in writing and reviewing—using specific grading standards—which they would need in the future.

This study was a retrospective analysis and not an a priori designed experiment. It addressed questions regarding the effect of CPR and student learning, and used instructor reflection to interpret data generated by CPR. Since CPR assignments are discipline-specific, and implementation of CPR assignments is strongly influenced by the context and structure of the course, the accumulation of studies in a range of disciplines and contexts will be needed for greater understanding of factors influencing the effectiveness of CPR as an educational tool. A future study may also include interviews with the students in order to get an understanding of their experience and perspective.

Acknowledgements

We wish to acknowledge the consultation and feedback from Dr. Victor Wilson and Dr. Stephanie Knight, Department of Educational Psychology, Texas A&M University, and Dr. Arlene Russell, Department of Chemistry, University of California, Los Angeles in preparing this manuscript. This material is based upon work supported by the National Science Foundation under Grant No. DUE-0243209.

References

- Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15(1), 101-111.
- Cross, K. P. (1998). Classroom research: Implementing the scholarship of teaching. In Angelo, T. (Fall, 1998). *New Directions for Teaching and Learning: Classroom assessment and Research: an update on uses, approaches, and research findings* (75). San Francisco: Jossey-Bass.
- Cutler, H., and Price, J. (1995). The development of skills through peer assessment. In A.

Edwards

and P. Knight (Eds.), *Assessing Competence in Higher Education* (pp. 150-159). London: Kogan Page.

Educause Learning Initiative (September 2005). Calibrated Peer Review: A writing and Critical-thinking instructional tool. *Innovations and Implementations: Exemplary Practices in Teaching and Learning*. UCLA, USC. Retrieved October 15th, 2005 from <http://www.educause.edu/ir/library/pdf/ELI5002.pdf>.

Ellis, G. (2001). Looking at ourselves – self-assessment and peer assessment: Practice examples from New Zealand. *Reflective Practice*, 2(3), 289-302.

Emerson, L., MacKay, B. R., MacKay, M. B., and Funnell, K. A. (2006). A team of equals: Teaching writing in the sciences. *Educational Action Research*, 14(1), 65-81.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International*, 32(2), 175-187.

Freeman, M. (1995). Peer assessment by groups of group work. *Assessment and Evaluation in Higher Education*, 20(3), 289-301.

Furman, B., and Robinson, W. (2003). *Improving engineering report writing with Calibrated Peer Review*. Paper presented at the 33rd ASEE/IEEE Frontiers in Education Conference, November 5-8, 2003, Boulder, CO, pp. F3E-14-F3E-15.

Gerdeman, R. D., Russell, A. R., and Worden, K. J. (2007). Web-based student writing and reviewing in a large biology lecture course. *Journal of College Science Teaching* (March/April 2007), 46-52.

Goody, J. (1994). *Entre l'oralite' et l'e'criture*. Paris: Presses universitaires de France.

Halliday, M. A. K., and Martin, J. R. (1993). *Writing science: Literacy and discursive power*. Pittsburgh, PA: University of Pittsburgh Press.

Holliday, W.G., Yore, L. D., and Alvermann, D. E. (1994). The reading-science learning-writing connection: Breakthroughs, barriers, and promises. *Journal of Research in Science Teaching*, 31, 877-894.

Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, 11(3), 203-270.

Kovac, J., and Sherwood, D. W. (1999). Writing in chemistry: An effective learning tool. *Journal of Chemical Education*, 76(10), 1399-1403.

Langer, J. A., and Applebee, A. N. (1987). *How writing shapes thinking* (Research Report No.

22). Urbana, IL: National Council of Teachers of English.

Lea, M. R., and Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, 23(2), 157-172.

Liu, J., Pysarchik, D. T. and Taylor, W. (2002). Peer review in the classroom. *BioScience*, 52(9), 824-829.

Lowman, J. (1996). Assignments that promote learning. In R. J. Menges, M. Weimer, and Associates (Eds.), *Teaching on solid ground: Using scholarship to improve practice*. San Francisco: Jossey-Bass.

Margerum, L. D., Gulsrud, M., Manlapez, R., Rebong, R., and Love, A. (2007). Application of calibrated peer review (CPR) writing assignments to enhance experiments with an environmental chemistry focus. *Journal of Chemical Education*, 84(2), 292-295.

McCarty, T., Parkes, M. V., Anderson, T. T., Mines, J., Skipper, B. L., and Greboksy. (2005). Improved patient notes from medical students during web-based teaching using faculty-calibrated peer review and self-assessment. *Acad Med*, 80, 67-70.

McGinley, G. A., and Tierney, R. J. (1989). Traversing the topical landscape: Reading and writing as ways of knowing. *Written Communication*, 6, 243-269.

McKeachie, W. (2002). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston: Houghton Mifflin Co.

National Council of Teachers of Mathematics. (1993). *Assessment Standards for School Mathematics: Working Draft*, Reston, VA: NCTM.

Orsmond, P., Merry, S., and Callaghan, A. (2004). Implementation of a formative assessment model incorporating peer and self-assessment. *Innovations in Education and Teaching International*, 41(3), 273-290.

Pelaez, N. J. (2002). Problem-based writing with peer review improves academic performance in physiology. *Advanced Physiology Education*, 26, 174-184.

Pope, N. K. (2005). The impact of stress in self- and peer assessment. *Assessment and Evaluation in Higher Education*, 30(1), 51-63.

Rivard, L. P., Stanley, B., and Straw, S. B. (2000). The effect of talk and writing on learning science: An exploratory study. *Science Education*, 84(5), 566-593.

Russell, A. (2001). *The evaluation of CPR. Prepared for HP e-Education; Business Development*. Los Angeles: UCLA.

Saavedra, R., and Kwun, S. K. (1993). Peer evaluation in self-managing work groups. *Journal of*

Applied Psychology, 78(3), 450-462.

Shafer, J.L. (1997). Software for multiple imputation. University Park, PA: The Pennsylvania State University Department of Statistics.

Sherwood, D. (1999). Writing in chemistry: An effective learning tool. *Journal of Chemical Education*, 76(10), 1399-1403.

Searby, M., and Ewers, T. (1997). An evaluation of the use of peer assessment in higher education: A case study in the school of music, Kingston University. *Assessment and Evaluation in Higher Education*, 22(4).

Sluijsmans, D., Brand-Gruwel, S., Van Merriënboer, J. (2002). Peer assessment training in teacher education. *Assessment and Evaluation in Higher Education*, 27(5), 443-454.

Sluijsmans, D., Dochy, F., and Moerkerke, G. (1999). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1, 293-319.

Sobral, D. T. (1997). Improving learning skills: A self-help group approach. *Higher Education*, 33, 39-50.

Stefani, L. A. J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education*, 19(1), 69-75.

Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.

Topping, K. J., Smith, E. F., Swanson, I., and Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education*, 25(2), 149-169.

College faculty and the scholarship of teaching: Gender differences across four key activities

Carrie B. Myers¹

Abstract. Prior work has spent much effort on what is the Scholarship of Teaching and Learning (SoTL) and little effort on who actually practices the SoTL. I begin to fill the research gap by examining two of the most salient characteristics that structure faculty dynamics and everyday lives—gender and years of experience—and four activities generally agreed to be indicators of the SoTL. Guided by identity theory in tandem with recent data on the characteristics and trends of higher education faculty in the U.S., I find that female faculty are more likely than their male counterparts to engage in the SoTL and that this gender gap increases with teaching experience. The data come from 2001-2002 survey data on 85 faculty at a large public research-intensive university in the Northwest.

Keywords: teaching, higher education, gender, identity theory

I. Introduction.

In *Scholarship Reconsidered* Boyer (1990) introduced the influential concept “the scholarship of teaching” into the academic community. This concept has been refined by others—most notably by the Carnegie Academy—both theoretically and substantively as the Scholarship of Teaching and Learning (SoTL), and has attracted widespread attention mostly in the UK, Australia, Canada, and the U.S. The SoTL is not entirely novel. Prior to Boyer (1990), higher education faculty have been historically urged to inform and improve their teaching by consulting experts and by using knowledge and resources generated by commissions, special offices, and practitioners. In the U.S., Cross and Steadman (1996) argue that efforts such as federally sponsored research and development in the 1960s, the emphasis on faculty development in the 1970s, and the publication of *A Nation at Risk* all stressed the role of experts and resources in improving faculty teaching practices.

Nearly all of the publications on the scholarship of teaching are aimed at (a) refining the concept with an eye toward multidimensionality, (b) elaborating models to indicate the set of activities that comprise the SoTL, and (c) discussing how universities should facilitate and reward the SoTL. That is, prior work has spent much effort on *what is* the SoTL and little effort on *who* actually practices the SoTL. In general, then, we are unaware of which faculty engage in the SoTL. In this study, I begin to fill the research gap with a focus on the U.S. and examine two of the most salient characteristics that structure faculty dynamics and everyday lives—gender and years of experience. I propose that female faculty are more likely than their male counterparts to engage in the SoTL and that this gender gap increases with experience. To develop these propositions, I use identity theory in tandem with recent data on the characteristics and trends of higher education faculty in the U.S. I test the research propositions with pooled data from surveys conducted in 2001-02 with 85 faculty at a large public research-intensive

¹ Department of Education, Montana State University, Bozeman, MT 59717-2880, cbmyers@montana.edu

university—exactly the type of institution that Boyer (1990) lamented that teaching is given far less weight in tenure and promotion decisions than is research.

II. Background.

A. The Scholarship of Teaching and Learning.

Exactly what comprises SoTL is, of course, debated widely. The discussion generally revolves around the definition of “scholarship” and manifests in two orientations: (a) Teaching-related outputs are the indicator of SoTL; and (b) Teaching and techniques are the indicator of SoTL. The former position was bolstered by Shulman (2000: 50), then president of the Carnegie Foundation, who stated that “[w]e develop a scholarship of teaching when our work as teachers become public, peer-reviewed and critiqued, and exchanged with other members of our professional communities so they, in turn, can build on our work.” Ritchlin (2001) forwards a similar sentiment when arguing that, unfortunately, the scholarship of teaching has become mixed up with the act of teaching itself. The notion that the SoTL is largely a public research- and output-oriented endeavor is further supported by Kreber and Canton (2000) and Richlin and Cox (2004).

The second orientation views the SoTL as teaching-related activities with terms ranging from teaching excellence to expert teaching to scholarly teaching. The term “scholarly teaching” not only involves teaching excellence but also developing expertise in the knowledge base of teaching and learning (Richlin, 2001). Bowden (2007) is especially passionate that scholarly teaching is an indicator of the SoTL. Through lexical statistics and rhetorical analysis, Bowden argues that the original intent of Boyer’s Scholarship of Teaching was the act of teaching itself, which contradicts Shulman’s (2000) view that excellent teaching is not indicative of the SoTL. Others argue that if SoTL is defined simply by peer reviewed research outputs then this further devalues the role of teaching in higher education and that teaching just becomes another research endeavor (Atkinson, 2001).

Boyer did not explicitly settle this debate within his writings and, to date, experts, practitioners, and faculty do not form an exact consensus on a definition of the SoTL (Kreber, 2003, 2005; Nicholls, 2004). A general consensus has emerged, though, where the SoTL is the development of teacher knowledge. As summarized by Trigwell and Shale (2004) and informed by the SoTL models of Trigwell et al. (2000) and Kreber (2002), the development of teacher knowledge includes the communication of ideas and practices, focused reflection, an awareness and use of information on teaching and learning, and the public dissemination of knowledge or pedagogic resonance. Indeed, Kreber (2005: 328) argues that “[t]here seems to be a consensus that the scholarship of teaching ... should be informed by the knowledge of the field, be inquiry-driven, involve critical reflectivity, and include scrutiny by peers.” The last element—scrutiny by peers—need not necessarily occur in journal articles (Richlin, 2001). The notion of “peer review” has been relaxed in more recent writings. Peer review now can occur in less formal contexts such as focused communication with colleagues, faculty development activities, public presentations, and even the Internet (Cambridge, 2000; Nicholls, 2004). The main notion of peer review is that one’s ideas and practices are disseminated in public (Trigwell et al., 2000; Trigwell and Shale, 2004).

In this present article, I examine four activities that dovetail with the general consensus of the SoTL elaborated in the above paragraph. Specifically, I measure the extent to which higher

education faculty (a) Review literature on teaching and learning issues, (b) Talk with colleagues about teaching and learning issues, (c) Consult campus experts on teaching and learning issues, and (d) Use assessment data to inform teaching and learning issues. Together, these four activities get at the heart of the SoTL—inquiry-driven, reflection, and peer communication and scrutiny. Further, these four activities are consistent with Shulman and Hutchings (1999) who argue that the scholarship of teaching requires “going meta” in which faculty frame and systemically investigate questions related to student learning—the conditions under which it occurs, what it looks like, how to deepen it, and so forth—and do so with an eye not only to improving their own classroom but to advancing practice beyond it. Going meta also includes an exchange of ideas and inquiry into student learning.

B. Gender, Theory, and the SoTL.

There exists much research on gender differences in activities within the classroom. In general, female faculty interact with students, use student-centered and collaborative techniques, engage students in higher-order cognitive activities, use diversity in their classes to a greater extent than male faculty, and spend more time preparing for their courses (Park, 1996; Umbach, 2006). Research to date, however, has not examined gender differences in activities outside of the classroom, especially activities indicative of the SoTL.

Given the lack of prior research on gender and the SoTL, we must develop the link theoretically. The theory I use is identity theory. Identity theories, especially Burke's (1991) identity control theory, argue that individuals act in a self-regulatory manner with the goal of matching perceptions of themselves with the standards of the role or situations in which they see themselves. Those who identify with a role feel a strong attraction and commitment to the role, even if the role has a relatively low status. As we form these identities and, often, group associations we develop role-based identities as well—we act to fulfill the expectations of a specific role and emphasize performance and competence in this role (Stets and Burke, 2000). Research shows that the more time we spent in a role and the more we identify with others in the same role the more that role becomes our identity. Further, once a role becomes a salient identity, the more we will work toward being competent in that role and reaffirming that identity (Stryker and Burke, 2000). This relationship is called the *identity-behavior* link (Stryker and Serpe, 1982).

The roles and statuses occupied by female faculty in higher education are, on average, different than from those of their male colleagues. By all accounts, educational data and research show that female faculty are more likely to be in a lower-status teaching role whereas male faculty are more likely to be in a research-oriented role. Data from the most recent *Digest of Education Statistics* and the 1998 *National Study of Postsecondary Faculty* find that female faculty (a) disproportionately hold lower status positions in higher education, (b) are less compensated than men for equal positions, (c) are more likely to be in teaching fields, and (d) spend more time teaching. These gender differences generally increase with years of experience (Harper et al., 2001; NCES, 2006). Females are overly represented in fields that place less emphasis on research and more on teaching, such as those found in humanities and health. Male faculty are overly represented in the natural sciences, fields that place more emphasis on research and grants. Female full-time faculty spent larger shares of their time in teaching or service activities, and smaller proportions in research or administrative activities, than male faculty. The net effect of these differences is that female faculty have access to fewer resources (e.g., office

space, computers), experience greater devaluation of their research, and have more concerns about job stability. As a result gender disparities increase with years experience and rank. Indeed, male faculty are promoted quicker, are more likely to be promoted, receive higher salary increases, and are more likely to be on the “research-track” instead of the “teaching track.”

Applied to higher education faculty, then, identity theory would predict that compared to their male colleagues female faculty will be more likely to identify with a teaching role given that they are more likely to have heavier teaching loads and be on the “teaching track.” This female-male disparity in identities will increase with experience as the teaching gap and track between female and male faculty increases with years of experience. In turn, identity theory would predict that in order to work toward being competent and reaffirming the teaching identity female faculty will spend more time and effort outside of the classroom working on their teaching; a pattern that will intensify with experience. This identity-behavior link, therefore, suggests that female faculty in higher education are more likely to practice the SoTL as they inform and improve their in-class behaviors with external pedagogical resources.

III. Methods.

A. Sample.

The data for this analysis come from formative assessment surveys at a large public university in the Northwest region of the United States. The purpose of the surveys is to provide information to the University’s teaching and learning unit in order to collaborate with participating faculty to help implement faculty goals and activities into the classroom. This unit operates an assessment division that systematically evaluates the use and effectiveness of teaching practices. All faculty who use the University’s centrally-supported online learning environment were invited to participate in the survey process.

This research uses data from 85 faculty who completed these formative survey questionnaires in either fall 2001 or spring 2002. The response rate for participating faculty was 90%. Ancillary analyses show (results available upon request) that there are no consistent patterns of nonresponse among the faculty. A random sample was not feasible because of the nature of the investigation; the data are from a convenience sample.

The present sample brings up two questions about generalizability. First, the 85 faculty are from a single institution. This does limit our ability to generalize to the larger population of college faculty. However, the faculty in the current research are not from selective colleges, disciplines, or departments and do represent a wide cross-section of the university’s overall faculty. Additionally, these faculty are from a 4-year public doctoral university, which according to the 2004 National Study of Postsecondary Faculty is the most common type of institution where faculty with instructional duties reside. Further, faculty at 4-year doctoral institutions also have the more typical faculty appointments that are a mix of teaching, research, and service instead of specializing in just one principal activity (Cataldi, Bradburn, and Fahimi, 2005).

The second issue is the extent to which these data from 2001-2002 still represent and generalize to the contemporary structure and culture of higher education. The most comprehensive and recently published data from the Higher Education Research Institute and their ongoing survey “The American College Teacher” suggest that the teaching and learning culture among faculty and higher education institutions has not changed substantially between 2001-2002 and 2004-2005 (Lindholm, Astin, Sax, and Korn, 2003; Lindholm, Szelényi, Hurtado,

and Korn, 2005). Data from the 2007-2008 faculty survey are not yet available. For example, the percentage of faculty who “agree somewhat” or “strongly” that most students are well-prepared academically was similar for 2001-2002 and 2004-2005 (32% vs. 36%, respectively). Likewise, there are nearly identical results for the percentage of faculty who report that it is “very important” or “essential” to develop a student’s ability to think critically (99% vs. 99%), to enhance a student’s knowledge of other racial and ethnic groups (58% vs. 60%), to provide for a student’s emotional development (36% vs. 38%), to prepare students for responsible citizenship (59% vs. 61%), and to prepare students for graduate or advanced education (58% vs. 61%). Faculty also reported little change in their teaching and evaluation methods. For example, the percentage of faculty who use multiple choice mid-terms and finals in “most” or “all” of their classes was identical for 2001-2002 and 2004-2005 (33% vs. 33%, respectively). Similar comparisons emerged for grading on a curve (17% vs. 19%), cooperative learning (43% vs. 48%), and student evaluations of each other’s work (15% vs. 16%). While some differences did emerge between 2001 and 2005, most of these were small and less than five-percentage points.

Overall, then, the institution from which the current sample is drawn is the same as the “typical” faculty with instructional duties in higher education. Also, given the lack of change in faculty teaching and learning approaches, it does not appear that the 2001-2002 data used for this present research is too dated or irrelevant to speak to today’s higher education faculty and institutions. Of course, the results presented here will be most useful when compared in context to other existing and future studies.

B. Variables.

Table 1 includes an inventory of all variables used in the analyses. The dependent variables approximating the Scholarship of Teaching were created from the following question: “To what extent do you engage in the following activities to inform your teaching practices?” The activities addressed were: (a) Review literature on teaching and learning issues, (b) Talk with colleagues about teaching and learning issues, (c) Consult campus experts on teaching and learning issues, and (d) Use assessment data to inform teaching and learning issues. The possible responses and coding of the responses are as follows: 1 = Never, 2 = Hardly ever, 3 = Seldom, 4 = Sometimes, 5 = Often, and 6 = Very often. In the analyses that follow, I examine each of these SoTL activities individually. I also created an index of the scholarship of teaching through confirmatory factor analysis. Using varimax rotation procedures, the analysis found that all four individual resources loaded on a single dimension, suggesting that these four resources can be combined conceptually and analytically because they are all part of a coherent latent variable. The load factors range from 0.68 to 0.84. To create the Scholarship of Teaching Index and make its metric equivalent to the individual teaching variables, I summed and averaged the four teaching resource variables where higher scores represent greater use of the resources (range: 1 – 6). The final index has a Chronbach’s reliability coefficient of 0.77. Two independent variables are used in the regression analyses. Gender is dummy coded where 0 = Male and 1 = Female. Years teaching is coded using the response categories on the original surveys where 1 = First course, 2 = 1 – 2 years, 3 = 2 – 5 years, 4 = 5 – 10 years, and 5 = Over 10 years.

Table1. Inventory of All Variables Used in the Analyses.

Variable	M	S	Coding / Range
<i>Dependent Variables</i>			
Review literature on teaching and learning issues	4.29	1.09	1 = Never; 2 = Hardly ever; 3 = Seldom; 4 = Sometimes; 5 = Often; 6 = Very often
Talk with colleagues about teaching issues	5.03	0.53	Same as above
Consult campus experts on teaching and learning issues	3.84	0.70	Same as above
Use assessment data to inform teaching practices	4.19	1.15	Same as above
Scholarship of Teaching Index	4.36	0.77	Same as above
<i>Independent Variables</i>			
Female faculty	59%	---	0 = Male; 1 = Female
Years teaching	3.40	1.76	1 = First course; 2 = 0 – 2 years; 3 = 2 – 5 years; 4 = 5 – 10 years; 5 = Over 10 years
<i>Control Variables</i>			
Upper-division course	40%	---	0 = Lower division; 1 = Upper division
Course for majors	16%	---	0 = Nonmajor; 1 = Major
Natural science course	9%	---	0 = Other science; 1 = Natural science

The original surveys constrain the extent to which I am able to control for additional faculty and course characteristics. Clearly not all faculty within the same university would be similar in their teaching expectations and activities. The data do include a limit set of variables in which teaching practices may differ. Research finds that teaching practices and expectations differ by department or academic field, the level of the course (e.g., lower or upper division), and whether the course is elective or required for the major (Colbeck, Cabrera, and Terenzini 2001; Kuh and Hu 2001; Paulsen and Wells 1998). To this end, I include three variables as controls. Academic field was measured on the original survey with six categories: natural science, humanities, social science, general education, business, and the arts. Research shows that differences in teaching and research expectations and behaviors are most between the natural sciences and all other fields (NCES, 2006). Thus, I code academic field as a simple dichotomous variable where 1 = natural science and 0 = all else (i.e., humanities, social science, general education, business, and the arts). The level of instruction is also captured by a dichotomous variable where 1 = upper-division (generally advanced junior and senior courses) and 0 = lower-division (generally introductory freshman and sophomore courses). Lastly, course composition is divided between those required for a major (coded 1) and those that are electives (coded 0).

C. Analytical Strategy.

To evaluate the research question I estimated two ordinary least squares (OLS) regression models for each of the four dependent variables and the SoTL Index. In the first model, the dependent variable was regressed on gender, years teaching, and the control variables. In the second model, a gender $\times$ years teaching interaction term was added to regression equation to determine if any effects of gender vary by experience. There are two possible statistical concerns that arose with using OLS regression. First, the dependent variables were measured with an ordinal-level scale. To test the effects of scaling on the regression estimates, all of the equations were re-estimated with ordered logistic regression models. The results were identical to those for the OLS analyses. Second, because five regression analysis were conducted (i.e., one for each dependent variable), there is an increase in the probability of making a Type I error. To estimate this probability, the models were estimated with a MANOVA procedure and with a simultaneous regression equation (i.e., all dependent variables). Again, the results were identical to those for the OLS analyses. In the end, I opted to use the results from the OLS analyses as they are more familiar than results from ordered logistic or simultaneous regression.

IV. Results.

Looking across the SoTL activities in Table 1, the averages suggest that faculty vary in their engagement. Faculty are most likely to inform their teaching practices by talking with colleagues about teaching issues; they do so “often.” On the other hand, the least engaged activity is consulting with campus experts on teaching and learning issues. They use this resource more than “seldom” but less than “sometimes.” Faculty report that they use the other two resources—reviewing literature on teaching and learning issues and using assessment data to inform teaching practices—only about “sometimes.” As a single construct, the SoTL Index shows that faculty engage in these activities as a group somewhere between “sometimes” and “often.” The remainder of the descriptive statistics provides a snapshot of the sample. About 60% of the sample is female faculty who have been teaching about 5 years. For all faculty, 40% were teaching an upper-division course mainly for nonmajors and not in the natural sciences.

The results from the regression analyses are in Table 2. The coefficients for Gender in Model 1 across the dependent variables return four statistically significant results. Clearly, female faculty are much more likely to engage in SoTL activities. Compared to male faculty, female faculty are significantly more likely to review literature, talk with colleagues, and consult campus experts in order to inform their teaching practices. As a result of these trends, female faculty are statistically more likely to score higher on the SoTL Index. These results support strongly the first research hypothesis that female faculty are more likely to engage in the SoTL.

The substantive effects of gender are moderate-to-large—all are greater than one-half of a unit in the teaching resource. For example, for reviewing literature, a gender coefficient of 0.80 suggests that if male faculty “sometimes” inform their teaching practices by reviewing literature (value = 4), then women faculty’s resource use approach “often” (where $4 + 0.80 = 4.80$, nearly a value of 5 or “often”). The effect of gender on talking with colleagues about teaching issues is equally large at 0.82, whereas the effect of gender on consulting with campus experts is smaller (0.62), as is the effect of gender on the Teaching Resource Index (0.65). The results show that, in general, teaching experience has little direct influence on the extent to which faculty use these resources to inform their teaching. The only significant result is for talking with colleagues about

Table 2. Unstandardized OLS Regression Coefficients from the Regression of Engaging in SoTL Activities on Faculty Gender and Years Teaching, 2001-2002.

Variables	SoTL Activities									
	Review literature on teaching and learning issues		Talk with colleagues about teaching issues		Consult campus experts on teaching and learning issues		Use assessment data to inform teaching practices		Scholarship of Teaching Index	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Gender (1 = female)	0.80*** (0.24)	1.00*** (0.43)	0.82*** (0.22)	0.78*** (0.15)	0.62** (0.24)	0.79** (0.27)	0.37 (0.24)	0.21 (0.29)	0.65*** (0.15)	0.69*** (0.18)
Years teaching	0.05 (0.05)	-0.09 (0.07)	-0.17** (0.06)	-0.28*** (0.08)	0.04 (0.05)	-0.04 (0.12)	0.07 (0.05)	0.09 (0.13)	0.01 (0.03)	-0.05 (0.04)
Sex x Years teaching	---	0.33** (0.11)	---	0.24* (0.11)	---	0.29** (0.12)	---	0.10 (0.11)	---	0.22** (0.08)
Upper-division course	-0.58** (0.21)	-0.58** (0.16)	-0.18 (0.16)	-0.18 (0.15)	-0.05 (0.23)	-0.06 (0.23)	-0.19 (0.24)	-0.18 (0.25)	-0.25* (0.09)	-0.25* (0.10)
Course for majors	-0.25 (0.29)	-0.22 (0.28)	-0.42* (0.19)	-0.46** (0.18)	-0.90** (0.28)	-0.88** (0.27)	-0.14 (0.29)	-0.15 (0.30)	-0.44* (0.19)	-0.43* (0.19)
Natural science course	-0.73** (0.30)	-0.81** (0.29)	-0.22 (0.19)	-0.19 (0.18)	-0.92** (0.17)	-0.94** (0.19)	-0.03 (0.11)	-0.02 (0.09)	-0.83** (0.27)	-0.86** (0.26)
Intercept	4.66	4.67	5.67	4.93	4.24	4.31	4.81	5.10	4.84	4.75
R ²	0.25	0.35	0.31	0.38	0.26	0.32	0.05	0.05	0.29	0.35
% change in R ²	---	40.0	---	22.6	---	23.1	---	0.0	---	20.7

Notes: Numbers in parentheses are standard errors.

*p < 0.05. ** p < 0.01. ***p < 0.001 (two-tailed)

teaching issues where greater teaching experience leads to less communication with colleagues. Even though the control variables were not theoretically included in the framework, several interesting trends emerged that are noteworthy. Faculty are more likely to engage in the SoTL in courses for nonmajors than in courses for majors and if they are in fields other than the natural sciences. Finally, in all but one of the models, the variables explain an impressive 25 – 31% of individual differences in the extent to which faculty engage in the SoTL.

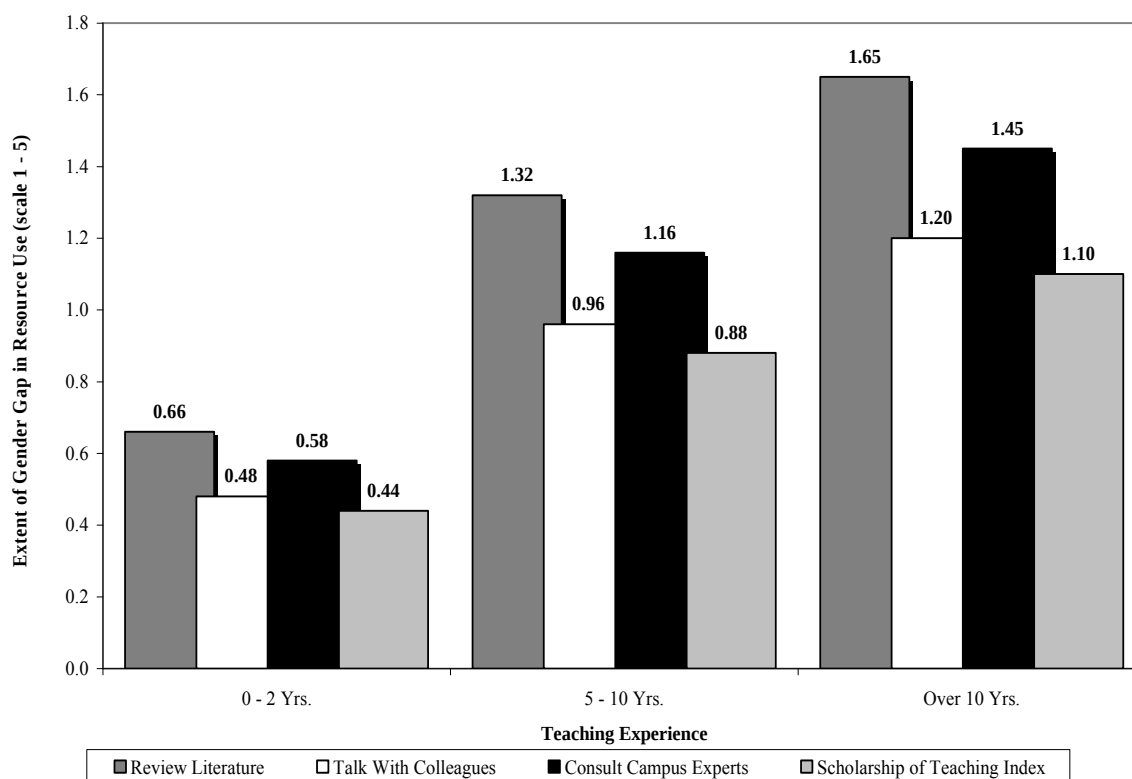
The results from the interaction equations in Model 2 support the second research hypothesis. Across the five dependent variables, four of the *Gender x Years teaching* interaction terms reach statistical significance implying that the positive association between gender and the SoTL activities (from Models 1) increases significantly with more years of experience. Specifically, the gap between female and male faculty widens with experience on reviewing literature, talking to colleagues, consulting campus experts, and the SoTL Index. Just as gender is not important to the use of assessment data, the interaction term also fails to reach statistical significance for this activity. The importance of teaching experience in the gender gap is captured by the significant increases in R-squared—between 21% and 40%.

How large are these interaction effects? The advantage of regression analysis is that interactions have a particular interpretation when one of the predictors is a categorical variable with two categories—such as Gender. In the interaction models in Table 2 (ignoring the control variables, which are generally used for prediction), the regression equation is equal to:

$Y = \alpha + b_1(\text{Gender}) + b_2(\text{Years teaching}) + b_3(\text{Gender} \times \text{Years teaching})$. The inclusion of the interaction term in the model allows the regression lines for men and women to have different slopes. For men (who are coded “0”), the model reduces to $Y = \alpha + b_1(\text{Years teaching})$, whereas for women (who are coded “1”) it is $Y = (\alpha + b_2) + [(b_1 + b_3)(\text{Years teaching})]$. Thus, b_3 is the gender difference in the effect of years teaching. That is, it estimates how the gender gap in using teaching resources increases by teaching experience.

Figure 1 graphically shows how the gender gap in SoTL activities widens with more years of teaching experience. For illustration, I present results at three experience levels: 0 – 2 years (coded “2”), 5 – 10 years (coded “4”), and over 10 years (coded “5”). These three levels roughly correspond to faculty who are just beginning their careers, faculty who are near-tenure or recently tenured, and faculty with the most post-tenure experience, respectively. As we can see in Figure 1, the gender gap is small for faculty with 0 – 2 years of teaching experience—ranging from 0.44 for the SoTL Index to 0.66 for Reviewing Literature. This gender gap widens significantly as faculty gain more teaching experience. At 5 – 10 years of teaching experience, the gap ranges from 0.88 for the SoTL Index to 1.32 for Reviewing Literature. Then with over 10 years of teaching experience the gap widens to 1.10 for SoTL Index, 1.20 for Talking with Colleagues, 1.45 for Consulting Campus Experts, and 1.65 for Reviewing Literature. Substantively, any gap that is 1.0 or greater represents a one-unit change in the engagement in SoTL activities, such as moving from “Sometimes” to “Often.”

Figure 1. Differences in the Use of The Scholarship of Teaching Between Female and Male Faculty by Teaching Experience



V. Discussion.

The purpose of this study was to provide an initial analysis of the extent to which higher education faculty engage in the Scholarship of Teaching; the focus was on gender and experience. The results do show that faculty differ in their levels of engagement in SoTL activities. Specifically, (a) female faculty are significantly more likely than male faculty to engage in the SoTL, and (b) this gender gap increases significantly with more teaching experience. The sample used in this study include faculty involved with the university's teaching and learning center. This suggests that these faculty may be more likely to engage in the SoTL. However, there is no a priori reason to anticipate the direction of the current findings. Further, the findings remain after controlling for type of course, type of students, and type of science. Because the faculty were from the same university the results are not a by-product of different types of higher education institutions (e.g., doctoral v. four-year). This study informs prior trends. Specifically, we already knew that female faculty, on average, teach more classes and are more likely to be on the teaching track and in teaching departments (NCES, 2006). Now we also know that female faculty, especially those experienced, are more likely to engage in the SoTL.

The implications of these results for student learning are important in two ways. First, based on the current sample, female faculty are more likely to approach teaching and learning reviewing the scholarly and pedagogy literature, discuss their ideas and experiences with other faculty and colleagues, and to consult and interact with experts. These activities generally

correspond to the SoTL model presented by Trigwell and Shale (2004), especially the role of communicating and making public one's experiences, understanding, and knowledge base of teaching. The greater use of experts by female faculty especially gets at Shulman's (2000) notion of "peer review." Second, therefore, if specific faculty are more likely to engage in the SoTL, this suggests that there are patterns in higher education classes where certain faculty, on average, create better learning contexts. According to Kreber (2002), these faculty who increasingly engage in the SoTL are not only excellent teachers but also expert teachers.

The benefits of the SoTL on student learning have been lauded in numerous publications (e.g., Trigwell and Shale, 2004). Yet, I argue that the actual benefits of the SoTL may be slightly underestimated. It is clear, on average, that teaching in higher education institutions—especially research institutions—is a devalued pursuit and a harder path to tenure and promotion (Braxton, Luckey, and Helland, 2002; Park, 1996). Given this, faculty who invest considerable time and energies into teaching vis-à-vis the SoTL may be especially driven to create outstanding teaching and learning contexts above and beyond those theoretically suggested by the SoTL. Indeed, as argued by Kreber (2000) and Palmer and Collins (2006), faculty who engage in the SoTL will need to be especially motivated and be driven by an intrinsic desire and reward structure given the low levels of external rewards and recognition. My results support this line of reasoning: faculty who are committed to the SoTL early in their career become increasingly committed to the SoTL throughout their career. That is, the SoTL becomes part of their *identity* and a core professional value (Nicholls, 2004).

Further research could expand and strengthen this study in several important ways. First, the present sample was limited to a single institution and a size under 100 faculty. Future studies should draw or analyze a national sample and, as important, a size that allows for the creation of subgroups. For example, it would be interesting to compare faculty within the same discipline or science or across different teaching responsibilities and loads. Second, I used identity theory as the organizing framework. Of interest would be the testing of the theory by gathering data on faculty identities. Both theoretical and empirical findings are necessary to further knowledge and understanding of important higher education issues. Explicit theoretical testing allows us to begin to answer the "Why" question of this research: Why do certain faculty engage in the SoTL? Third, the present study used four indicators of the SoTL. Other indicators should be examined to allow a full understanding of faculty differences in pedagogy and teaching-related practices. Finally, student-centered studies are needed to understand how and if students recognize the SoTL and their assessment of such practices.

by

References

- Atkinson, M. P. (2001). The scholarship of teaching and learning: Reconceptualizing scholarship and transforming the academy. *Social Forces*, 79(4), 1217-1230.
- Bowden, R.G. (2007). Scholarship reconsidered: Reconsidered. *Journal of the Scholarship of Teaching and Learning*, 7, 1-21.
- Boyer, E.L. (1990). *Scholarship Reconsidered: Priorities of the professoriate*. The Carnegie Foundation for the Advancement of Teaching. New York: John Wiley and Sons.

Braxton, J. M., Luckey, W., and Helland, P. (2002). Institutionalizing a Broader View of Scholarship Through Boyer's Four Domains (ASHE-ERIC Higher Education Report, 29, 2). San Francisco: Jossey-Bass.

Burke, P. J. (1991). Identity processes and social stress. *American Sociological Review*, 56, 836-849.

Cambridge, B.L. (2000). The scholarship of teaching and learning: A national initiative. In M. Kaplan and D. Lieberman (Eds.), *To Improve the Academy*, 56-68. Bolton, MA: Anker.

Colbeck, C. L., Cabrera, A. F., and Terenzini, P. T. (2001). Learning professional confidence: Linking teaching practices, students' self-perceptions, and gender. *The Review of Higher Education*, 24, 173-191.

Cross, K. P., and Steadman, M. H. (1996). Classroom Research: Implementing the scholarship of teaching. San Francisco: Jossey-Bass.

Harper, E. P., Baldwin, R. G., Gansneder, B. G., and Chronister, J. L. (2001). Full-time women faculty off the tenure track: Profile and practice. *The Review of Higher Education*, 24, 237-257.

Cataldi, E.F., Bradburn, E.M., and Fahimi, M. (2005). 2004 National Study of Postsecondary Faculty: Background characteristics, work activities, and compensation of instructional faculty and staff: Fall 2003 (NCES 2006-176). Washington, DC: National Center for Education Statistics

Kreber, C. (2000). Becoming an expert university teacher: A self-directed process. In H. Long (Ed.), *Practice and Theory in Self-Directed Learning*, 131-142. Schaumburg, IL: Motorola University.

Kreber, C. (2002). Teaching excellence, teaching expertise, and the scholarship of teaching. *Innovative Higher Education*, 27, 5-23.

Kreber, C. (2003). The scholarship of teaching: A comparison of conceptions held by experts and regular academic staff. *Higher Education*, 46, 93-121.

Kreber, C. (2005). Reflection on teaching and the scholarship of teaching: Focus on science teachers. *Higher Education*, 50(2), 323-359.

Kreber, C., and Cranton, P.A. (2000). Exploring the scholarship of teaching. *Journal of Higher Education*, 71(4), 476-495.

Kuh, G. D., and Hu, S. (2001). The effects of student-faculty interaction in the 1990s. *The Review of Higher Education*, 24, 309-332.

Lindholm, J. A., Astin, A., Sax, L. J., and Korn, W. S. (2002). *The American College Teacher: National norms for 2001-2002 HERI Faculty Survey*. Los Angeles: Higher Education Research Institute, UCLA.

Lindholm, J.A., Szelényi, K., Hurtado, S. and Korn, W.S. (2005). *The American College Teacher: National norms for the 2004-2005 HERI Faculty Survey*. Los Angeles: Higher Education Research Institute, UCLA.

National Center for Educational Statistics. (2006). *Digest of Education Statistics, 2005*. Washington, DC: U.S. Department of Education.

Nicholls, G. (2004). Scholarship in teaching as a core professional value: What does this mean to the academic? *Teaching in Higher Education*, 9, 29-42.

Park, S. M. (1996). Research, teaching, and service: Why shouldn't women's work count? *Journal of Higher Education*, 67, 46-84.

Paulsen, M. B., and Wells, C. T. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education*, 39, 365-384.

Richlin, L. (2001). Scholarly teaching and the scholarship of teaching. *New Directions for Teaching and Learning*, 86, 57-68.

Richlin, L., and Cox, M. D. (2004). Developing scholarly teaching and the scholarship of teaching and learning through faculty learning communities. *New Directions for Teaching and Learning*, 97,

Shulman, L.S. (2000). From Minsk to Pinsk: Why a scholarship of teaching and learning? *The Journal of Scholarship of Teaching and Learning*, 1, 48-53.

Shulman, L.S., and Hutchings, P. (1999). The scholarship of teaching: New elaborations, new developments. *Change*, 31(5)

Stets, J. E., and Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, 63, 224-237.

Stryker, S., and Burke, P. J. (2000). The past, present, and future of identity theory. *Social Psychology Quarterly*, 63, 284-297.

Stryker, S., and Serpe, R. T. (1982). Commitment, identity salience, and role behavior: A theory and research example. In W. Ickes and E. S. Knowles (Eds.), *Personality, Roles, and Social Behavior*, 199-218. New York: Springer-Verlag.

Trigwell, K., Martin, E. Benjamin, J., and Prosser, M. (2000). Scholarship of teaching: A model. *Higher Education Research and Development*, 19(2), 155-168.

Trigwell, K, and Shale, S. (2004). Student learning and the scholarship of university teaching. *Studies in Higher Education*, 29, 523-536.

Umbach, P. D. (2006). The contribution of faculty of color to undergraduate education. *Research in Higher Education*, 47(3), 317-345.

Putting theory into practice: a professional development school/university co-teaching project

W. Corry Larson and Abigail J. Goebel¹

Abstract: This investigation measured the impact of co-teaching on pre-service teachers' sense of efficacy in classroom management and student engagement. The study utilized a Professional Development School partnership between a university and an elementary school to make a theory-to-practice connection for pre-service teachers enrolled in an applied behavior analysis course. Instruction was delivered by a university professor and an elementary school teacher using collaborative consultation. Changes in pre-service teachers' attitudes were measured by a pre/post administration of the Teachers Sense of Efficacy Scale (TSES) (Tschannen-Moran and Wolfolk, 2001). Results suggested that the pre-service teachers' sense of efficacy increased during the course.

Keywords: Co-Teaching, Professional Development School, Classroom Management, Self-Efficacy

I. Introduction.

Interest in school consultation and collaboration originated at least as early as 1970 (Caplan, 1970), and although research on school consultation has continued apace (see Dettmer, Thurston and Dyck, 2005, pp. 43-46 for a concise history of the evolution of consultation in schools), relatively little attention has been paid to research that has actually brought together university faculty with PreK-12 (preschool to 12th grade) teachers to co-teach university school of education courses. While much has been written describing what such co-teaching projects should look like (see, for instance, the National Council for Accreditation of Teacher Education, Standard III: Collaboration, 2001, or Addal-Haqq, 1988), actual data resulting from research projects that have employed university faculty and PreK-12 teachers in co-teaching pre-service courses has not been common. (For some exceptions to this scarcity see Evans, 1996, and Dallmer and Baker, 2002) There appear to be several reasons for this paucity of published research.

First, the pressure for schools of teacher education to collaborate with PreK-12 schools is so recent, that little research has yet to make its way into the literature (V. H. Pilato, Director, Teacher Quality, Maryland State Department of Education; N. Allen, Program Approval Specialist, Maryland State Department of Education; personal communications, September 9, 2005). It might be noted that early efforts to link schools of teacher education and PreK-12 schools by means of the Professional Development School model did not appear until the late 1980s in the first and second Holmes Group Report (Isher, 1995). It was not until the third Holmes Group Report's call in 1995 for raising standards for teacher education by increasing "...the numbers of university faculty who are as at home working in the public schools as on the university campus..." (Isser, 1995, p. 1), and by suggesting that "[B]oard-certified teachers and

¹ Special Education Program, University of Maryland Eastern Shore; Elementary Special Education Teacher, Wicomico (Maryland) Public Schools

other qualified practitioners will join these faculties as colleagues in conducting important research and in better educating the nation's educators...." (Isser, 1995, p. 2) that the idea of PreK-12 teachers and university faculty co-teaching courses began to firmly take root in the firmament of schools of teacher education.

Second, there are several studies in the literature that suggest that teachers resist participating in education and educational psychology research. McBee (2004) reports that the small body of literature that exists describing PreK – 12 teacher attitudes towards educational research finds that "... (T)eachers unfortunately find much of existing research to be inaccessible, irrelevant to their daily experiences in the classroom, or counterintuitive" (p. 1). McBee (2004) also argues that "... (I)n addition to the aforementioned factors of perceived irrelevance of research and inaccessibility, the exclusion of teachers' voices from the research literature, educational policy decisions, and curriculum decisions, may lead teachers to distrust 'outsiders'" (p. 2). Perhaps, as McBee (2004) notes,

... one reason that teachers find research irrelevant is that almost none of the educational research is written by practicing teachers. The teachers are considered as objects of study and as implementers of research results, but the actual concerns, questions, and perspectives of teachers are conspicuously absent.

Since much educational research is written by and for academic researchers, it is easy to understand why classroom teachers may not find it relevant or accessible (p. 2).

Third, Bryan (2001) states that although

... the National Council for Accreditation of Teacher Education (1997), in its Draft Standards for Professional Development Schools, emphasized the integration of collaborative, practice-based inquiry within PDS, as well as the importance of school and university faculty learning to work together to produce research that enhances student learning and improves the organizational environment... PDS research in general has been slow to emerge and is underemphasized in relation to other PDS goals such as pre-service and inservice teacher development (p. 1).

Furthermore, "... (T)he National Council for Accreditation of Teachers Education standards for accreditation of teacher education programs does not list basic familiarity with statistics or experimental design as a goal for teacher training programs...." (McBee, 2004, p.2). Basic familiarity with statistics or experimental design might serve to make research studies more accessible and relevant to PreK-12 teachers. It might also encourage teachers to become better consumers of research.

Finally, reports of research on co-teaching projects involving PreK–12 teachers and university faculty may be scarce because of cultural differences between PreK–12 and faculty. Bryan (2001) states that the incentives

... for participating in collaborative PDS research are different for teachers and university faculty because the two groups represent distinct and, at times, clashing cultures with different norms, roles, and expectations. For university faculty, the incentive of publication plays a central role in promotion, tenure, and merit decisions and is likely to constitute a major factor affecting their willingness to participate and sustain involvement in PDS collaborative research. In addition, PDS research represents a form of professional development for university faculty and enables them to use field-based methodologies (p. 2).

However, as Bryan (2001) continues, "... publication [of research] does not affect promotion, tenure, and merit decisions of (Pre – K) teachers and other public

school personnel” (p. 2).

If improvement in both PreK-12 schools and schools of education is to be expected, “[A] research agenda must be jointly established to improve teaching and learning in the schools. To be meaningful and useful, educational research should be designed in collaboration with K-12 teachers, integrated into schools over time and address real classroom issues.” (National Association of State Boards of Education, 1994, p. 2). As this report continues,

“Guidelines and incentive systems should promote exchanges between higher education faculty and teachers for the continuous professional development of all. Higher education/public school partnerships that involve co-teaching and incentives for interaction and exchange among education program faculty and classroom teachers provide excellent professional development opportunities” (p. 1).

In line with these concerns and suggestions, the authors decided to apply for permission to conduct an investigation into the effect co-teaching might have on a university teacher education program. After considering the mandates for collaboration that current models of Professional Development Schools place on school-university partnerships, the authors decided to petition the first author’s school of education to allow both authors to co-teach a university course. Since the second author pointed out that her own teaching experience, as well as an extensive literature review that she conducted (see Mitchell, 2004; Orr, Thompson and Thompson, 1999, and Giallo and Little, 2003) indicated that classroom management is of critical concern to most new teachers, the authors requested to co-teach an applied behavior analysis course. Permission was granted, and planning for the co-taught course began immediately. All Institutional Review Board policies of the university were followed, as were confidentiality and privacy policies of the school district in which the second author taught.

II. Method.

A. Participants.

The participants were eight individuals who were enrolled in a small (total student enrollment: approximately 3800), Historically Black College or University (HCBU) in the Eastern United States, and enrolled in an applied behavior analysis undergraduate/graduate course. The course was required for certification and graduation from the University’s Special Education program. The Special Education program was, and still is, accredited by the National Council for Accreditation of Teacher Education (NCATE). All students enrolled in this course had been admitted to the University’s Teacher Education Program, which required completion of undergraduate general education requirements, with a 2.75 grade point average. Two of the students in the course were African-American, and the rest were Caucasian. Six were undergraduate students (three juniors, and three seniors); five of these undergraduates were females and one was male. Two of the participants were graduate students, both female, and both Caucasian. The course included a field experience placement in a public school for 15 hours in addition to class meetings. Six of the eight students in the course completed their field experiences in the second author’s elementary school.

B. Materials.

Using a pretest/posttest design, The Teacher Sense of Efficacy Scale (TSES) (Tschannen-

Moran and Wolfolk, 2001) was administered at the beginning and at the end of the course. Research evidence (Giallo and Little, 2003) suggests a correlation between a pre-service teacher's sense of efficacy in classroom management and successful implementation of these management techniques.

Bandura (1986) defines efficacy as "...a judgment of one's capability to accomplish a certain level of performance" (cited in Emmer and Hickman, 1991, p. 755). Emmer and Hickman (1991) assert that self-efficacy is a more specific construct than self-concept or self-esteem because it describes the individual's self-conception of performance capability rather than a more global self-evaluation (p. 756). Giallo and Little (2003) suggest that a significant relationship exists between an individual's sense of self-efficacy and confidence in behavior management (p. 29).

The TSES was chosen because it had been used in a similar investigation (Tschannen-Moran and Wolfolk, 2001). This scale measures three components of individuals' sense of efficacy towards three moderately correlated factors: Efficacy in Student Engagement, Efficacy in Instructional Practices, and Efficacy in Classroom Management (Tschannen-Moran and Wolfolk, 2001). In addition to administering the scale, deriving scores for the entire test, and examining differences in pre- and posttests, the authors examined the results of the Efficacy in Classroom Management and Efficacy in Student Engagement subscales for differences between pre- and posttest scores. Scores from The Efficacy in Instructional Practices subtest were not examined because the authors felt that the results from this subtest did not bear sufficiently on the purposes of the investigation.

Other materials used in the project included two textbooks the students were required to purchase (Miller, 1997, and Wolfgang, 1999), and miscellaneous teaching materials brought in by the second author. Students also used a logbook to record their observations in the second author's school, and their reflections.

C. Design and Procedures.

Considering McBee's (2004) suggestion that PreK-12 teachers often do not feel included or consulted as equals by university faculty conducting research projects, the authors felt that the choice of co-teaching model was important. In considering the six most common types of school consultation, that is, Triadic, Stephens/Systems, Vermont Consulting Program, School Consultation Committee, Resource/Consulting Teacher Program, and Collaborative Consultation (Dettmer, Thurston and Dyck, 2005), the authors decided that the Collaborative Consultation model best fit requirements for collegiality between university and PreK-12 teachers. This model also seemed to be designed to help assure that the research resulting from the investigation would be relevant and accessible to PreK-12 teachers.

The Collaborative Consultation model is derived from Tharp and Wetzel (1969) and Tharp (1975) and includes three components – (C)onsultant, (C)onsultee/(M)ediator, and (T)arget (Dettmer, Thurston and Dyck, 2005). Dettmer, Thurston and Dyck (2005) conceptualize the consultant and consultee/mediator as equal partners with diverse experience. Communication is not hierarchical or one-way. Rather, there is a sense of parity that blends the skills and knowledge of both consultant and consultee/mediator, with disagreements viewed as opportunities for constructive extraction of the most useful information (Dettmer, Thurston and Dyck, 2005, p.57). In this investigation, the authors decided that both would hold and exchange roles as consultant and consultee/mediator, as situations warranted, and the students in the course

would be the clients, or targets.

Dettmer, Thurston and Dyck (2005) quote Pryzwansky (1974) as suggesting that the basic structure of the collaborative approach emphasizes the need for mutual consent on the part of both consultant and consultee/mediator, mutual commitment to the objectives, and shared responsibility for implementation and evaluation of the plan. The consultant, consultee/mediator, and target have reciprocally reinforcing effects on one another. Each collaborator, as part of the team, contributes a clearly defined portion of the effort so that all comes together to create a complete plan or solution (Dettmer, Thurston and Dyck, 2005, p. 57).

The authors, as co-teachers, served as both the consultant and consultee/mediator. These roles were often exchanged, as described by Dettmer, Thurston, and Dyck, (2005) in their definition of collaborative consultation. One of the co-teachers, the first author, taught in the special education program of the university in which the applied behavior analysis course was taught. The other co-teacher, the second author, taught special education at a local elementary school. Both instructors agreed to, and attended, weekly planning meetings to outline instructional goals, review the current course syllabus in order to delineate tasks and define scope of teaching responsibilities, make any necessary changes in the course structure, and grade student projects/exams. The instructors felt that these weekly planning sessions were an essential component to support the co-teaching model as well as to give feedback, brainstorm, and guide instruction throughout the course.

In addition, although a possibly confounding variable might be introduced, the authors agreed that the existing format for the required school observation was not providing as adequate a theory-to-practice connection as might be hoped, and sought to change it. Informal conversations with pre-service teachers who had completed school observations appeared to indicate that such required observations were somewhat useless, since cooperating teachers rarely gave observing students responsibility at any level. These students felt that they had no direction from cooperating teachers, and often felt left out of any school activities. Traditionally, each pre-service teacher was assigned one cooperating teacher for the required school observation. Assignment was random, and often did not result in successful, theory-to-practice experience for the pre-service teacher, according to the students. Given the dissatisfaction students seemed to experience with the school observation, the authors considered a possible redesign of the traditional role the cooperating teacher plays during pre-internship school observation.

Under this redesign, most of the pre-service teachers in the course were placed at the same school. At this school, they all had access to the same mentor teacher, who also was one of the co-teachers in the university course the students took. This meant that most students had access to the mentor teacher at the observing school as well as during the course at the university, and this teacher could provide immediate feedback and additional resources to students in both settings.

The second author consulted with, and got permission from, the school administration to place and mentor the students from the behavior analysis course, assuring him that her participation in this project would not interfere with her instructional responsibilities. The second author also obtained cooperation from the faculty and staff at her school. She obtained this cooperation so that the students could observe a variety of instructional situations and classroom management styles.

The university in which the course was conducted was a Professional Development School (PDS). As defined by the Maryland State Department of Education, a Professional Development

School (PDS) is

... a collaboratively planned and implemented partnership for the academic and clinical preparation of interns and the continuous professional development of both school system and institution of higher education (IHE) faculty. The focus of the PDS partnership is improved student performance through research-based teaching and learning. (Maryland State Department of Education, 2003, p. 3).

The Professional Development School model is intended to improve student performance at both the IHE and the PreK-12 school. It is intended that students in schools of education benefit by a closer exposure to actual teaching experiences and that PreK-12 schools benefit by exposure to current educational research in teaching interventions (Maryland State Department of Education, 2003).

The course, which met once per week in 2-3/4 hour sessions, was divided into two segments per session. The first section was taught by the first author, and presented applied behavior analysis theory and classroom management models. The second section, taught by the second author, was a weekly reinforcement section designed to highlight concepts introduced the previous week by utilizing resources used in local school districts. In this section the second author facilitated a "Classroom Connection" discussion for the last thirty minutes of the course period so students could discuss field placement observations, ask additional questions, and role-play discipline situations. The instructors believed that having most of the students in the course observing at the same school, in similar situations, encouraged more depth in these course discussions, with more authentic learning experiences.

III. Results.

The pre-service teachers' attitudes, as measured by pre/post test administration of the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran and Wolfolk, 2001) suggested that students developed a greater confidence in their ability to handle classroom management situations over the course of the course. Total Scale score means increased from 153.6 ($SD = 15.9$) in the pretest to 174.5 ($SD = 22.3$), a difference of 20.8 points, an increase in scores suggesting an increase in feelings of efficacy, and therefore increases in confidence.

Pretest/posttest difference scores suggested that students' confidence also may have improved in the factors measured by the Efficacy in Classroom Management and Efficacy in Student Engagement subscales of the TSES. Aggregated student scores in the Efficacy in Classroom Management subscale went from a mean of 53.5 ($SD = 7.1$) in the pretest to a mean of 59.8 ($SD = 7.2$) in the posttest. The Efficacy in Student Engagement subscale aggregated pretest scores went from a mean of 50.2 ($SD = 4.9$) to a mean of 56.9 ($SD = 6.7$). As stated previously, difference scores were not calculated for the Efficacy in Instructional Practices subscale because the authors felt that scores from this subscale did not bear directly or substantially on this investigation.

Table 1. Results from Total Score Scale pre-test/post-test administration of Teachers' Sense of Efficacy Scale (TSES).

Total Score Scale	<u>M</u>	<u>SD</u>	<u>n</u>
Pre-test	153.6	15.9	8
Post-test	174.5	22.3	8

Table 2. Results from Efficacy in Classroom Management subscale pre-test/post-test administrations of Teachers' Sense of Efficacy Scale (TSES).

Efficacy in Classroom Management Subscale	<u>M</u>	<u>SD</u>	<u>n</u>
Pre-test	53.5	7.1	8
Post-test	174.5	22.3	8

Table 3. Results from Efficacy in Classroom Engagement subscale pre-test/post-test administrations of Teachers' Sense of Efficacy Scale (TSES).

Efficacy in Classroom Management Subscale	<u>M</u>	<u>SD</u>	<u>n</u>
Pre-test	50.2	4.9	8
Post-test	56.9	6.7	8

IV. Discussion.

There is a great deal of evidence in the literature that classroom management is a major concern of new teachers (see for example Johannessen, 2004; Mitchell, 2004; Orr and Thompson, 1999, and Giallo and Little, 2003). This level of concern suggests a lack of confidence on the part of these teachers in their ability to successfully handle classroom management situations.

One means to measure the confidence a pre-service teacher has in his or her ability to handle classroom management is to use the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran and Wolfolk, 2001). In the present investigation, the authors employed this scale in a pretest/posttest design to examine any change in this sense of efficacy over the course of a university course in applied behavior analysis, taught by collaborative consultation between a PreK-12 teacher and a university instructor.

The authors conducted data analysis on both the whole-scale level and on the subscale levels. A comparison of pretest scores and posttest scores suggested that the pre-service teachers' overall sense of efficacy, in general, may have increased over the course of the course. Analysis of the subscale data suggested that the pre-service teachers' sense of efficacy may have also increased in classroom management and in student engagement, as measured by the scale.

Factors that may have influenced a change in the pre-service teachers' sense of efficacy in the areas tested included the content and process of the applied behavior analysis course, the co-teaching format, the presence on the co-teaching team of a teacher employed in a local school district and having day-to-day responsibilities and experiences in a public school, the school

observation experience, and the group format of student school observation experience. Some of these factors are common to every course taught in this subject, at this level, and at this university, so exposure to them was not unique. The other factors (the co-teaching format, presence on the co-teaching team of an individual employed as a PreK-12 teacher, and the group format of the pre-service teachers' school observations) are new to this university environment, and may have had influence in the difference scores between pre- and post-administrations of the scale.

There were several limitations to this investigation. The limited number of subjects would preclude firm conclusions from the data. In addition, the course that was taught was novel in the sense that its format had never before been used in this university education department. The need to flexibly change components of the course in light of unanticipated challenges might therefore have influenced the results. Further, the variety of new interventions used in this course (e.g., the co-teaching format, the group nature of the school observations) make it difficult to tease out the influence of the individual components. Finally, there was no control course to provide comparisons.

Future investigation might seek to evaluate the relative importance of the aforementioned components, in order to determine which were important ones and which were relatively unimportant or less powerful ones.

Finally, investigation into PreK-12 changes in teachers' attitudes towards both research and PreK-12 teachers' and university faculty using collaborative consultation might be fruitful.

In conclusion, the establishment of the Professional Development School model, with its emphasis on the improvement of student learning at both the PreK-12 and university level has brought with it the need to evaluate those factors that influence that improvement. University and PreK-12 teachers co-teaching school of education courses is one means of providing opportunities to determine what instructional practices influence that improvement. Replication of this investigation, with larger sample sizes, comparison across courses that do, and do not, use co-teaching, and better control of confounding variables might provide better insight into which components influence pre-service teachers' sense of efficacy in classroom management in a co-teaching environment.

References

Abdal-Haqq, I. (1989). *The nature of professional development schools*. ERIC Digest 4-89. Washington, D.C.: ERIC Clearinghouse on Teacher Education. (ERIC Identifier: ED316548)

Bandura, A. (1986). *Social foundations of thought in action*. Englewood Cliffs, N.J.: Prentice-Hall.

Dettmer, P, Thurston, L.P. and Dyck, N.J. (2005). *Consultation, collaboration, and teamwork for Students with Special Needs* (5th ed.). Boston: Allyn and Bacon.

Caplan, G. (1970). *The theory and practice of mental health consultation* (2nd ed.) New York: Basic Books.

Dallmer, D., and Baker, C. (2002). Collaborative co-teaching as professional development *American Reading Forum*, 22. Retrieved September 25, 2005, from http://americanreadingforum.org/02_yearbook/html/05_Dallmer.htm

Emmer, E. T. and Hickman, J. (1991). Teacher efficacy in classroom management and discipline. *Educational and Psychological Measurement*, 51, 755-765.

Giallo, R., and Little, E. (2003). Classroom behavior problems: The relationship between preparedness, classroom experiences, and self-efficacy in graduate and student teachers. *Australian Journal of Educational Psychology*, 3(2003), 21 - 34.

Isser, R.E. (1995, Fall). Tomorrow's teachers, schools, and schools of education. *National Forum*. Retrieved September 24, 2005, from <http://www.nationalforum.com/index.htm>

Johannessen, L.R. (2004). Why do new teachers cry? *The Clearing House*. Retrieved October 30, 2005 from <http://www.highbeam.com>

Maryland State Department of Education. (February 2004). *Professional development schools: An implementation manual*. Baltimore, MD: Program Approval and Assessment Branch.

McBee, M.T. (2004, September 9). The classroom as laboratory: An exploration of teacher research. *Roeper Review*. Retrieved September 25, 2005, from <http://www.highbeam.com>

Miller, L. K. (1997). *Principles of Everyday Behavior Analysis, 3rd Ed.* Pacific Grove, CA: Brooks/Cole.

Mitchell, A. (2004). Behavior management skills as predictors of retention among South Texas special educators. *Journal of Instructional Psychology*. Retrieved January 31, 2005, from <http://www.looksmart.com/>

Mitchell, B.E. (2001, November 1). Perceptions of research from a second-year, multisite professional development schools partnership. *The Journal of Educational Research*. Retrieved September 19, 2005, from <http://www.highbeam.com>

National Association of State Boards of Education (1994). Report Review. Retrieved October 25, 2005 from http://www.nasbe.org/Educational_Issues/Reports/Sum_learn_Lifetime.pdf

National Council for Accreditation of Teacher Education (2001, Spring). *Standard III: Collaboration*. Author.

Orr, B., Thompson, C., and Thompson, D.E. (1999). Pre-service teachers perceived success of classroom management strategies. *Journal of Family and Consumer Sciences Education*, 17 (1), 38 – 49.

Pryzwansky, W.B. (1974). A reconsideration of the consultation model for delivery of school-based psychological services. *American Journal of Orthopsychiatry*, 44, 446-583.

Tharp, R. (1975). The triadic model of consultation. In C. Parker (Ed.), *Psychological consultation in the schools: Helping teachers meet special needs*. Reston, VA: The Council for Exceptional Children.

Tharp, R.G., and Wetzel, R.J. (1969). *Behavior modification in the natural environment*. New York: Academic Press.

Tschannen-Moran and Wolfolk (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.

Wolfgang, Charles. H. (1999). *Solving Discipline Problems: Methods and Models for Today's Teachers, Fourth Edition*. Boston: Allyn and Bacon.

Promoting student engagement through scholarship in a teacher preparation program

Claudia Sanchez, Ali Olson-Pacheco, Liliana Grosso, and Elizabeth Hanley¹

Abstract: A project entitled “Academic Presentations and Publications by Leaders in Education” (Project APPLE) was developed to offer pre-service teachers opportunities to grow professionally outside traditional coursework requirements. Project APPLE seeks to engage students in teacher education programs in two types of scholarly activities: professional conference presentations (Phase A) and publications (Phase B). This article describes the project’s goals and evaluates Phase A of the project. Practical suggestions are presented to develop professional development projects in teacher education programs that wish to promote early student engagement in scholarly activities.

Keywords: student engagement, scholarship, institutions of higher education.

I. Introduction.

Project APPLE (Academic Presentations and Publications by Leaders in Education) was developed to support pre-service teachers’ professional development outside traditional coursework requirements. The project’s main goal is to involve educators—who keep abreast of new developments in their field and seek to develop their skills—in profession-enhancing scholarly activities. Project APPLE has two objectives: First, to encourage student participants to identify and research their own areas of interest. Second, to engage students in two types of scholarly activities: conference presentations (Phase A) and publications (Phase B).

In pursuing these objectives, Project APPLE fosters the scholarship of teaching and learning (Isaacson, 2000) in three ways; faculty and student participants (a) expand and reflect on their knowledge and experience as pre-service or in-service teachers, (b) integrate their knowledge through collaborative work, and (c) share their knowledge publicly via presentations and publications. The project also responds to three critical issues in the field: the preparation of highly qualified teachers in grades K-12 (U.S. Department of Education, 2005b); the increasing need for educators able to integrate diversity, multiculturalism, and higher levels of technological literacy into instruction (National Board for Accreditation of Teacher Education [NCATE], 2002; Languages Other Than English [LOTE] Center for Educator Development, 2006); and the need for professional development in undergraduate programs (U.S. Department of Education, 2005a; Higher Education Act, 1989, section 201; U.S. Department of Education, 2004).

This article describes APPLE’s inception, development, and evaluation of Phase A of the project, which focused on pre-service teachers’ conference presentations on regional, state, and national levels. In addition, the plan of action for Phase B of the project and practical suggestions for those interested in developing similar professional enhancement projects in teacher education programs are presented.

¹ Texas Woman’s University, P.O. Box 425769, Denton, TX, 76204-5769, (940) 898-2009, csanchez@twu.edu, ali_olsonpacheco@hotmail.com, lgrosso@twu.edu, lizhanley@gmail.com.

II. Theoretical Framework.

Student engagement is generally considered to be among the better predictors of learning and personal development (Hu and Kuh, 2002, 2003; Kuh, Schuh, Whitt, and Associates, 1991). It is defined as the quality of effort students themselves dedicate to educationally purposeful activities such as studying, interacting with faculty members and peers on an intellectual level, and using libraries and technology for academic purposes (Astin, 1993; Chickering and Reisser, 1993; Pascarella and Terenzini, 1991). The premise behind student engagement as a predictor of learning is obvious; students learn from what they do in college. Therefore, the more students study, practice a subject, write, analyze, or problem solve, the more they will learn (Carini, Kuh, and Klein, 2006; Kuh, 2003). Empirical studies have supported this assumption, indicating a positive correlation between student engagement and critical thinking (Pike and Kuh, 2005), higher grades (Astin, 1993; Carini, Kuh, and Klein, 2006; Pike, Schroeder, and Berry, 1997), self-reported gains (Hu and Kuh, 2003), and persistence rates (Pike, Schroeder, and Berry, 1997).

Critics of the quality of undergraduate education in the U.S. encourage institutions to provide more hands-on, inquiry-related experiences to empower students as learners through their engagement in the exploration and discovery of knowledge (Association of American Colleges and Universities, 2002; National Commission on the Future of Higher Education, 2006). This recommendation is also extended to teacher preparation programs in American colleges of education. A recent study by Carini and Kuh (2003) reported results on the 2001 National Survey of Student Engagement (NSSE). Findings from their randomly-selected sample of 31,000 seniors from 317 four-year U.S. colleges and universities suggest that college students' lack of high quality experiences is more pronounced in teacher preparation programs than in other undergraduate programs. The study compared students from various university majors in terms of their level of engagement as defined by behaviors associated with high levels of learning (i.e., study time, student/faculty interaction, active and collaborative learning, enriching educational experiences). Results showed that the overall level of engagement in effective educational practices reported by future teachers was well below what may be generally desirable at the college level. When compared with their peers in other majors, future teachers also reported lower levels of student/faculty interaction, writing a larger number of short papers, taking courses that required little analytic thinking, and having fewer academic conversations with culturally diverse peers. Given this scenario, how can we promote pre-service teachers' engagement as well as active and collaborative learning in educationally enriching activities? One way to attain this goal may be through professional development opportunities that encourage students' active participation in conference presentations and the production of scholarly publications.

Aspects of the professional development of young scholars include publishing, networking in professional settings, preparing a resume, becoming a member of professional organizations, and participating in annual meetings within these organizations (Holloway, Sanchez, and Olson-Pacheco, 2004; Scott and Symens, 1997; Silvia, 2007). Some studies in the published literature describe workshops and programs dedicated to professional development for undergraduate and graduate students, especially in the areas of publishing (Figgins and Burbach, 1989; Lumsden, 1984) and membership in professional organizations (Scott and Symens, 1997). In contrast, there is a lack of studies that explore ways to promote students' participation in scholarly activities such as presentations at annual conferences of professional organizations,

where young scholars network, are able to receive valuable feedback from experienced researchers, and stay abreast of the latest developments in their field. This study describes and evaluates a project whose goal is to engage students in teacher education programs in professional conference presentations (Phase A) and publications (Phase B).

III. Project APPLE.

Project APPLE was conceived as a professional development opportunity with a focus on the integration of language diversity, multiculturalism, and higher levels of technological literacy in a teacher preparation program. The project originated at a time when the enforcement of standards for higher teacher quality and higher standards in teacher preparation were on the rise at the national level under No Child Left Behind (NCLB) and the Higher Education Act (HEA), respectively (U.S. Department of Education, 2004; 2005a). At the time, internal institutional efforts to align teacher preparation curricula with state and national standards (NCATE, 2002; LOTE Center for Educator Development, 2006) had also resurfaced at the college level.

A. Project's Inception.

The university's teacher preparation program developed Project APPLE with funds from the mentoring component of a former Title VII grant, which was designed to support pre-service teacher's professional development experiences. The funds were mainly used to purchase office supplies required for the preparation of conference presentations. Six leaders (three faculty members and three graduate students) with extensive experience in conference presentations and/or publications spearheaded the project with the purpose of disseminating information on how to present at professional conferences and publish scholarly work. Thirty-two students (18 undergraduate and 14 graduate students) participated during the project's first phase. The students were representative of the non-traditional student population on campus, whose vast majority (91%) were female and between the ages of 20 and 24. The ethnic diversity among leaders and participants included African-American, Asian, Caucasian, and Hispanic.

B. The APPLE Model.

The APPLE model consists of two main components: (a) thematic foci that respond to areas of knowledge expected from teachers, and (b) mentoring as a tool to encourage student engagement and the development of students' skills as professional presenters and writers.

Thematic foci. Themes originate from students' individual interests in the curriculum. Their chosen areas of interest include the teaching of mathematics, reading, English as a second language, native/second language acquisition and development, dual language teaching, parental involvement, grant writing for teachers, and applications of technology in instruction.

Mentoring. Faculty members participating in Project APPLE communicate high expectations for student performance, both inside and outside the classroom. The role of these faculty members is to serve as role models for scholarship and professional growth, open professional opportunities for mentees, and help protégés navigate in the culture of academe (Espinoza-Herold and Gonzalez, 2007; Gonzalez, 2003). Faculty members and students are paired based on their research interests. During Phase A of the project, mentors provide protégés

counsel in the preparation of conference proposals and conference presentations at regional, state, and national levels.

C. Project Implementation.

In order to engage students in scholarly activities, Project APPLE carries out six main activities:

Recruitment. Faculty members share the project's goals with students in their classes and promote APPLE as an optional opportunity for professional growth. Students are usually attracted to the project because of their interest in doing research, collaborating in academic projects with peers and faculty, preparing for graduate school, and enhancing their resumes through professional experiences. Other than the student's interest and commitment to the project, there are no conditions to become accepted as a participant. Teacher candidates interested in participating are then mentored outside their coursework requirements as to project specifics and completion of initial tasks. A directory is created with essential contact information of graduate and undergraduate recruits for the project, which includes students' areas of interest.

Identification of areas of interest. Faculty and participants are matched based on potential areas of interest. Faculty members help teacher candidates focus on specific topics (manageable units) for research that are related to their own personal interests and/or projects in their teacher education courses. These specific topics and projects become the focus of students' research. During the academic year, students generally devote around two hours a week to work on their projects and attend APPLE meetings.

Links between research and practice. Faculty and staff mentor teacher candidates on a one-on-one basis in using electronic databases to conduct research. Faculty members usually mentor an average of 3 students each semester with the help of staff running grant projects in the department. Teacher candidates evaluate the reviewed literature by questioning the extent to which their proposed projects are relevant to actual classroom practice and could be applied to real classroom settings. Participants are also encouraged to prepare their projects with a particular audience in mind – practitioners for grades K-12 who attend conferences in the field in search of best practices applicable to their daily work in the classroom.

Proposal writing. APPLE faculty and participants join efforts to produce high-quality conference proposals that emphasize practical applications derived from research on teaching and learning. In addition, participants look for and retrieve calls for proposals on-line, as well as write and submit proposals. Written proposals are initiated by the teacher candidates and revised by faculty as needed until the product is appropriate for submission.

Enhancement of presentation skills. The presentation delivery training available for APPLE participants includes instruction in PowerPoint formatted presentations. To ensure that teacher candidates are ready for their conference presentation, APPLE participants are screened by the project's Advisory Committee. The committee consists of three faculty members and two members of the staff who evaluate participants' presentation strengths and weaknesses using the project's evaluation materials for professional presentations. The evaluations focus on the participants' ability to introduce, present, and close their presentation, as well as engage and interact with the audience, create and use a professional PowerPoint presentation, and adhere to the time-frame.

Publication preparation. The purpose for Phase B of the project is that APPLE participants prepare publications on their areas of research based on the projects presented at professional conferences.

D. Project Evaluation.

Using qualitative methodology, an effort was made to evaluate APPLE's preliminary results by exploring students' perceptions of their engagement in the project. Semi-structured interviews (consisting of seven questions) were conducted with six program participants. The questionnaire items consisted of open-ended questions on participants' experiences, their perceptions of the project's features and effectiveness, and recommendations for improvement. The sample for the evaluation was an available pool of project participants (faculty mentors and students) working or attending classes on a full-time basis. Two faculty members and four undergraduate students supplied the data for this study. Following the transcription of the interviews, findings were analyzed using the constant comparative method (Lincoln and Guba, 1985; Merriam, 1998). Three main themes emerged from the study: students experienced increased confidence in giving professional presentations, they expressed strong interest in their continued involvement in APPLE, and also perceived benefits from the project beyond their newly acquired skills as presenters.

First theme: Students experienced increased confidence in giving professional presentations. Faculty-student mentoring and pre-conference screenings enabled students to prepare and produce professional presentations. After a meeting with her mentor, one program participant decided to present her work in her native language, Spanish. By realizing she could build on her strengths in her native language, this student became confident to perform her best in front of a group.

I prepared everything in English and they told me: "Why you don't do it in Spanish? Wow!, I said. So, of course you know it changed it all... and my concerns about my shyness were gone. I felt very confident [Speaking in Spanish] and I knew what I had to say, I knew what I had to do, I knew how to approach the people, how to get their attention, how to give my presentation a special appeal. I think I did a great job.

Nora, Undergraduate Student

[Note: All written quotes from participants have not been edited for spelling or grammar to preserve authenticity.]

Nora, a teacher candidate seeking a bilingual teacher certification became one of the strongest Spanish language presenters in the project, and her native language proficiency served as a model during presentations targeted to bilingual educators, minority language parents, and other Spanish speaking audiences.

Ana, another undergraduate student, recalled the practice and feedback her group received through the APPLE screening as beneficial to building the group's confidence and improving their presentation.

It was a lot of pressure...thinking you're going to present at a conference and don't know how to do it, having never done that before. [At the screening] you get to present in front of people you know, so you have more confidence and their feedback is honest feedback.

It helped us a lot. So, then the day we presented we were confident ourselves. People told us it was a good presentation.

Ana, Undergraduate Student

Another undergraduate student expressed how APPLE supported and encouraged her to progress from regional to national-level conferences.

There is such a great support and advice system, all of which is there for you to utilize to the full extent. As a student, it can be hard to know how to continue on and work toward more and higher goals, but they [APPLE staff and mentors] are there to help push you in the right direction to continue further... Naturally, you get a little nervous during the screenings but that is soon replaced by excitement at the conference presentations.

Nicole, Undergraduate Student

Second theme: Participants expressed strong interest in future developments of Project APPLE. When asked to offer suggestions for future project developments, the participants responded with enthusiasm regarding their interest in advanced technology training and participation in international conferences.

I think technology is a good tool to improve presentations and it's very good to know. I think I just need some more practice. I really love PowerPoint and I would like to incorporate multi-media and learn more advanced applications.

Carmen, Undergraduate Student

The more exposure we can give our students (nationally, internationally), the better. Factoring in an international component into the project would certainly be a plus. This would promote recruitment of new students... as well as give more visibility to our institution. Inter-institutional collaboration with universities abroad would be a dream come true.

APPLE Faculty Member

I think that international conferences would be really, really good. Personally, I would love to talk and work with others internationally, including going to and/or speaking at international conferences. What a better way to learn more viewpoints and perspectives, not to mention meet, learn from, and correspond with other professionals in your field.

Nicole, Undergraduate Student

Third theme: Students perceived key benefits from the program beyond their newly acquired confidence as presenters. When asked to add any information to the individual interviews conducted, participants responded with insightful comments on how the APPLE project engaged students in active and collaborative learning experiences, extended departmental visibility, and encouraged long-term professional development.

I have always been interested in higher leadership roles and activities, but have never known what direction or steps to take to do so. APPLE and its leaders break down those steps and guide you through so they don't feel intimidating and they actually feel within reach, even for an undergraduate student... I believe working with Project APPLE has helped me get to the professional level that is needed for these conventions and my hope is to go even further to the next levels. This project is about relationships, collaboration, and group work.

Nicole, Undergraduate Student

I think the program brings out the department of teacher education. This is a good project, so other people will see and be interested in it, and the university will shine because of this.

Clara, Undergraduate Student

I have become more excited about the idea of presenting at conferences and publishing. I believe this project could help other faculty members feel the same way. I love mentoring students, and helping them to start building their resumes. In doing so, I feel the urge to continue to build my own vita. I believe I have become a more productive scholar since I am involved in APPLE.

APPLE Faculty Advisor

Evaluation summary. Project APPLE was evaluated to identify participants' perceptions of their involvement in the project. The analysis of the interviews revealed three thematic findings. First, the faculty-student mentoring and pre-presentation screenings fostered students' confidence as presenters. Second, participants expressed their interest in future project developments concerning advanced technology training and conference presentations at international forums. Third, participants identified additional benefits of the project, which included expanded learning experiences through active and collaborative work, institutional visibility, and professional/scholarly motivation. The present study suggests that the preliminary outcomes of the program are directly in line with its goal. In addition, the findings from the evaluation indicate high levels of engagement by the student participants in the project. The confidence gained by students as conference presenters and their reported interest in growing with the project into its next phases (specifically, through training in advanced technology applications and international conferences) reveal participants' sense of empowerment and ownership within the project. In the following section, we outline recommended steps for teacher education programs interested in implementing a profession-enhancing project such as APPLE.

IV. Recommendations for Practice.

This step-by-step "How To" section summarizes best practices we have identified while conducting the project, and may serve as a guide for the implementation of this unique professional enhancement program in other educational settings. Figure 1 shows a summary of this section in the form of a checklist with the steps to take in order to promote pre-service teacher involvement in presentations and publications.

Figure 1. Professional Enhancement Project Checklist.

Professional Enhancement Project Checklist	
Phase A: Conference Presentations at Regional, State, and National Levels	
<i>1. Preparing for the Project</i>	
☐ A. Create/define team to conduct program	☐ D. Organize tasks in timelines with deadlines
☐ B. Develop clear goals and objectives	☐ E. Appoint Advisory Committee
☐ C. Identify available resources to run the project	☐ F. Produce recruitment plan ☐ G. Promote project

2. Training for the Conference

- | | |
|--|---|
| ☐ A. Identify students' areas of interest | ☐ F. Develop PowerPoint presentations |
| ☐ B. Produce training materials based on individual needs | ☐ G. Develop conference evaluation forms |
| ☐ C. Conduct literature reviews on specific areas of interest | ☐ H. Screen and evaluate presentations with Advisory Committee |
| ☐ D. Identify calls for proposals | ☐ I. Integrate feedback and prepare final version |
| ☐ E. Prepare conference proposals | |

3. Presenting at the Conference

- | | |
|--|--|
| ☐ A. Go through conference logistics | ☐ C. Have presentations evaluated by the audience |
| ☐ B. Organize presenters to observe/evaluate each other's performance | |

4. Debriefing after the Conference

- | | |
|---|---|
| ☐ A. Organize a meeting or a luncheon | ☐ D. Show presenters how to update their resumes |
| ☐ B. Give recognition to presenters | ☐ E. Discuss project's next steps: next conference(s), timeline(s), and expectations |
| ☐ C. Discuss presenters' experiences and evaluations | |

Phase B: Publications

5. Moving from Conference Presentations to Publications

- | | |
|---|--|
| ☐ A. Identify possible forums for publication | ☐ E. Submit manuscript according to publication guidelines |
| ☐ B. Study publication guidelines | ☐ F. Acknowledge accomplishment and record in resumes and vitas |
| ☐ C. Expand presentations into scholarly articles | ☐ G. Evaluate project and identify areas for improvement |
| ☐ D. Edit articles with participants until manuscript is ready | |

A. First Step: Preparing for the Project.

Clear goals and objectives anchored in the program, college, and university philosophies constitute the foundation of a project like APPLE. A team should be formed to run the project and monitor its development based on the stated goals and objectives. Also, members' roles should be defined and resources to run the project should be identified. Among these resources are funds for professional development activities available at the college or university level (i.e., student travel funds, and private, federal, or state program development grants). As well, the

APPLE concept can be integrated as a professional development/mentoring component into the design of many grant proposals for program development.

In preparing for the project, it is also important to establish timelines with deadlines to organize tasks. Among the key tasks to schedule are planning meetings, researching conference calls for proposals and guidelines, preparing conference proposals, and screening presentations. It is also important to appoint an Advisory Committee to govern and evaluate the project. This committee can be selected from faculty, department administrators, and student teacher representatives. Once the project's main components have been established, a plan for the recruitment of student participants should be devised. The plan should aim to recruit faculty members willing to promote the project in their classes as well as serve as mentors of teacher candidate participants.

B. Second Step: Training for the Conference.

It is vital to produce training materials based on the needs of project participants. In developing these training materials, several themes to consider may include: identifying students' areas of interest, conducting literature reviews, preparing successful presentation proposals, developing effective PowerPoint presentations, delivering effective presentations, and listing conference presentations in resumes.

In preparing for conference presentations, four main activities deserve special attention, since they demand basic technical skills as well as considerable time and effort. The first activity is the preparation of literature reviews related to the student's chosen area of research. Students often need training, which involves the identification of research themes, use of electronic databases, and use of the American Psychological Association (APA) guidelines for writing. The second key activity is the preparation of successful conference proposals, which are the result of multiple edits by participants and mentors who exchange an average of six drafts before arriving at the final product. Teacher candidate participants and mentors often use the reviewing tools in Microsoft Word to track changes and edits to proposals and use email to exchange files in an expeditious manner. The third activity that deserves special attention is the development of PowerPoint presentations. A common misconception is that all college students can effortlessly create a PowerPoint presentation. However, this is not the case. In fact, we found that 60 percent of the students in our project reported a need for basic technical guidance in using technology to produce their presentations. As a result, three staff members with strong technical skills conducted training sessions on PowerPoint for project participants. It is recommended that in cases where students need technical guidance, training be individualized when possible and designed based on a needs assessment. The fourth activity that merits special attention is the participants' presentation screenings. Approximately three weeks before the conference, mentors and student participants screen each other's work in order to assess their skills as presenters, receive feedback from an audience, and refine their conference presentations.

C. Third Step: Presenting at the Conference.

Prior to the day of the presentation, it is important to go through the presentation schedules and logistics with the teacher candidate presenters (i.e., building layouts, room assignments, and directions to the conference location). Also, the number in attendance to the presentation should be estimated so that each presenter is provided with a packet of evaluation

forms to hand out. Project participants often offer each other support by observing/evaluating each other's presentations. At the end of the presentations, presenters collect the evaluation forms from the attendees and deliver them to an Advisory Council member at the conference.

D. Fourth Step: Debriefing after the Conference.

A follow-up meeting with project participants, Advisory Council members, department administrators and possible student recruits is scheduled after the conference to present a short project review and give recognition to the students' achievements. This meeting may be planned as a luncheon and serves as an opportunity for presenters to debrief as well as share experiences and the results from presentation evaluations.

E. Fifth Step: Linking Conference Presentation's Outcomes to Next Steps.

The debriefing luncheon is a good opportunity to discuss the project's next steps. However, if an on-line interface such as Blackboard or WebCT is available, it may also serve as a viable on-going means of communication among project participants. In discussing the project's next steps, it may be helpful to address application and proposal guidelines for future conferences as participants move up to the next level of presenting (i.e. from regional to state, from state to national).

V. Conclusion.

Institutions seeking to improve the quality of undergraduate education should consider ways to encourage higher levels of student engagement (Hu and Kuh, 2003; Hu, Kuh, and Gayles, 2007) since it is, perhaps, the best predictor of learning and personal development for college students. As the literature indicates, students in teacher preparation programs who have considerable firsthand experience in educationally meaningful pursuits tend to be more effective teachers after graduation because they learn more during college (Carini and Kuh, 2003). Thus, professional enhancement programs that encourage pre-service teachers' participation in scholarly activities may be a way to promote higher student engagement and support future teachers' effectiveness.

This article described APPLE's inception, development, and evaluation of the preliminary outcomes from its first phase. Three main findings related to student engagement surfaced from the project's evaluation: First, student participants reported increased confidence as presenters, especially thanks to faculty's mentorship and pre-presentation screenings. The mentor/protégé interaction played a significant role in the achievement of this outcome. This finding comports with literature indicating that undergraduate participation in research through collaboration with faculty members is deemed highly desirable and effective in promoting student engagement (Association of American Colleges and Universities, 2002; Hu and Kuh, 2007).

The second finding from the evaluation suggests another effect of student engagement: student participants expressed their interest in expanding the project internationally and training in more advanced applications of technology. This finding suggests participants' interest in scholarly activities abroad. It also confirms the idea that the use of information technology for educational purposes is linked to how today's college students engage in active and collaborative

learning (Nelson-Laird and Kuh, 2005). Thus, it can be expected that technology will continue to play a significant role in the retention of current project participants and the recruitment of new students.

The third finding from the project evaluation outlined three main benefits beyond students' newly acquired skills as presenters: expanded learning experiences through active collaboration, institutional visibility, as well as students' and faculty's professional motivation to further their scholarly work. Given these additional benefits, project participants may be likely to remain engaged in the scholarship of teaching and learning, therefore, developing dispositions for continuous learning and personal development (Carini, Kuh, and Klein, 2006; Shulman, 2002).

The results of this study are based on a small available sample where students outnumbered faculty mentors in the interviews. Future studies evaluating professional development programs such as APPLE should further explore effects on student engagement by incorporating objective productivity indicators (i.e., participants' number of conference presentations and publications per year, as well as results from evaluations on presenters' performance at conferences). In addition, further studies should explore mentors' perceptions of their roles in this kind of professional development programs, along with the extent to which the project impacts mentors' scholarly productivity.

References

- Association of American Colleges and Universities. (2002). *Greater expectations: A new vision for learning as a nation goes to college*. Washington, DC: Association of American Colleges and Universities.
- Astin, A. W. (1993). *What matters in college: Four critical years revisited*. Jossey-Bass: San Francisco.
- Carini, R. M., and Kuh, G. D. (2003, January). Tomorrow's teachers: Do they engage in the 'right things' during college? *Phi Delta Kappa*, 391-398.
- Carini, R. M., Kuh, G. D. and Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, 47(1), 1-32.
- Chickering, A. W. and Reisser, L. (1993). *Education and identity*. Jossey Bass: San Francisco
- Espinoza-Herold, M. and Gonzalez, V. (2007). The voices of senior scholars on mentoring graduate students and junior scholars. *Hispanic Journal of Behavioral Sciences*, 29, 313-333.
- Figgins, M. A. and Burbach, H. J. (1989). Preparing our students to publish: Lessons learned. *Innovative Higher Education*, 14, 15-23.
- Gonzalez, V. (2003). Daring to be: Identity, healing, and mentoring of minority scholars. *NABE Journal of Research and Practice*, 1, 149-170.

Holloway, C., Sanchez, C., and Olson-Pacheco, A. (2004). Minority faculty in the new millennium: Five assets that increase marketability. *NABE News*, 28 (2), 11-13.

Hu, S. and Kuh, G. D. (2002). Being (dis)engaged in educationally purposeful activities: The influences of student and institutional characteristics. *Research in Higher Education*, 43(5), 555-575.

Hu S. and Kuh, G. D. (2003). Maximizing what students get out of college: Testing a learning productivity model. *Journal of College Student Development*, 44(2), 185-203.

Hu, S., Kuh, G. D. and Gayles, J. G. (2007). Engaging undergraduate students in research activities: Are research universities doing a better job? *Innovative Higher Education*, 32, 167-177.

Isaacson, R. (2000). Why JoSoTL and why now? *Journal of Scholarship of Teaching and Learning (JoSoTL)*, 1(1), 1-7.

Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change*, 35(2), 24-32.

Kuh, G. D., Schuh, J. H., Whitt, E. J. and Associates (1991). *Involving colleges: Successful approaches to fostering student learning and development outside the classroom*. Jossey Bass: San Francisco.

Lincoln, Y. S., and Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

LOTE Center for Educator Development (2006). *LOTE Teacher Competencies for Professional Development*. Austin, TX: State Board for Educator Certification.

Lumsden, D. B. (1984). Preparing future educational leaders to write for scholarly publications. *Innovative Higher Education*, 9, 36-41.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.

National Board for Accreditation of Teacher Education (NCATE). (2002). *Professional Standards for the Accreditation of Schools, Colleges, and Departments of Education*. Washington, DC: NCATE.

National Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: U.S. Department of Education.

Nelson-Laird, T. F. and Kuh, G. D. (2005). Student experiences with information technology and their relationship to other aspects of student engagement. *Research in Higher Education*, 46(2), 211-233.

Pascarella, E. T. and Terenzini, P. T. (1991). *How collage affects students*. Jossey Bass: San Francisco.

Pike, G. R. and Kuh, G. D. (2005). A typology of student engagement for American colleges and universities. *Research in Higher Education*, 46(2), 185-209.

Pike, G. R., Schroeder, C. C., and Berry, T. R. (1997). Enhancing the educational impact of residence halls: The relationship between residential learning communities and first-year college experiences and persistence. *Journal of College Student Development*, 38, 609-621.

Scott, A. D. and Symens, A. M. (1997). Promoting graduate students' membership in professional organizations. *Teaching Sociology*, 25, 176-182.

Silvia, P. J. (2007). *How to write a lot. A practical guide to productive academic writing*. Washington, DC: American Psychological Association.

U.S. Department of Education (2004). *Partnerships for Reform: Changing Teacher Preparation through the Title II HEA Partnership Program, Interim Report*. Washington, D.C: U.S. Dept. of Education. Retrieved on September 12, 2007 from www.ed.gov/rschstat/eval/teaching/title2hea/changing-teacher-prep.pdf

U. S. Department of Education (2005a). *A Highly Qualified Teacher in Every Classroom. The Secretary's Fourth Annual Report on Teacher Quality*. Washington, D.C: U.S. Dept. of Education. Retrieved on September 11, 2007 from www.ed.gov/about/reports/annual/teachprep/2005Title2-Report.pdf

U. S. Department of Education (2005b). *No Child Left Behind Act. Annual Report to Congress*. Washington, D.C: U.S. Dept. of Education. Retrieved on September 11, 2007 from www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/27/fa/c8.pdf

Discrepancy dinosaurs and the evolution of Specific Learning Disability assessment

**Whitney Moores-Abdool, Caridad H. Unzueta, Dolores Vazquez Donet,
and Eduard Bijlsma¹**

Abstract: Identifying Specific Learning Disability (SLD) has been an arduous task. Until IDEA 2004, diagnosing SLD was limited to IQ discrepancy models lacking in both empirical evidence and contributing to minority over-representation. This paper examines the history of SLD assessment, the phenomena of minority over-representation, and the implementation of Response to Intervention, (RtI) and other alternate assessment for SLD.

Key Words: Specific Learning Disability Assessment, Response-to-Intervention, Culturally and Linguistically Diverse

I. Introduction.

Identifying students with Specific Learning Disability (SLD) has been a difficult and arduous process for many Local Education Agencies (LEAs). Diagnosing SLD was limited to discrepancy models, due in large part, to the federal definition of learning disability which states that “a severe discrepancy between achievement and intellectual ability,” must be present for student identification (Re-Authorization, 2004). Lack of variation in SLD assessment is a key factor in the over-representation of culturally and linguistically diverse (CLD) students in special education (SPED) programs (Harry and Klingner, 2006).

CLD student over-representation and the lack of empirical support for test-based discrepancy models in identification of students as Learning Disabled (Ysseldyke, 2005) revealed that Intelligence Quotient (IQ) combined with school failure was an inadequate measure for identification of SLD students (Fletcher, Coulter, Reschly, and Vaughn, 2004). Addressing this concern, the *Re-Authorization of IDEA* (2004), widened the assessment process to include scientific, research-based intervention in the SLD evaluation/assessment procedure. A historical overview of CLD over-representation in SPED programs, evaluation/assessment procedures and alternate SLD assessments will be discussed in this manuscript.

II. Exigency of Special Education.

Brown v. Board of Education (1954) laid the foundation for the sixties’ civil and disabilities rights movements, bringing to light social injustices and prejudices suffered by people with disabilities. Special educator training programs in the 1960s were scarce and there were relatively no specialists available to address infant and pre-school aged children with special needs (Kirk, Gallagher, Anastasiow, and Coleman, 2006). Congress created the Bureau for the Education of the Handicapped in 1967 and passed the Early Childhood Special Education

¹ Florida International University, University Park Campus, College of Education, 11200 SW 8th Street Miami, Florida 33199

Assistance Act in 1968 (Kirk et al.) to address this issue. However, many students with special needs continued to be neglected in schools. As a result, in 1975 the Education for All Handicapped Children Act created federal funding for states to offer free and appropriate public education (FAPE) to all students, including those with special needs. This act was later re-named and amended to become the *Individuals with Disabilities Education Act* (2004). Countless students with special needs have received the adaptations and accommodation they require to be academically successful as a result of this legislation.

III. SPED Referral Process.

Classification as a student with SLD requires an assessment that may be initiated by a parent, teacher or other school professional who works directly with the student (Kirk et al., 2006). Typically consisting of a referring teacher, an administrator, a counselor, a parent, and a general education teacher, the multidisciplinary team (MDT) collaborates to recommend intervention strategies for the child (Kirk et al.). This process attempts to ensure that the student's teacher implements research-based instructional strategies, collects student performance data, and documents results to curtail unnecessary SPED referrals (Kirk et al.). Samples of the student's writing, homework, in-class assignments, and teacher-made tests/quizzes are presented for discussion in the MDT meetings (Waterman, 1994). Although information of this nature tends to be subjective (Barth, 2004), the MDT analyzes and discusses it to determine whether or not a formal SPED referral is necessary (Kirk, et al.). Many LEAs have established a pre-referral process, minimizing the subjectivity and differentiating between low achievers and students with SLD (Overton, 2003). The goal is to have the teacher attempt different teaching strategies to address multiple learning styles before referring a child for testing (Overton, 2003).

IV. Issues of Over-representation of Minorities in SPED.

Over-representation of CLD low academic achieving students is a causal factor of disproportional numbers in special education (Daunic, Correa, and Reyes-Blanes, 2004). In 1992, Black students accounted for 16% of the total U.S. population, but represented 32% of students in programs for mild mental retardation (Burnette, 1998). During the 1997-1998 academic year, more than 2.9 million, or slightly less than five percent of school aged children, received special education services throughout the United States. The Council for Exceptional Children (CEC, 2002), states that Black students are approximately three times as likely as White students to be labeled Mentally Retarded and twice as likely to be labeled Emotionally Disturbed or SLD.

The National Academy of Sciences reports that there are disproportionate numbers of CLD students in poverty. Schools with high concentrations of low-income CLD students tend to have fewer well-trained teachers and resources (Donovan and Cross, 2002). Inadequate quality teacher resources and curriculum hinders equal educational opportunities (Henderson, Kohler, Ritter, Simmons, Skiba, and Wu, 2003). School resource distribution systems most often favor the White students, and additional funding for tutoring CLD students may not be a priority. Teacher's lack of multicultural training, unclear referral processes (Burnette, 1998), ineffective instruction, and poor behavior management (Cartledge, 2002) influence inappropriate SPED program placement for CLD students. These factors hinder cultural mainstreaming for CLD

students, and serve as obstacles in gaining proficiency in the dominant language. Insufficient multi-cultural teacher training results in (Festus and Utley, 1997), a mismatch between home, community and school environments which then becomes a contributing factor in the continued over-representation of students with SLD in SPED programs. Often times, teachers have difficulty in discerning the difference between true SLD characteristics and behaviors that are a manifestation of diverse cultural backgrounds and morays (Duren, Green, McIntosh, Cook-Morales, and Robinson-Zanartu, 2005). This blurring of the two causes an increase in the number of students referred for testing.

V. Exceptional Student Education Assessment Procedures.

Understanding how minority over-representation in SLD occurs, requires one to analyze the process of student assessment for those suspected of having SLD. Choosing an appropriate assessment for SLD is vital to achieving accurate and valid results. SLD assessment requires one to review the Federal law procedures. The *Re-Authorization of IDEA* (2004) offers guidelines for identifying and testing students suspected of having disabilities. It states:

when determining whether a child has a specific learning disability, ... a local educational agency shall not be required to take into consideration whether a child has a severe discrepancy between achievement and intellectual ability in oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematical calculation, or mathematical reasoning ... a local educational agency may use a process that determines if the child responds to scientific, research-based intervention as a part of the evaluation procedures...(IDEA, 2004).

States and LEAs are offered latitude in assessing students who may meet the criteria of SLD by various methods.

Although the law stipulates that MDTs must review documentation from the child's teacher, parents, observations, and class assignments (IDEA, 2004); it does not offer specific measurements or guidelines for selecting additional intelligence assessments such as the Wechsler Intelligence Scale for Children (WISC) or Woodcock Johnson. The law, as it is currently written, entrusts LEAs to ultimately determine appropriate assessment and placement of children with SLD. What are states considering now for SLD assessment? Is there consensus on SLD definitions and on IQ measurements? What procedures are states following regarding student assessment? Studies in the 1990s yielded a wide array of factors needing further investigation. As far as the name and definition of SLD, a vast majority of states use and agree with the federal terminology previously mentioned (Kidder-Ashley, Deni, and Anderton, 2000; Reschly and Hosp, 2004). All agree on having a team evaluate and determine student placement, as it is part of *IDEA* 2004 (Re-authorization 2004). It is in the tools used to assess the students that the discrepancy exists. SLD classification criteria, including intellectual ability requirements and achievement areas, vary significantly between states (Reschly and Hosp). When congress reauthorized *IDEA*, schools were no longer required to take into consideration whether a child has a severe discrepancy in these areas, yet many states still maintain this practice. Discrepancy criterion is potentially harmful to students, as it delays intervention until student achievement is sufficiently low and a discrepancy is achieved (Lyon, Fletcher, Shaywitz, Shaywitz, Torgesen, and Wood, 2001).

In determining which assessment to use the LEA first considers the skill areas needed to be assessed (Waterman, 1994). LEAs then find suitable tests and select the appropriate ones (Waterman). Tests must be reliable and valid, as well as age and grade appropriate. Most importantly, assessments should be culturally and linguistically appropriate for the students being tested (Waterman). This is where the major problem exists for SLD assessment.

For over 30 years, the most common assessment given to school-aged children is the *Wechsler Intelligence Scale for Children- Revised* (WISC-R). This IQ test is composed of two separate tests, the Verbal and the Performance. These two sets reveal important information about the subtests useful in finding student deficiencies (Overton, 2003). However, recent studies indicate that these tests are not practical for establishing who has a learning disability and do not reveal who will most benefit from remediation (D'Angiulli and Siegel, 2003). Studies done on the WISC-R have determined that there is no clear evidence that relates performance on the WISC-R to SLD (D'Angiulli and Siegel). Prior studies have failed to include children without a SLD in their study (D'Angiulli and Siegel). This has caused a lack of comparison between students with a SLD and average students. When trying to formulate patterns, the means are used instead of individual scores (D'Angiulli and Siegel). This indicates that no other individual patterns of needs are developed other than group means. Since there is a range, looking only at the average of particular groups does not help determine future outcomes. Many now are skeptical of even using IQ testing for SLD classification. This is not to say that it may not be useful for diagnosing other disorders.

According to a recent study of the *Woodcock Reading Mastery Test- Revised*, the *Wide Range Achievement Test- Revised*, and others, the presence of significantly low scores on achievement tests is the best predictor for SLD placement (D'Angiulli and Siegel, 2003). The purpose of an achievement test is to measure a child's knowledge and retention of information. These norm-referenced tests are currently used to assist educators in making both placement and eligibility decisions (Overton, 2003). Unlike IQ tests, assessment tools measure what children should be learning in their grade level. Much revision and review of current policies for assessing students with a SLD is still needed.

After the assessments have been made, the LEAs study the results to determine eligibility criteria. LEAs look for a discrepancy between intellectual ability and achievement (Reschly and Hosp, 2004). However there is no agreement as to the size of the discrepancy nor on a formula to measure said discrepancy. Some states use standard deviations ranging from 10 to 25. This wide range could mean the difference in a child's education. When looking at the specific academic areas, a majority of the states only ask for a deficit in one, some not at all (Kidder-Ashley et al., 2000). This means a child with no notable deficiency may still be deemed as having SLD in some states. Sixteen states rely on their multi-disciplinary teams to determine eligibility of a child (Kidder-Ashley et al.). This means that the LEAs are the ones determining placement; hence there are intra-state inconsistencies in child placement and in assessment results used.

VI. Alternate Assessment for SLD.

Generally, alternate assessments aim at ameliorating student learning challenges through interventions and tiered approaches. Some examples of these are Response to Intervention (RtI), Dynamic Assessment, Curriculum Based Assessment (CBA), and the Learning Disability Evaluation Scale (LDES).

RtI has generated the greatest interest among educators. RtI is defined as the altering of activities or achievement as a result of intervention which utilizes before and after levels of measurement (Gresham, 2001). According to Mellard, Byrd, Johnson, Tollefson, and Boesche (2004), there are three tiers of RtI distinguished by the focus of intervention. The first tier is characterized by preventative activities directed toward the whole group, mainly emphasizing strong instructional intervention strategies (Mellard et al.). The second tier continues to address the group, but includes an increase in intensity, time, and frequency of instructional interventions (Mellard et al.). What changes in the third tier is that the intervention is no longer conceptualized as a preventative intervention for the group, but more as a prescriptive intervention that focuses on the individual student (Mellard et al.).

As with any new construct, there appear to be positive and negative aspects in the model. Positive effects of RtI point to a combining of functions for special and general education, a reduction of inappropriate special education referrals (Fletcher et al. 2004), and elimination of ineffective instructional strategies that may contribute to student learning difficulties (Mellard et al., 2005). Fletcher et al., indicate that some advantages of RtI are that it does not depend on teacher referral, which may at times be biased as shown by Donovan and Cross (2002). Additionally, it does not require a rigorous evaluation process. Another benefit of the RtI approach is that unlike the IQ discrepancy model which builds from student failure, RtI is preventative in nature (Mellard et al.). These qualities may sway some to embrace RtI, but many questions remain about its effectiveness as a viable alternate assessment for SLD.

Issues that arise with the RtI model are based on its effectiveness in identifying a wide age-range of students, and many ask whether all aspects of SLD can be addressed by this type of assessment (Kavale, Holdnack, and Mostert., 2005). Other questions include: (a) Will minority over-representation decrease; and (b) Can RtI measures be applied in a technically efficient manner (Kavale et al.)? Furthermore, there are implementation issues that may hinder the effectiveness of RtI. For instance, there is no consensus on procedural issues of RtI. Also there is no formulation for secondary student intervention, and the quantity of tiers needed for correct SLD identification have not been determined (Mellard et al., 2004).

There are those that propose that Dynamic Assessment may be the answer to the dilemma of accurate SLD identification. According to De la Cruz, Dynamic Assessment incorporates actual student learning as part of the SLD assessment process; the students' ability to learn is analyzed. Dynamic Assessment comes in many forms, but the basic procedure that is followed is: (a) test the student; (b) train the student; and (c) then re-test the student (De la Cruz, 1996). Swanson and Howard (2005) cite work done by Embretson (1987) that identified three goals of Dynamic Assessment, including increasing estimates of ability, evaluating new ideas, and enhancing true ability.

Swanson and Howard (2005) conducted a research study to determine if Dynamic Assessment could differentiate between poor readers and readers with a disability. Their study showed that poor readers improved with Dynamic Assessment. However 60% of those with a reading disability and 70% of those with a reading and math disability were unable to maintain performance when interventions were stopped and were unresponsive to dynamic testing conditions (Swanson and Howard, 2005). Non-response to this type of intervention was a clear identifier of learning disability (Fuchs and Fuchs, 1998; Swanson and Howard, 2005; Torgeson, 2000).

According to De La Cruz (1996), disadvantages of Dynamic Assessment mainly include concern with construct validity. However there is some disagreement over this. Some say that

changes in student scores due to intervention actually “increase construct validity,” (Carlson and Wiedl, 1979; Elliot and Lauchlan, 1997; Swanson, 1992; Swanson and Howard, 2005, p 18). Dynamic Assessment may assist in the identification of SLD by virtue of non-response to intervention; however it is unclear how reliable and practical it would be to adopt this method on a large scale.

The Learning Disability Index (LDI) developed by Lawson and Inglis (1984) is another alternate assessment being proposed in lieu of the IQ discrepancy model. LDI uses the Wescheler Intelligence Scale for Children III (WISC III), but focuses on specific subtests that indicate brain dysfunction (Watkins, Kush, and Schaefer, 2002). According to Watkins, Kush, and Schaefer, LDI “has been hypothesized to relate to specific neuropsychological deficits of students with learning disabilities,” (Lawson and Inglis, 1985, p. 98). Watkins et al. investigated the diagnostic utility of LDI using a sample from Arizona school districts of 2,053 students currently labeled SLD. They determined, based on their results that the LDI assessment resulted in accurate diagnostic results only 55% to 64% of the time (Watkins et al.). Additionally, results for diagnosing students in both reading disability and math disability revealed minimal accuracy (Watkins et al.). It can be concluded that the LDI assessment does not appear to be a scientifically supported identifier of SLD. After thirty years, we are still seeking the accurate assessment for the elusive diagnosis of SLD.

VII. Future of SLD Alternate Assessment and Minority Over-representation.

Advocates for children such as educators, counselors, social workers, parents, and guardians refer students for assessment in order to determine whether or not the student has SLD. Assessments utilized by school districts around the country aim to identify the problem, as well as the student’s eligibility for receiving necessary special education and related services. In essence, as stated earlier, students suspected of having SLD are tested, diagnosed, and are given interventions as needed to ensure that individuals have access to and make progress in the general education curriculum (Danielson, Doolittle, and Bradley, 2005).

It is apparent that based on research, the current SLD identification assessment practices, and other models that are being implemented or proposed to identify SLD, require further refinement and research to be true indicators of SLD. Keogh (2005) argues that “after decades of effort, we still struggle with questions and problems of classification and identification...” (Keogh, p. 101). Additionally, according to Reschly (2002), “current disability constructs have dubious reliability and validity, particularly treatment validity,” (Reschly, p. 124).

Validity issues and problems of classification, does not minimize the pressing concern about SLD minority over-representation. Since one is not able to determine the true validity of assessments for SLD identification, it is unclear whether one assessment or another would be more useful in reducing minority over-representation in the SLD category. However that does not mean educators and researchers have not been actively trying to address the issue of minority over-representation in SLD. In fact, since teachers are usually the first to refer a student for assessment, many have focused on issues that may be attributed to disproportionate numbers of minorities being identified for SLD. Ineffective teaching, lack of cultural understanding, failure to encourage student motivation, and using a deficit model when assessing students are a few reasons cited for minority over-representation (Malhome, and Brown, 2002; Meyer, and Patton, 2001). Addressing these and similar concerns, combined with continued research and efforts to

infuse authentic SLD assessments, hold the key to opening improved outcomes for students with SLD.

References

- Barth, J. M., Dunlap, S. T., Dane, H.; Lochman, J. E. and Wells, K.C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology* 42(2), 115-133.
- Burnette, J. (1998). Reducing the disproportionate representation of minority students in special education. *The ERIC Clearinghouse on Disabilities and Gifted Education: The Council for Exceptional Children*. Retrieved November 4, 2005, from ERIC/OSEP Digest #E566 database.
- Carlson, J.S. and Wiedl, K.H. (1979). Toward a differential testing approach. Testing the limits employing Raven matrices. *Intelligence*, (3), 323-344.
- Cartledge, G. (2002). *President's commission on excellence in special education assessment and identification task force: Minority over-identification and misidentification*. Retrieved November 4, 2005, from <http://www.charityadvantage.com/aaclld/cartledgepresentation.asp>
- D'Angiulli, A. and Siegel, L.S. (2003). Cognitive functioning as measured by the WISC-R: Do children with learning disabilities have distinctive patterns of performance? [Electronic version] *Journal of Learning Disabilities*, 36(1), 48-58.
- Danielson, L., Doolittle, J. and Bradley, R. (2005). Past accomplishments and future challenges. *Balt Learning Disability Quarterly*, 28(2), 137-139.
- Daunic, A.P., Correa, V.I. and Reyes-Blanes, M.E. (2004). Teacher preparation for culturally diverse classrooms: Performance-based assessment of beginning teachers [Electronic Version]. *Teacher Education and Special Education*, 27(2), 105-118.
- De la Cruz, R. (1996). Assessment-Bias issues in special education: A review of literature. EC 304 511.
- Donovan, M. and Cross, C. (2002): *Minority students in special and gifted education*. Washington, D.C.: National Academy Press.
- Green, T. D., McIntosh, A. S., Cook-Morales, V. J. and Robinson-Zanartu, C. (2005). From Old Schools to Tomorrow's Schools: Psychoeducational Assessment of African American Students. *Remedial and Special Education* 26(2), 82-92.
- Elliot, J. and Lauchlan, F. (1997). Assessing potential- the search for the philosopher's stone? *Educational and Child Psychology*, 51, 6-16.
- Festus, O. and Utley, C. A. (1997). Rethinking preservice preparation for teachers in the learning disabilities field: Workable multicultural strategies. *Learning Disabilities Research and Practice* 12(2), 100-106.

- Fletcher, J., Coulter, W., Reschley, D., and Vaughn, S. (2004). Alternative approaches to the definition and identification for learning disabilities: Some questions and answers. *Annals of Dyslexia*, 54(2), 304-331.
- Fuchs, L.S. and Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing identification of learning disabilities. *Learning Disabilities Research and Practice*, 13, 204-219.
- Gresham, F. (2001). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. Paper presented at the Learning Disabilities Summit: Washington, D.C.: ED 458 755.
- Henderson, M., Kohler, K., Ritter, S., Simmons, A. B., Skiba, R. J. and Wu, T. (2003). The context of minority disproportionality: Local perspectives on special education referral, a status report. *Indiana Educational Policy Center Indiana University*. Retrieved November 4, 2005, from <http://www.indiana.edu/~safeschl/contextofmindisp.pdf>
- Lawson, J. S. and Inglis, J. (1984). The Psychometric Assessment of Children with Learning Disabilities: An Index Derived from a Principal Components Analysis of the WISC-R.. *Journal of Learning Disabilities* 17 (9), 517-22.
- Jitendra, A. and Kameenui, E. (1993). Dynamic assessment as a compensatory assessment approach: A description and analysis. *Remedial and Special Education*, 14(6), 6-18.
- Kavale, K., Holdnack, J. and Mostert, M. (2005). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly*, 28(1), 2-16.
- Keogh, B. (2005). Revisiting classification and identification. *Learning Disability Quarterly*, 28(2), 100-102.
- Kidder-Ashley, P., Deni, J. R. and Anderton, J. B. (2000). Learning disabilities eligibility in the 1990's: An analysis of stat practices. *Education* 121(1), 65-71.
- Kirk, S. and Gallagher, J. (2006). *Educating Exceptional Children (11th ed.)*. New York: Houghton Mifflin.
- Lyon, G. R., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., et al. (2001). *Rethinking Learning Disabilities*. In C. E. Finn, Jr., A. J. Rotherham, and C. R. Hokanson, Jr. (Eds.), *Rethinking special education for a new century* (pp.259-287). Washington, DC: Thomas Fordham Foundation. Retrieved September 29, 2007, from http://www.edexcellence.net/library/special_ed/index.html
- Maholmes, V. and Brown, F. (2002). Over-representation of African American students in special education: The role of a developmental framework in shaping teacher's interpretations of African American students' behavior. *The Trotter Review*, 14(1), 45-61.

Meyer, G. and Patton, J. (2001). *On the nexus of race, disability, and overrepresentation: What do we know? Where do we go? On Point...Brief discussions of critical issues in urban education*. Newton, MA: National Institute for Urban School Improvement.(ERIC Document Reproduction Service No. ED 462 487)

Mellard, D., Byrd, S., Johnson, E., Tollefson, J. and Boesche L. (2004). Foundations and research on identifying the model responsiveness-to-intervention sites. *Learning Disability Quarterly*, 27(4), 243-256.

Overton, T. (2003) *Assessing Learners with Special Needs: An Applied Approach* (4th ed.). 9- 20. New Jersey: Pearson Education. Practical Applications of Special Education. Retrieved November 4, 2005 from http://www.ouhsd.k12.ca.us/schools/instructional_programs/sped_hand

Re-Authorization of the Individuals with Disabilities Education Act, Pub. L. No.108-446, §614, 118 Stat. 2647 (2004).

Reschly, D. (2002). Change dynamics in special education assessment: Historical and contemporary patterns. *Peabody Journal of Education*, 77(2), 117-136.

Reschly, D.J and Hosp, J.L. (2004). State SLD identification policies and practices. *Learning Disability Quarterly*, 27(4), 197-213.

Swanson, H.L. and Howard C.B. (2005). Children with reading disabilities: Does dynamic assessment help in the classification? *Learning Disability Quarterly*, 28(1), 17-34.

Swanson, H.L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, 86, 473-488.

Torgeson, J.K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice*, 15, 55-64.

Waterman, B.B. (1994). LDOnline: Assessing Children for the Presence of a Disability. *NICHEY News Digest* 4(1). Retrieved October 16, 2005, from http://www.ldonline.org/ld_indepth/assessment/assess-nichcy.html.

Ysseldyke, J. (2005). Assessment and decision making for students with learning disabilities: What if this is as good as it gets? *Learning Disability Quarterly*, 28(2), 125-128.

Instructional accommodations: Impact of conventional vs. social constructivist view of disability

Sarah M. Ginsberg and Karen Schulte¹

Abstract: This study examined the relationship between university faculty members' perspectives on students with special learning needs and their views on providing accommodations. Findings demonstrate that professors who view students with disabilities from a conventional, deficit perspective feel ill-equipped to provide necessary accommodations. Professors who hold a social constructivist view of disabilities view all learners, including those with identified needs, on a continuum and see accommodations for special learners as being an extension of their good teaching. Implications for faculty development and increased use of universal instructional design are discussed.

Key Words: accommodations, social constructivism, interactionist perspective, view of disability, college classroom, university teaching

I. Background.

We have seen a significant increase in the numbers of students with disabilities attending post-secondary education programs in the past decade. Current statistics suggest that approximately 10% of students in postsecondary education institutions in the United States have a disability (Rickman, 1995, Scott, McGuire and Foley, 2003). Unlike the K-12 programs, however, teachers in colleges and universities have little support or knowledge of how to effectively support and educate students with special needs. While there are many institutions that offer support programs to students who have identified disabilities, few of these focus on the improvement of the quality of educational experience from the faculty perspective. As one professor from a College of Business stated, "most of us around this university do not . . . even have any training in education per se. I mean you know for most of us teaching, exam writing, I mean, we've learned off-the-cuff, on our own." He goes on to explain that

I've never had training in the needs of special education students so unless the student tells me there's some kind of additional thing that I need to be doing in class or there is something that I could be doing differently, then I just basically have to assume that everything is going okay.

Limited familiarity with a wide variety of pedagogical techniques and with students whose needs are out of the mainstream of average learners is common in post-secondary institutions across the country (McGuire and Scott, 2006). There is a "persistent myth" (Roche and Marsh, 2001, p. 445) that college faculty who have completed terminal degrees are knowledgeable about teaching and are adequately prepared with pedagogical skills for teaching positions. Very few educators receive any formal preparation for their roles as teachers (Hativa, 2000; Hativa, Barak, and Simhi, 1999; Roche and Marsh, 2000). Instead most teachers report that they have learned their skills in on-the-job training and by trial and error with little support

¹ Department of Special Education, Eastern Michigan University, 128 Porter, Ypsilanti, MI 48197, sginsberg@emich.edu, kschulte@emich.edu.

or direct supervision. Faculty may have “fragmented knowledge and unfounded beliefs about what makes instruction effective” because of their limited formal training (Hativa, Barak, and Simhi., 1999, p. 3). Limited knowledge about the variety of learning needs present in their classroom may be one such area of deficit in their understanding of effective teaching. Many university instructors emphasize content over pedagogy (Shaw, Scott and McGuire, 2001).

The conventional view of students with special learning needs or disabilities as being “defective” (Ghersiequiere, Maes, Vandenberghe, 2004) is not unheard of in society in general nor amongst university faculty. This view of disabilities and special needs focuses on the disability labels assigned to students. The emphasis is on the impairment as it resides within the individual—This is perceived as the dominant characteristic and represents a deviance (Porter, 1994). The underlying assumption in this view is that the student who has a learning disability, for example, is a student who has an underlying condition that is the root cause of his success or failure (Dudley-Marling, 2004). This view of a disability sets the learner off to the side, creating a sense of otherness or being outside the spectrum of learners present in the class. It may be stigmatizing and overwhelming for the student as he feels that the disability represents his load alone to bear. Teachers holding this view may be “on the lookout for deficits in need of remediation” (Dudley-Marling, 2004) rather than seeing themselves as an educator seeking an opportunity for learning.

In contrast, the *interactionist* point of view suggests that disabilities are not “rooted in persons, but instead arising from interactions between persons and their environment” (Porter, 2004, p. 71). Also known as *social constructivism*, this perspective suggests that learning and learning problems are rooted with the context of human interactions and relationships (Dudley-Marling, 2004). A learning disability, for example, is not a reflection of the learner’s malfunction, but rather is evident when the learning activities and environment, including the teaching, do not sufficiently support the learner in such a way as to facilitate successful learning. Inherent in the *interactionist* perspective is the view of the student with special learning needs as one who may have more unique or extreme needs than the average learner, but whose needs are related to those other students have as well. In other words, they are not separate and apart from a body of learners, but perhaps some of their needs represent a farther point on the continuum of needs all learners have. In some learning scenarios, they may be as successful as or more successful than other classroom learners. They are not a deficient learner at all times, rather their abilities, performance and learning will change with alterations in tasks, environments, and teachers. The “disability” is not static and omnipresent, but dynamic and fluctuating across interactions (Porter, 1994).

Given the lack of formal education regarding college teaching and limited exposure to individuals with disabilities in combination with the rising numbers of students with disabilities of all types in higher education, we began to wonder how university faculty members were responding to students with special learning needs in their classrooms. We set out to learn more about the faculty perspective on “What do you do for students with disabilities in your classroom?”

II. Method.

A. Study Design.

This qualitative study was conducted by interviewing individual faculty members to gain their self-report regarding their approach to teaching students with disabilities in their own classrooms. The participants' perspective was solicited in an effort to understand the multiple perspectives of those situated in the context of this particular problem (Glesne, 1999). We attempted to create meaning of the similarities and differences in the individual stories (Denzin and Lincoln, 1998). The use of qualitative research methods is possibly one of the most appropriate "because it allows researchers to unravel the complex school and classroom realities" (Ghesquiere, Maes and Vandenberghe, 2004, p. 172). This study used a multiple case study approach to aid in drawing comparisons and differences between of a small group of university faculty.

B. Participants.

This study was conducted at a large, Midwest, public four year university. Faculty volunteers were solicited from all departments across campus. Faculty volunteers were purposefully chosen in an attempt to have broad disciplinary representation of a typical case sampling (Glesne, 1999). Of the respondents, 12 full-time, tenure track faculty members were interviewed. Participants' experience with university teaching ranged from untenured assistant professors with only two years of experience to full professors who had been teaching for more than 30 years. No teachers who were in their first year of teaching were included in the study. Of the 12 faculty members, 25% of the volunteers were men. The average age of participants was 49 years old, ages ranged from 32 to 65 of age. Purposively sampling (Glesne, 1999) was used to obtain representation from all of the university's colleges, including Business, Technology, Education, Health and Human Services, and Liberal Arts. Faculty disciplines included teacher education, math, psychology, accounting, art, communications, health education, occupational therapy, geology, and engineering. No faculty from the Department of Special Education was included.

C. Data Collection.

As noted above, all participants were individually interviewed. Interviews averaged approximately 60 minutes in length. Questions were designed to elicit the perceptions and attitudes of the teachers (Glesne, 1999) regarding their teaching of students with disabilities. The term "disabilities" was broadly defined as any condition that affected the students' learning or limited functional abilities (Belch, 2004-5). The semi-structured interviews began by asking the professors the open-ended question "Tell me about your experience teaching students with disabilities or special learning needs at this university." In addition to spontaneous questions that arose during the course of the interview, all faculty were all asked to explain what accommodations they made in their classrooms for learners with special needs.

D. Data Analysis.

We audio taped each interview with the participant's permission. The audio recordings were transcribed. Creswell's (1994; 1998) guidelines for the analytic processing of data using reduction and interpretation were followed. We used the "progressive process" of sorting, defining, and relating the data with codes to interpret into relevant units to make better sense of their meanings (Glesne, 1999). In the initial coding stages, all data from faculty interviews were reviewed individually. All data was sorted sentence by sentence into categories related to content or message associated with immediacy behaviors reported and observed. These categories were then compared and consolidated into broader categories. Each chunk of data was then compared with these broader categories and compared with each other. As we analyzed the codes for patterns and themes, we "linked [them] together" to begin forming theoretical models (Denzin and Lincoln, 2003, p. 279). In addition, data analysis and coding was reviewed by study collaborators and co-authors for the purpose of validating codes and conclusions (Creswell, 2003; Lincoln and Guba, 1985).

III. Results.

Participants' responses to the question "Tell me about your experience teaching students with disabilities or special learning needs at this university" demonstrated that faculty tended to hold either the deficit, conventional view of students with disabilities or the social constructivist perspective, though none identified their views in these terms. Faculty whose responses to students with special learning needs indicated that they viewed these students as separate and unrelated to the range of learners in their classrooms were identified as having the conventional view of disabilities. Comments from these faculty suggested that the student with disabilities posed significant challenges to them as teachers, often because they did not have enough knowledge of the specific disability to educate adequately the student. This group also described the special needs learner as being the person primarily, if not solely, responsible for identifying how they should be effectively educated. Faculty holding the conventional view also noted that students with physical and other obvious disabilities, such as being in a wheelchair, were easier to educate because they understood those disabilities better.

The volunteers whose perspectives were consistent with the conventional view of disabilities advocated accommodations only for those students with appropriate documentation, and the accommodations mentioned focused on assessment. Providing "extra time to take the exam" was often the first accommodation reported by this group of faculty and was occasionally reported as the sole commonly offered accommodation. It was reported that it is "not faculty responsibility" to offer accommodations and that they prefer that the learner comes to the class "when they already have adaptations." The accommodations focused on specific activities within the context of the course, such as note taking, test taking, or periods for work completion. Additionally, all of the faculty who held the conventional view of disability expressed concerns that if they offered students with disabilities too many accommodations, it would not be "fair" to the other students in the class. Most of these faculty members indicated that they would not provide any accommodations to students who failed to produce the appropriate university paperwork documenting their disability. This belief was not explicitly connected to a strictly enforced university policy and was, in fact, quite dissimilar to other faculty interviewed at the same university.

In contrast, some faculty described their view of students with special learning needs in a manner consistent with the interactionist view. These faculty see special needs students as students on another point of the same continuum as all other students. Rather than describing their understanding of the specific disability causing learning difficulties, they focused on their interactions with the students. Many participants in this group described sitting down and talking with the students about what they need in order to be successful learners, independent of disability label. In fact, few of these teachers required students to present formal documentation. Rather, they viewed their role in teaching to be one in which they were able to facilitate learning for all students, no matter that their particular needs were. This perspective was best described by one participant when she said that what she does “isn’t an accommodation, it’s just good teaching.” Samples of faculty verbatim quotes are included in Table 1 to illustrate the two perspectives.

Table 1. Faculty conventional vs. social constructivist view of disabilities.

Conventional View of Students with Disabilities	Interactionist/Social Constructivist View of Students with Disabilities
• “I don’t have a tool kit of adaptations.” • “I don’t have that knowledge of how to teach kids with special needs.” • “How do you adapt your teaching? I haven’t because I don’t know what it should be.” • “I’m treating the students [with disabilities] differently.” • “I’ve never had training in the needs of special education students, so unless the student tells me there’s some kind of additional thing that I need to be doing in class or there is something that I could be doing differently, then I just basically have to assume that everything is going o.k.”	• “I’ll do whatever the person needs.” • I do for them “what I do for all students.” • I talk “about the wonder of our differences” in class. • “There aren’t any magic wands.” • “It is the students who have success . . . not me.” • What matters to me is their mastery of the material, not how long it takes them to communicate that.” • “The difference between a problem and a crisis is about 24 hours of neglect. We just need to take care of things.” • “This isn’t an accommodation, it’s just good teaching.”

When asked how they accommodate for learners with special needs, the focus of faculty who held the social constructivist view was on the interactions between themselves and the learners in their classrooms. Accommodations were developed following interactions with the students. The power of the teacher-student interaction was frequently mentioned. They noted that it was important to them to create a “safe environment” for learners in their classes, including those with special learning needs. In contrast to the faculty who had a more conventional view, the teachers who had a social constructivist view of disabilities seldom raised concerns regarding accommodations being “fair” to other students. Little mention was made regarding accommodations made for specific disabilities or some disabilities being easier to work with than others. These faculty members did not require documentation of student disability prior to

providing accommodation. They espoused a view of teaching that invited the students to interact with them so that together they could be sure that the students' needs were met. Table 2 summarizes common accommodations and responses to the question of what faculty do for students with special learning needs.

Table 2. Faculty conventional vs. social constructivist use of accommodations.

Conventional View of Students with Disabilities	Interactionist/Social Constructivist View of Students with Disabilities
• "Extra time to take the exam" • Lecture notes in advance • Extra time for work • "Supplemental Instruction" provided by learning center G.A. • "It is best when they already have adaptations." • It is "not faculty responsibility."	• Paraphrase readings to teacher to check understanding • Break content down into "smaller pieces" • "Cafeteria exams" giving students choices of variety of question styles • Give exams "orally" • Peer support facilitated by instructor • Create a "safe environment" • "I would be happy to meet with you as much as necessary." • "Come and see me . . . so we can work things out."

IV. Discussion.

Ghesquiere, Maes and Vandenberghe noted that viewing students "with special education needs to be a special, 'defective' group hinder[s] the development of a truly inclusive vision and practice" (p. 182). The data from faculty interviews suggest that indeed the faculty whose beliefs in the social constructivist view of students with disabilities espoused very inclusive ideas about how to educate the entire class, including those with special needs. In contrast, faculty who held a conventional view of students with disabilities viewed the educational practices for these students to be very separate from the whole of the class.

The qualitative nature of this study means that we cannot generalize the results to all groups of faculty in all situations. Rather it gives us insights into the perspective of faculty from one university who were chosen based on representative sampling for the wide-ranging disciplinary background, age, and years in higher education. These results give us a new framework that may be transferable to broad faculty development issues in efforts to improve the quality and accessibility of higher education to all students, particularly those with special learning needs. As this study was fairly preliminary in nature, it did not seek to gain first hand observations of the professors teaching, nor were students in their classrooms interviewed regarding their perceptions of the teachers' views of students with disabilities. Inclusion from these sources would be a next logical step for the progression of this line of research. It is a limiting factor for this work. Additional work might take into account greater representation of diverse faculty, including, but not limited to, different types of post-secondary institutions and cultural diversity. Nonetheless, the results reported here do allow us to reflect on the value of

understanding teachers' perspectives on the students in their classroom and the impact on how they facilitate their learning.

For students with teachers who hold the conventional view of disabilities, the sense of being divided from the main learning community has the potential to not only put undue burden on the learner for their role in their education, but to create a sense of alienation in the learning environment. Students who do not feel that they belong are likely to have decreased retention in post-secondary education programs (Belch, 2004-5). A common and successful approach to remediating this problem in college classrooms is the use of universal instructional design (Belch, 2004-5; Burghstahler, 2001; Mino, 2004; Ouellett, 2004; Scott, McGuire and Foley, 2003). In universal instructional design, education is designed to be accessible to all learners, including those with and without disabilities. It is accepted among those following universal instructional design principles that the educator recognizes that the differences between learners is a continuous range and that no one group represents a discreet, disconnected set of students within the classroom (Belch, 2004-5; Scott, McGuire and Foley, 2003). Flexibility is a key component that allows teachers to adjust and adapt to the varied learning styles and needs in their classroom.

In recent years, there has been an increasing call for expanding the awareness of faculty in higher education concerning issues of diversity in our classrooms as well as implementing universal instructional design (Belch, 2004-5; Burghstahler, 2001; Mino, 2004; Ouellett, 2004; Rickman, 1995; Scott, McGuire and Foley, 2003). This inclusive approach to education has been mandated in public K-12 education since Public Law 94-142 Education for All Handicapped Children Act was passed by Congress in 1975 (Smith, 2004). This legislation has been reauthorized numerous times, most recently in 2004, and is now known as the Individuals with Disabilities Act (IDEA). While access to higher education has been a more recent mandate, legislation such as the Americans with Disabilities Act (ADA) of 1990 and the Section 504 of the Rehabilitation Act of 1973 requires education to be nondiscriminatory and equally accessible to all students. (Belch, 2004-5; Rickman, 1995; Scott, McGuire and Foley, 2003).

The use of universal instructional design is one way to ensure the adherence of colleges and universities to the above laws and is beneficial to all students, not just those identified with special learning needs. However, if we expect professors to be able to provide an accessible education for all students, faculty development centers will need to begin not only with increasing faculty awareness of disabilities, but by teaching the underlying premises central to social constructivism before moving forward to address universal instructional design. Instruction in pedagogical accommodations, while a necessary component of change is not sufficient. This must be accompanied by experiences that will change the perspectives teachers hold related to disabilities. Without addressing the underlying view of students with disabilities, they will continue to be seen by faculty as a distinct and separate group of learners within the classroom.

References

- Belch, H. A. (2004-5). Retention and students with disabilities. *Journal of College Student Retention*, 6(1), 3-22.
- Burghstahler, S. (2001). Universal design of instruction. DO-IT. Seattle, WA; University of Washington. (ERIC Document Reproduction Service No. EC309179).

Denzin, N. K. and Lincoln, Y. S. (2003). *Collecting and interpreting qualitative materials*. Thousand Oaks, CA: Sage.

Denzin, N. K. and Lincoln, Y. S. (1998). *Strategies of qualitative inquiry*. Thousand Oaks, CA: Sage.

Dudley-Marling, C. (2004). The social construction of learning disabilities. *Journal of Learning Disabilities*, 37(6), 482-489.

Ghesquiere, P., Maes, B., and Vandenberghe, R. (2004). The usefulness of qualitative case studies in research on special needs education. *International Journal of Disability, Development and Education*, 51(2), 171-184.

Glesne, C. (1999). *Becoming Qualitative Researchers: An introduction* (2nd ed.). New York: Longman.

McGuire, J. M. and Scott, S. S. (2006). An approach for inclusive college teaching: Universal design for instruction. *Learning Disabilities*, 14, 1, 21-31.

Mino, J. J. (2004). Planning for inclusion: Using universal instructional design to create a learner-centered community college classroom. *Equity and Excellence*, 37, 154-160.

Ouellett, M. L. (2004). Faculty development and universal instructional design. *Equity and Excellence*, 37, 135-144.

Porter, J. (1994). Disability in higher education: From person-based to interaction-based. *Journal on Excellence in College Teaching*, 5(1), 69-75.

Rickman, J.C. (1995). The microculture of disability. In J.Q. Adams and J.R. Welsch (Eds.), *Multicultural education: Vol. 4. Strategies for implementation in colleges and universities* (pp.177-186). Macomb, IL: Western Illinois University.

Scott, S. S., McGuire, J. M., and Foley, T. E. (2003). Universal design for instruction: A framework for anticipating and responding to disability and other diverse learning needs in the college classroom. *Equity and Excellence in Education*, 36, 1, 40-49.

Shaw, S. F., Scott, S. S., and McGuire, J. M. (2001). Teaching college students with learning disabilities. ERIC Digest. Arlington, VA: ERIC Clearinghouse on Disabilities and Gifted Education. (ERIC Document Reproduction Service No. ED459548).

Smith, D. D. (2004). Introduction to special education (6th Ed.). Boston, MA: Pearson.

Why I love/hate Wikipedia: Reflections upon (not quite) subjugated knowledges

John Aycock and Alan Aycock¹

Abstract: Wikipedia is a well-known online encyclopedia, whose content is contributed and edited by volunteers. Its use by students for their research is, to be polite, controversial. Is Wikipedia really evil, or is it a teaching opportunity in disguise, a representation of some deeper cultural change? We present first-hand accounts from two different disciplines, computer science and anthropology, to illustrate how experiences with Wikipedia may be crossdisciplinary. We use these to reflect upon the nature of Wikipedia and its role in teaching.

Keywords: Computer Science, Anthropology, Social Sciences, teaching, learning, scholarship, World Wide Web, Wikipedia

Wikipedia exists. That is not in dispute; the question is what role (if any) Wikipedia should play in the classroom, in assignments, and in research. We think that Wikipedia itself is not really at issue, but it provides a convenient lightning rod for a debate that, deep down, is really about a restrictive view of research sources and underlying cultural values, and their eventual emergence in the classroom.

We present our argument in a nontraditional manner. We begin with two first person narratives, experiences the authors have had with Wikipedia in their courses, one in computer science and one in anthropology. Those narratives are used in the discussion that follows to draw out our argument, about why we love/hate Wikipedia and what can be done about it.

I. John's Story.

'Please permit me to introduce my humble self to you. I am Mr. James Williams the Director of International Relation with Access Bank of Nigeria Limited. I'm 47yrs old and I got your email address on the Chambers of Commerce address list and my confidence reposed on you to contact you. I hope you read this message Carefully and reply me immediately, although we have not met before but I suggest that this transaction will bring us together.'

This excerpt is all too familiar to anyone who has ever checked their email. These emails are currently known as Nigerian scams, or 419 scams, and they are a type of what is generally called advance fee fraud. I wanted to find out when these scams originated, and Wikipedia sent me on a search that has lasted, off and on, over two years.

Since 2005, I have taught a 4th-year computer science course on spam and spyware (J. Aycock, 2006). This was, and to the best of my knowledge still is, the only course of its kind in the world. One of the topics I cover is fraud that's perpetrated via spam, like the aforementioned

¹ Department of Computer Science, University of Calgary, 2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4, aycock@ucalgary.ca; Department of Anthropology, University of Wisconsin at Milwaukee, P.O. Box 413, Milwaukee, WI 53201, aycock@uwm.edu.

advance fee fraud. When preparing my first lecture on the topic, I thought it would be useful to put this fraud into its historical context. When was the first advance fee fraud, anyway?

I started with Wikipedia. I view Wikipedia in the same way that I view Google, as a labor-saving device subject to noise and misinformation. However, Wikipedia articles sometimes cite reasonable, academically sound sources of information that can provide good starting points, and that's what I was after.

The Wikipedia entry on advance fee fraud (Wikipedia, a) observes that the Nigerian scam is related to an older variant, called the "Spanish Prisoner" scam. So far, so good. The Spanish Prisoner Wikipedia entry (Wikipedia, d), in turn, begins by dating the scheme back to the year 1588. And a citation beside it! A lead on a primary source, I thought, as I eagerly clicked on the link to find... some guy's web page, with no mention of where the date came from. Some Guy was very helpful when I emailed him, however, and said he got the date from a long-lost newspaper article. A dead end.

Much pawing through web search results and library databases later, I realized that there were many references to that date, but no one seemed to know where it came from. Yet it appeared, without citation, in a wide variety of places: a law journal article (Morton, 1996); the *New Yorker* (Zuckoff, 2006); the UK Metropolitan Police website (Metropolitan Police Service, n.d.); an academic journal article (Glickman, 2005). Finally I stumbled across a well-buried bibliography that mentioned a book by Jay Robert Nash, *Hustlers and Con Men* (Nash, 1976). Page 26 has the magic number, 1588, but there are no excerpts from this ancient letter, no mention of which source it came from. Trying likely-looking references in the book led nowhere. Another dead end.

The same well-buried bibliography also pointed to an article about the Spanish Prisoner dating back to 1909; this article even contains an excerpt of the letter from 1908 (Author Unknown, 1909). A little more poking around through the 1900s turned up several more examples, one from 1940 (U.S. Department of State, 1940) and one from 1909 (U.S. Department of State, 1909), both warnings issued by the U.S. State Department. The latter is interesting because it notes that the scams have been going on 'For nearly twenty years.'

My next attempt was made on the premise that legal records go back hundreds of years, and if the scam did go back to the 16th century, then surely someone had been caught and tried. Unfortunately, I couldn't find any such records, but it did cause me to stumble across some criminology books from 1899 (Griffiths, 1899; PowerBerrey, 1899). Both make it sound as if receiving the scam letters is a regular event, and one contains an 1898 scam letter, noting the scam 'has been practiced for half a century or more' (PowerBerrey, 1899, p. 170). Even allowing for some exaggeration by the author, it seems fairly certain that these scams existed in the mid-to late 1800s.

But as early as 1588? I'm skeptical. Despite spending far too much time trying to trace the letters further, I've found no evidence before the 1800s. The post in England was not open for the public to use before 1635 (Beale, 1998), making it difficult for scam letters to have been sent in 1588, and Elizabethan fraud seemed to have been conducted mostly face-to-face (see, for example, Aydelotte, 1913).

Thanks to one Wikipedia entry, I was sent on a wild goose chase while preparing my lecture for a computer science class. And after all this, 1588 still persists in the Wikipedia universe – the digging I've done constitutes original research, which is prohibited content according to Wikipedia policy (Wikipedia, c).

II. Alan's Story.

During the spring semester of 2006, I prepped and taught a freshman seminar course in cultural Anthropology called "Buy Me! Ads and Shopping in American Culture." Freshman Seminars are especially designed to help traditional first-year students with their transition to university work. The idea is that if the course is kept small enough and interesting enough, first-year students are more likely to stay enrolled during their first semester, which typically has the highest attrition rate. I had only taught this course at the graduate level previously, but I believed that it would be very attractive to students, since most of them are avid shoppers and consumers of popular culture.

As an advertisement for the course, I was asked to send the first-year student advisors a sprightly two- or three- paragraph blurb to entice students to take the course. At the time I wrote this, I had not completely finished writing my syllabus. However, I knew that there were certain topics I wanted to cover, and particular theoretical models that I intended to use during the course. In this subfield of Anthropology, the theoretical models are often named after the cultural situation that inspired them: for instance, there is a model of "McDonaldization" (Ritzer, 2004) and another of "Disneyfication" (Bryman, 2004). I also wanted to spend some time in the course talking about Walmart as a cultural phenomenon, so pressed for time, I spontaneously made up a theoretical model that I termed "Walmarting" to add to the blurb. I emailed this little advertising blurb to the first-year advisors, and promptly forgot about it.

A few months later, about halfway through the spring semester, we were ready to start on the portion of the course that included an exploration of Walmart. I had made it my practice to use Web sites as often as possible to add resources to each course topic, so I Googled "Walmart," then as an afterthought, "Walmarting." To my surprise, "Walmarting" turned up a Wikipedia entry, so I went there to see what was available online.

The Wikipedia entry (Wikipedia, e) was very short, not really suggesting a developed theoretical perspective. There was a brief statement of some criticisms of Walmart, but little else. However, my attention was drawn to the links provided at the end of the Wikipedia article. One of the links was to "a course syllabus that involves Walmarting" – I followed this link, and what did I find? Of course, the link was to my own advertising blurb for the course I was teaching! Talk about circularity, I felt as if I were in an Escher drawing!

I decided that this had to become a "teaching moment," so I told my students about the Wikipedia link at our next class meeting. But I also wanted to use this unexpected Web manifestation as an opportunity for active learning. So I asked my students to help develop the model of "Walmarting" by writing a short assignment in which they identified two or three key features of Walmart that they considered central to the culture of Walmart. The students took to this assignment quite happily. Not only had they been reading about Walmart online, they had also seen a video critical of Walmart. In addition, many of them were themselves dedicated Walmart shoppers, so they had a fund of their own cultural experience to draw upon.

After the class was over, I went back to my office and compiled the students' responses, then wrote half a dozen paragraphs – making certain to observe the Wikipedia standard of NPOV (neutral point of view (Wikipedia, b)) – that constituted a preliminary model of "Walmarting" that would be very similar in scope to others in this field. I also compiled an up to date bibliography, including a surprising number of books and videos on Walmart: it's really been assimilated into American popular culture in the past few years! The Wikipedia editor I contacted about the Walmarting article was extremely helpful with technical matters, and within a couple of days the article was complete.

I revisited my class to show them the article and invite comments. Interestingly, although they agreed that it was pretty cool to be part of a wiki collective, an “exit assignment” (A. Aycock, 2006) at the end of the class suggested that they weren’t nearly as amazed as I was by what had just happened. I had thought that this was an extraordinary demonstration of how globalization affects us in ways we can never predict or even control. As well, it seemed to me to illustrate nicely the effect of time-space compression that makes our pace of life so frenetic (Harvey, 1990). At first I was a little disappointed at the students’ reaction (or lack of reaction), but then as I thought about it I realized that for them, it had always been this way – in the same sense that for them, John Lennon had *always* been dead. It was I who was the cultural outsider here, for them globalization and the postmodern experience of “placelessness” were merely the way it’s supposed to be.

III. Discussion.

Taken together, we have two distinct experiences with Wikipedia and teaching. Alan’s story is a positive interaction, showing how Wikipedia can feed class discussion and stimulate students to begin contributing to research in a supervised manner. John’s story may *seem* like a negative interaction with Wikipedia, but this was beneficial too. In the classroom, John’s “hunt for 1588” was actively used to illustrate potential pitfalls with Wikipedia and Internet sources in general, and then led into a discussion of research methods.

Even though the Wikipedia experiences were distinct and across two different disciplines, they yielded similar results with respect to the classroom. This indicates that there may be deeper issues here to explore.

A. History Matters.

Is it easy to dismiss Wikipedia as lacking formal academic peer review, taking academic peer review to be a gold standard, but in fact that grossly oversimplifies matters for two reasons. First, academic peer review suffers from its own problems – indeed, it is hard to find a researcher without a laundry list of peer review horror stories – but a complete discussion of this point is outside the scope of this paper. Second, the rich availability of Internet sources, of which Wikipedia is only one example, already forces students and researchers alike to be historians.

Historians often use non-peer reviewed material as primary sources: government records, letters, memoirs, diaries (Tosh, 2002). The problem becomes one of interpreting the sources. The source may not be representative of its time, for instance, and may only give a look into the reality for upper-class society, or for non-minority males (Hudson, 2000). There may also be biases and intentional or unintentional errors (Tosh, 2002). The historian must interpret the source to determine its veracity and its meaning in context.

But this is no longer the refuge of historians; anyone using an Internet source must now do the same. Furthermore, the problem is magnified because it is not just the elite who can publish web pages, and there are new sources of data. Instant messaging and text messages are the conversation transcripts, emails the letters, blogs the diaries of old.

One could argue that nothing pertinent to computer science research would ever appear in these new forums, but in fact that is not true in some areas. Computer history is one obvious example, but instead we look at what should be far removed from history: computer security.

In 2005, Sony was revealed to be installing a “rootkit” on its customers’ computers.

While the definition of a rootkit is outside the scope of this paper, suffice it to say that this was a shocking revelation: rootkits are properly considered to be malicious software. The primary source that broke the rootkit news? Mark Russinovich's blog (Rusinovich, 2005). Recently, another rootkit used in a Sony product was made public, this time by FSecure's blog (Mika, 2007).

And there are more nonacademic, non-peer reviewed primary sources; computer security has heavy industry involvement. Security companies, large and small, produce white papers and have online encyclopedias of malicious software containing sound technical analyses. These are sources that should be, and are, cited in security research. The written output of security companies tends to be at its best and most detailed when talking about malicious threats, however. The details about how a commercial security product operates are rarely more than hinted at in vendor documents, and such details are treated by security company employees as proprietary secrets. The details of commercial products are not secret, as it turns out, but are available to security researchers in another nonacademic form: patent applications. Patent applications are "reviewed" in some sense of the word, but it is often a simple matter for an domain expert to find overlapping claims or prior art. Although clearly biased – the patent applicant has great interest in being awarded a patent – patents (available online) can also be primary sources of research material.

Computer security also has a large underground involvement, resulting in many more potential sources of information with a wide variance in quality. Some, like Aleph One's "Smashing the Stack for Fun and Profit" (Aleph One, 1996) are heavily cited in peer-reviewed academic work; this particular example appeared in *Phrack*, an underground electronic magazine (Phrack, n.d.). There are a plethora of underground, or at least non-mainstream, security conferences as well, like DEFCON (DEFCON, n.d.).

Other underground sources are challenging to a researcher not only because of interpretation problems, but because of their highly ephemeral nature. Often multiple exact copies of an underground document will exist on multiple sites, and there is no clear authoritative location to cite. "The ELF Virus Writing HOWTO" (Bartolich, 2003), for instance, is available from many different web sites.

Other times, the availability of underground sites is unstable, like VX Heavens (VX Heavens, n.d.) (a web site with information about computer viruses) and Google's cache or the Internet Archive's Wayback Machine (Wayback Machine, n.d.) become the only means of acquiring underground source material. This impermanence is a theme we return to in the next section.

One approach to dealing with the lack of academic peer review is to entice the alternative to become mainstream. A 2007 workshop called WOOT (USENIX Workshop on Offensive Technologies) provided 'a forum for high-quality peer reviewed papers for discussing tools and techniques for attack' (USENIX, n.d.), computer security material that might otherwise appear underground. The papers in the workshop program, however, were almost exclusively from universities and technology companies. We conclude from this that attempts at making the underground conform to academic peer review are ultimately doomed to failure, and a wiser strategy is to learn to handle alternative sources.

Given that there are already areas of computer science where non-peer reviewed sources are important, and the growing number of Internet sources, fighting their use is likely a losing battle and a disservice to students. Instead, it makes more sense to embrace the change, and properly train computer science students to use and interpret Internet sources like Wikipedia.

B. Isn't It Romantic?

Equally to the point in our view is the fact that these Wikipedia experiences illustrate an interesting and important dilemma in our understanding of how knowledge is produced and how learning can occur in our culture.

On the one hand, modernity is imbued through and through with the residue of a Romanticism which identifies knowledge as a deep upwelling of an internally located individuality that then manifests itself somehow as art, as science, as leadership or even divine inspiration (Heath and Boreham, 2000). This Romanticism is reflected in such disparate realms as hero worship in politics or sport, the Hollywood cult of celebrity, the winner-take-all compensation of CEOs, prohibitions against “cheating” (collaboration?) in university, the movie stereotype of the mathematical or artistic genius or of the downtrodden loner struggling to overcome all obstacles to success, even the familiar story of tormented star-crossed lovers who intuit themselves to be soul mates (David Black, 2002).

Yet on the other hand, anyone who has spent time in a high-energy particle physics setting or served on a clinical lab sciences editorial review board, to name only a couple of disciplines, knows that sole authors can be few and far between. In the real world, we typically work in groups or teams; those who work alone do so often more out of misanthropy than genius (Aycock and Buchignani, 1995). So which is it: is “knowing” an act of solitude or rather one of sociability?

Our Wikipedia experiences suggest that it is both and neither, but that this equivocation is worth exploring further in the classroom. Fraud emails are no less a product of the imagination than a sociological model of Walmart, yet both are products of a certain collective tradition in which knowledge is recognized to have been sustained and transmitted across generations or even centuries. One happens to be illegitimate while the other is scholarly, but the structures of knowledge and their conditions of production are remarkably analogous (McCarthy, 1996). We contribute to this tradition and even embellish it from time to time, but the body of knowledge peculiar to our culture – Wikipedia is case in point – remains both awkwardly intact yet constantly under construction, whether we are aware of it or not. Are we then summarily to dismiss Wikipedia as an Orwellian memory hole, or instead to exalt it as a “technology of the self” (Aycock, 1995) meant to shape ourselves to the very real burdens of modernity? Foucault’s famous phrase, *subjugated knowledges*, that appears in the title of this paper, alludes precisely to the manner in which knowledge refuses to be entirely the one or the other, continuity or interruption, institution or intuition (Foucault, 1972).

In the classroom, we reinforce this equivocation when we tell our students that they must locate their own work in the larger scholarly tradition that has preceded it, yet if they replicate that work too seamlessly, we prosecute them for plagiarism. The scientific rule of peer review is an institutionalized version of this dilemma; the NPOV rule of Wikipedia is yet another. If I ask my students to cite their references, and they consult Wikipedia to find references that may literally have been fabricated the day before, not much has been gained (Aycock, 1993b). Wikipedia’s notion of authoritative sourcing without authorial accountability is scarcely in our view a satisfactory substitute for peer review, but peer reviewers themselves are routinely acknowledged (if anonymously) as collaborators in the final product. And so it goes. We have constantly to remind ourselves that just as “Walmarting” was self-referential in Alan’s example, so is the quest for disciplinary knowledge, evinced in the historicity of John’s fraud email.

IV. Implications.

In a recent work (Dron, 2007), Dron outlines the play of constraint and choice in E-Learning, with particular reference to the social networking infrastructure that is addressed by such Web 2.0 technologies as wikis, P2P file sharing, Flickr and MySpace, MMOGs and eBay. Popular volumes celebrate the revolutionary impulse of collaboration and the wisdom of crowds (Surowiecki, 2005, Tapscott and Williams, 2006). Like it or not, our students are creatures of a culture in which image and truth, experience and fantasy are difficult to disentangle (Aycock, 1993a). Does this mean that we should simply acquiesce? Not necessarily, but cultural movements must be taken seriously to achieve that teachable moment in the classroom. We cannot really expect students to adopt a more thoughtful approach to the academy's version of knowledge unless we also engage them in the dialog that addresses theirs. The late Richard Rorty's admonition to scholars that above all, we must keep the conversation going, is very much a recognition of the dilemma we have sketched above (Rorty, 1979).

What does this mean? Simply that good teachers ignore changes in technology and culture at their peril. Instead of taking a hardline stance on Wikipedia (Waters, 2007), we should be using Wikipedia as an opportunity (Aycock, 2003). It is an excellent vehicle to discuss with our students not only the use and interpretation of Internet sources, but the management of rapidly changing source material and the role of collaborative work within academic research.

V. Some Pedagogical Strategies.

Here are several strategies that I (Alan) use in my courses to bring the Web into the dialog of teaching and learning:

- I require the students to introduce themselves online, including the usual – where are you from, what do you do for fun – but also ‘what is your favorite YouTube clip?’ The student must then include a link to their YouTube clip and justify it in the context of their interests and background.
- I often use ads from sources such as iFilm to illustrate cultural patterns in American society. Since ads are ubiquitous, they constitute a medium that students are familiar with, yet often taken for granted. The brevity of the ads, e.g., 30-second Super Bowl ads, presents data in a form that allows students to whet their empirical skills.
- Whenever possible, I post a YouTube clip to our own course Web site for the students to view. I choose these clips deliberately for their “edginess” in terms of the current course material. For instance, I teach a module on Islam about which my students are typically both ignorant and curious. I include a YouTube clip on Clotaire K, a French-Lebanese hip-hop artist whose protest songs provide my students a perspective that they don't get in the American news media.
- I may give my students a Wikipedia reference to a concept that we're studying – e.g., moral panic, alienation, panopticon, Neopaganism – but then ask them to use their assigned hardcopy reading or other Web-based materials to identify key elements of the concept that have been omitted or misconstrued in the Wikipedia article. This encourages students to discern what is most central or merely peripheral to an idea, and to learn how to examine different discourses and use them critically against one another.

For my (John's) part, computer science students at the senior undergraduate and even the graduate level are adept at finding material on the Internet, but are not always skilled in assessing the value of that material. References to Wikipedia are legion in written work even at this

advanced level. This is a reminder that familiarity with, and skill using, the Internet does not imply being able to filter what flows back. Furthermore, seeing this happen so late in students' academic programs suggests that these skills need to be introduced immediately upon entry to university, if not sooner. (Alan re-enters the dialog.) I agree, and typically give my students training in the assessment of Web-based materials by requiring them to use a standard rubric such as Beck (2007) to demonstrate their ability to filter Web sources of their choice accurately.

Acknowledgements

The first author's research is supported in part by a grant from the Natural Sciences and Engineering Council of Canada. Thanks to Some Guy, a.k.a. Curtis Hoffmann, for his help; also thanks to the anonymous reviewers for their helpful comments.

References

- Aleph One. (1996). Smashing the stack for fun and profit. *Phrack*, 7(49).
- Aycock, A. (1993a). Derrida/Fortda: Deconstructing play. *Postmodern Culture*, 3(2).
- Aycock, A. (1993b). Virtual play: Baudrillard online. *Arachnet Electronic Journal on Virtual Culture*, 1(7).
- Aycock, A. (1995). Technologies of the self: Michel Foucault online. *Journal of Computer-mediated Communications*, 1(2).
- Aycock, A. and Buchignani, N. (1995). The email murders: Reflections on 'dead' letters. In Jones, S. G., editor, *CyberSociety: Computer-mediated communication and community*, pp. 184–231. Thousand Oaks, CA: Sage.
- Aycock, A. (2003). *Serendipity and SoTL: An ethnographic narrative*. In C. Schroeder and A. Ciccone, (Eds.), *Learning More About Learning*, University of Wisconsin-Milwaukee Center for Instructional and Professional Development.
- Aycock, A. (2006). "This way to the egress": Using exit assignments to improve communication in your course, in P. Rice and D. McCurdy (Eds.), *Strategies in Teaching Anthropology* (4th Ed.). Upper Saddle River, NJ: Prentice-Hall.
- Aycock, J. (2006). Teaching spam and spyware at the University of C@1g4ry. In *Third Conference on Email and AntiSpam*, pp. 137–141. Short paper.
- Aydelotte, F. (1913). *Elizabethan rogues and vagabonds*, (vol. 1). London: Oxford University Press.
- Bartolich, A. (2003). The ELF virus writing HOWTO.
- Beale, P. (1998). *A history of the post in England from the Romans to the Stuarts*. Ashgate.
- Beck, S. (2007). *The good, the bad, and the ugly: or, why it's a good idea to evaluate web resources*. Retrieved March 20, 2008 from <http://lib.nmsu.edu/instruction/eval.html>.

- Bryman, A. (2004). *The Disneyization of society*. Thousand Oaks, CA: Sage.
- David Black, J. (2002). *The politics of enchantment: romanticism, media, and cultural studies*. Wilfrid Laurier University Press.
- DEFCON. Defcon. <http://www.defcon.org>.
- Dron, J. (2007). *Control and constraint in elearning: Choosing when to choose*. Idea Group Publishing.
- Foucault, M. (1972). *The archaeology of knowledge*. London: Routledge.
- Glickman, H. (2005). The Nigerian “419” advance fee scams: Prank or peril? *Canadian Journal of African Studies*, 39(3):460–489.
- Griffiths, A. (1899). *Mysteries of police and crime: a general survey of wrongdoing and its pursuit*, (Vol. 2). Cassell.
- Harvey, D. (1990). *The Condition of postmodernity: An enquiry into the origins of cultural change*. Cambridge, MA: Blackwell.
- Heath, D. and Boreham, J. (2000). *Introducing romanticism*. Totem Books.
- Hudson, P. (2000). *History by numbers: An introduction to quantitative approaches*. Arnold.
- McCarthy, E. (1996). *Knowledge as culture: The new sociology of knowledge*. London: Routledge.
- Metropolitan Police Service. Fraud alert – 419 fraud. Retrieved April 19, 2007 from <http://www.met.police.uk/fraudalert/419.htm>.
- Mika (2007). Double whammy! Another sony case (and it’s not BioShock). *FSecure: News from the Lab*. <http://www.fsecure.com/weblog/#00001263>.
- Morton, J. (1996). The Spanish Prisoner swindle. *New Law Journal*, 146(6728):102.
- Nash, J. R. (1976). *Hustlers and con men*. M. Evans and Company.
- Phrack. <http://www.phrack.org>.
- PowerBerrey, R. J. (1899). *The byeways of crime: With some stories from the Black Museum*. Greening.
- Ritzer, G. (2004). *McDonaldization of society* (rev. ed.). Thousand Oaks, CA: Sage.
- Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton University Press.
- Russinovich, M. (2005). Sony, rootkits and digital rights management gone too far. Mark’s Blog. <http://blogs.technet.com/markrussinovich/archive/2005/10/31/sonyrootkitsanddigitalrights-managementgonetooofar.aspx>.
- Surowiecki, J. (2005). *Wisdom of crowds*. New York: Random House.

Tapscott, D. and Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything*. New York: Portfolio.

The uncovering of a Spanish swindle. (1909). *Green Bag*, 21:396–397.

Tosh, J. (2002). *The pursuit of history*, (rev. 3rd ed.). New York: Longman

U.S. Department of State (1909). The so-called “Spanish swindle”. *Central Law Journal*, 69:452–453.

U.S. Department of State (1940). Warnings to American citizens against the “Spanish swindle”. *Department of State Bulletin*, 2:612–613.

USENIX. WOOT ’07 call for papers. Retrieved September 6, 2007 from <http://www.usenix.org/events/woot07/cfp/>.

VX Heavens. VX Heavens. <http://vx.netlux.org>.

Waters, N. L. (2007). Why you can’t cite Wikipedia in my class. *Communications of the ACM*, 50(9):15–17.

Wayback Machine. Internet Archive Wayback Machine. <http://www.archive.org/web/web.php>.

(Wikipedia, a) Wikipedia. Advance fee fraud. Retrieved April 25, 2007 from http://en.wikipedia.org/w/index.php?title=Advance_fee_fraud&oldid=125842839.

(Wikipedia, b) Wikipedia. Neutral point of view. Retrieved September 6, 2007 from http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=155896552.

(Wikipedia, c) Wikipedia. No original research. Retrieved April 25, 2007 from http://en.wikipedia.org/w/index.php?title=Wikipedia:No_original_research&oldid=124181241.

(Wikipedia, d) Wikipedia. Spanish Prisoner. Retrieved April 25, 2007 from http://en.wikipedia.org/w/index.php?title=Spanish_Prisoner&oldid=121873688.

(Wikipedia, e) Wikipedia. Walmarting. Retrieved September 6, 2007 from <http://en.wikipedia.org/w/index.php?title=Walmarting&oldid=41712078>.

Zuckoff, M. (2006). The perfect mark. *The New Yorker* (May 15, 2006), pp. 36–42.