

Journal of Scholarship of Teaching and Learning



Volume 6
Number 2
October 2006

*JoSoTL is sponsored by
The Mack Center at Indiana University for Inquiry on Teaching and Learning
in association with the Faculty Colloquium on Excellence in Teaching (FACET).*

Journal of Scholarship of Teaching and Learning



Volume 6

Number 2

October 2006

<u>Carol Hostetter & Monique Busch</u>	<u>Measuring Up Online: The Relationship between Social Presence and Student Learning Satisfaction</u>	1
<u>Timothy M. Franz & Tam M. Spitzer</u>	<u>Different Approaches to Teaching the Mechanics of American Psychological Association Style</u>	13
<u>Anthony N. Docan</u>	<u>Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation</u>	21
<u>Jack Bookman & David Malone</u>	<u>Negotiating Roles and Meaning While Learning Mathematics in Interactive Technology-Rich Environments</u>	41
<u>Norman R. Spivey & Andrea Cuthbert</u>	<u>Reciprocal Teaching of Lecture Comprehension Skills in College Students</u>	66
<u>Mark G. Urtel, Rafael E. Bahamonde, Alan E. Mikesky, & Jeff S. Vessely</u>	<u>On-line Quizzing and its Effect on Student Engagement and Academic Performance</u>	84
<u>Maria Mikolchak</u>	<u>Service Learning in English Composition: A Case Study</u>	93
<u>Brian D. Brost & Karen A. Bradley</u>	<u>Student Compliance with Assigned Reading: A Case Study</u>	101

Copyright © 2006 The Board of Trustees of Indiana University

To view documents in this electronic journal, you will need a copy of Adobe Acrobat Reader.
A free copy can be downloaded from the [Adobe website](#).



Measuring Up Online: The Relationship between Social Presence and Student Learning Satisfaction

Carol Hostetter¹ and Monique Busch²

Abstract: The study examined students' perceptions of social presence in online and face-to-face course environments. Data from surveys of 112 undergraduate students (80 in online, 32 in face-to-face classes) are presented. Statistical tests include t-tests and ordinary least squares regression tests. Students' perceptions of social presence were similar in the online and face-to-face sections, and predicted their learner satisfaction scores. Experience in online courses had a statistically significant effect on online students' perceptions of social presence. Students' social presence scores did not have a significant effect on their learning outcomes, perhaps due to the small amount of variation in learning outcomes.

Keywords: Online, social presence, learner satisfaction, learning outcomes.

I. Introduction.

The evolution in higher education from a traditional to a computer-mediated environment creates challenges and opportunities for educators and researchers. As technology transforms the way we teach and learn (Bonk and Cunningham, 1998), computer-mediated options range from email communication with students (Flinn, 1995; Folaron, 1995), to online course environments for posting course materials and facilitating discussions, and finally to totally asynchronous Internet-based course delivery (Twigg, 2003). Computer-mediated education is popular: in 2003, nearly two million U.S. college students took an online class (Carlson, 2004). In an annual survey of university information technology officers, Green reports that helping faculty “integrate technology into instruction” is the top priority on campuses (2001, p.2). Further, Green finds that one-fifth of all college courses utilize technology for course management. Entire degrees may be earned online at some accredited universities (Indiana University, 2003).

Some educators have been hesitant to embrace online education (Mama, 2001; Schoech and Helton, 2002), due in part to questions about the soundness of its pedagogy. At the forefront are concerns about accreditation, which at present applies standards for traditional courses to online courses rather than establishing standards specific to computer-mediated environments (Benson, 2003). More research is needed to thoroughly understand excellent pedagogy in online education, including establishing the criteria by which excellence should be judged. A review of the scholarship of teaching and learning (SoTL) literature begins with excellence in face-to-face education, as a foundation for excellence in online education.

Educational Excellence

A classic in SoTL literature on educational excellence is Chickering and Gamson's *Seven Principles for Good Practice in Undergraduate Education* (1987). The seven principles are contact between students and faculty, cooperation among students, active learning, prompt feedback, time on task, high expectations, and respect for diverse talents and ways of learning (1987, p.3). More recently, research has been conducted on student engagement, developing

¹ Assistant Professor, School of Social Work, Indiana University Bloomington, chostett@indiana.edu.

² Doctoral Candidate and Outcome Project Coordinator, IARCCA, moniquebusch@sbcglobal.net.

Chickering and Gamson's seven principles into measurable variables and evaluating college education across the country. Examples of items about which students have been surveyed are contributing to class discussions, making class presentations, participating in community-based projects as part of course work, working harder than they thought they could to meet an instructor's standards, and critical thinking skills, to name a few (Kuh, 2002). The literature indicates that educational excellence is becoming well-studied.

Related to the above principles of excellence in education is that of *teacher immediacy behavior*, based on Wiener and Mehrabian's early work in psychology (1968). Immediacy can be defined as the amount of "perceived physical and/or psychological closeness between people" (Christophel, 1990, p. 325). Smiling, having a relaxed body posture and position, speaking to the students rather than to the chalkboard, using humor, and modulating the voice are examples of teacher immediacy behaviors. Researchers have linked teacher immediacy behaviors with positive student learning outcomes (Kearney, Plax, and Wendt-Wasco, 1985; Gorham, 1988).

The notion of teacher immediacy behaviors has been further developed into the concept of social presence as first identified by social psychologists Short, Williams, and Christie (1976). Social presence is "the degree to which a person is perceived as a 'real person' in mediated communication" (Gunawardena, 1995, p.151). Immediacy behaviors increase social presence. In examining whether "real" relationships can be established in a "virtual" medium (Jones, 1995, p. 14), studies of social presence and computer-mediated communication have shown that "despite the low social bandwidth of the medium, users of computer networks are able to project their identities whether 'real' or 'pseudo,' feel the presence of others online, and create communities with commonly agreed on conventions and norms ..." (Gunawardena, 1995, p. 151). An examination of the effectiveness of teaching methods to increase social presence in computer-mediated communication is clearly warranted if educators are to learn how to create excellent online course environments.

A. Social Presence in Online Education.

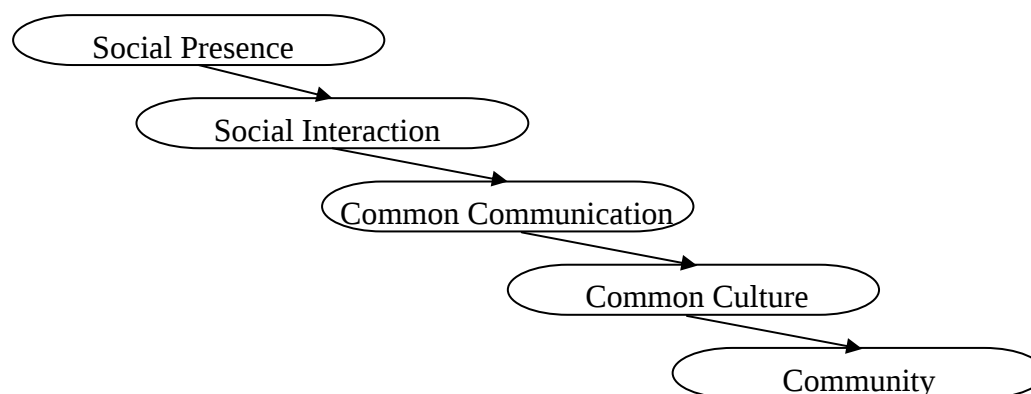
Several studies address social relationships in online education. Mama (2001) compares students' attitudes regarding a site-based and a web-based (with three face-to-face meetings) class, finding that the students in the web-based class felt it was more personal than site-based courses they had taken before. Swan (2002) reports that students perceived online discussions as more equitable and more democratic than traditional classroom discussions, and that there was a positive relationship between levels of interaction among students and student satisfaction in the course. Gunawardena and Zittle (1997) surveyed graduate students in a computer-mediated inter-university conference. They found that social presence predicted student's learning satisfaction. Richardson and Swan (2003) adapted Gunawardena and Zittle's survey instrument for use with mostly nontraditional-aged students in several online courses. Their results indicated that social presence was positively correlated with students' perceived learning.

Further research examines the connection between relationships and learning. Two studies claim to have found a positive correlation between social presence and students' perceptions of their learning (Christophel, 1990; Richardson and Swan, 2003). Support is found for the positive correlation between the level of students' perceptions of social presence in their courses and higher results on learning measures (Picciano, 2002). In a study involving random assignment of students to online or traditional classes, Schutte (1997) finds that the students in the online course perceived a greater amount of peer contact than the traditional classroom-based students and also earned significantly higher grades than the traditional students. Rodriguez, Plax

and Kearney (1996) claim that teacher immediacy behaviors influenced students' affective learning, which in turn influenced students' cognitive learning.

B. The Link between Social Presence and Community.

The importance of human relationships in computer-mediated education, then, is well-supported, but do these relationships translate to an educational community? Before exploring the possible connection between social relationships and community, it is helpful to understand what is meant by community. Researchers do not always provide a definition of community or an explanation of the link between social presence, community, and learning (Jones, 1995; Gunawardena, 1995; Rourke, Anderson, Garrison, and Archer, 1999; Reid, 1995; Swan, 2002; Tu and McIsaac, 2002). Brueggemann (2002) describes community as shared experiences in which both individual and group needs are met, and holds that community can be linked to a place and time but can also transcend them. Rovai (2002) sees community as a group of individual members of formal and informal organizations, interacting and connecting with each other. With this understanding of community, the link between social presence and community can be explored. Wise, Chang, Duffy, and del Valle (2004) state that the concepts of social presence and community both transmit the sense of relating and caring among participants. Gunawardena (1995) asserts that "The development of social presence and a sense of online community becomes key to promoting collaborative learning and knowledge building" (p. 164). She believes that collaborative learning is possible only if participants have social presence, a sense of community, and a common goal. Rovai (2002) evaluated online and traditional classes of 14 professors, finding that certain online classes had significantly higher ratings from students on feelings of community in the classroom. He asserts that the method of teaching, not the environment for delivering the course, is what influences feelings of community. In a qualitative study, Wegerif (1998) found that students' sense of community affected their success in the course. Students who felt more like insiders in the learning community were more likely to achieve success. In a computer-mediated environment, feelings of community and social presence may be considered to be strongly connected to each other and to online interaction (Tu and McIsaac, 2002, p.131). Reid posits another model³ of the link between social presence and community (1995, p. 156):



Learning takes place in a social environment, and cognitive understanding and personal construction of knowledge depend on relations with others (Fung, 2004; Richardson and Swan,

³ Model diagram by Valerie Decker

2003; Vygotsky, 1978). Creating a safe environment for a learning community in class allows students to take risks and collaborate in an authentic manner (Bonk and Cunningham, 1998). Wegerif (1998) contends that it is essential for students to feel that they are members of a community in order to collaborate and learn, and that computer-mediated communication can provide support for the development of feelings of community.

The nature of an asynchronous online course, with educational opportunity available 24 hours a day, seven days a week, can provide a more rapid responsiveness to students' questions and ideas than the traditional face-to-face class (Wise et al., 2004). Teacher immediacy behaviors and social presence among students can be enhanced in an online course, thus creating a learning community that facilitates educational excellence. However, more studies are needed regarding the relationship between online social presence and student learning satisfaction and outcomes, in order to understand more fully what facilitates excellence in computer-mediated education (Rourke, et al., 1999).

II. Purpose of Study.

The study is designed to understand whether social presence can be established in an online college class, at rates similar to a face-to-face class. In addition, the study is designed to understand the relationship between social presence and student learning satisfaction and outcomes. The paper presents data comparing four sections of an online course with two sections of the same course, delivered in a face-to-face environment. All sections used the identical syllabus and assignments. Social presence is selected as a means to investigate online course delivery because, as Rourke and colleagues wrote, "social presence supports cognitive objectives through its ability to instigate, sustain, and support critical thinking in a community of learners" (1999, p. 52). Understanding the social presence aspect of an online class may help educators better understand how to construct an excellent, engaging online teaching/learning environment. Both quantitative and qualitative data were used to explore students' perceptions of social presence in their classes (the qualitative data are reported elsewhere).

Research Questions

The study addresses the questions:

- 1) Are students' perceptions of social presence in an online class similar to those of students in a face-to-face class?
- 2) Do students' perceptions of social presence in the course have an effect on learner satisfaction?
- 3) What variables affect students' perceptions of social presence in the course?
- 4) Do students' perceptions of social presence in the course have an effect on learning outcomes?

III. Method.

A. Procedure.

The study was conducted over two semesters on two campuses, one urban and the other rural. In each semester, two sections of an undergraduate online course were taught by one author (offered on both the urban and the rural campus), and one section of a face-to-face version

of the same course was taught by the other author (on the urban campus). For all sections, surveys were distributed in the last week of the course. The online students were provided the survey electronically, to complete and email to a research assistant. The face-to-face students were given the survey during class, with an envelope to be delivered by the students to the departmental secretary. It was made clear to both online and face-to-face students that their answers would be anonymous. Completed surveys were returned by 112 out of 128 students, for a response rate of 88%.

B. Participants.

Of the 128 potential participants, 124 were female, and the vast majority were of traditional age. Seventeen percent of the students in all classes were students of color, self-identifying either as African American or Hispanic American. Of the 112 students responding to the survey, 15% were students of color. This number is slightly higher than the nine percent of the student body on both campuses who identify as African American or Hispanic American. Data was not collected on gender or any other demographic information for respondents, due to the risk that such information might inadvertently identify the students.

C. Instrument.

The study used a survey instrument originally designed by Gunawardena and Zittle (1997) and modified by Richardson and Swan (2003)⁴. The wording of the questions was slightly altered to reflect the specific course. Students in the online sections were asked about the number of online courses they had taken. Ten questions were asked about students' perceptions of the class and their expectations for learning, using Likert-scale responses (ranging from 1=Strongly Disagree to 6=Strongly Agree). Two sections of the survey are not presented here. One asked about specific course activities, and the other presented 12 open-ended questions about students' satisfaction, learning, and feelings of community.

IV. Results.

A. Factor Analysis.

A factor analysis was performed to help determine the factors underlying the 10-item scale; two factors were identified. One factor (Social Presence) has to do with students' perceptions of social presence, which accounts for the items:

- I felt comfortable conversing through this medium/in class.
- I felt comfortable participating in course discussions.
- I felt comfortable interacting with other participants in the course.
- I felt that my point of view was acknowledged by other participants in the course.
- I was able to form distinct individual impressions of some course participants.

The second factor (Learner Satisfaction) was students' satisfaction with their expectations of the learning community. The following questions are related to this factor:

- The instructor created a feeling of community

⁴ Permission was given by Gunawardena and Zittle, and Richardson and Swan, for use of the instrument.

- The instructor facilitated discussions in the course.
- My level of learning that took place in this course was of the highest quality.
- Overall this course met my learning expectations.
- Overall the instructor for this course met my expectations.

The reliability for the 10-item scale is a Cronbach's alpha of 0.87. Table 1 shows the means for the two factors. The mean response for the individual Social Presence statements was 5.17, and the mean response for the individual Learner Satisfaction statements was 5.34. Both sets of means indicate that students, on average, rated the items between "agree" (5) and "strongly agree" (6).

Table 1. Means and Standard Deviations for Social Presence Scale.

Factor	Mean	Std. Dev.
Social Presence overall	25.83	3.28
Social Presence per item	5.17	0.62
Learner Satisfaction overall	26.69	3.08
Learner Satisfaction per item	5.34	0.66

N=112

B. Results for Research Questions.

The first question asks if online students' perception of social presence is similar to those of face-to-face students. An independent samples t-test was conducted, with no significant difference found between the online and face-to-face classes ($t = 1.87$, $p > 0.05$). This is important information not in terms of which course environment is superior, but in terms using the face-to-face version as a benchmark for how much social presence is needed to facilitate a feeling of community in an online course.

The second research question is concerned with whether Social Presence scores have an effect on Learner Satisfaction scores. Table 2 shows an ordinary least squares (OLS) regression test of the effect of Social Presence on Learner Satisfaction, indicating that Social Presence is a predictor of Learner Satisfaction (0.60, $p < 0.001$). Social Presence explained 40% of the variance in the Learner Satisfaction scores. This finding suggests that facilitating social presence in an online class is important for students' satisfaction in their learning.

Table 2. OLS Regression Analysis for the Effect of Social Presence on Learner Satisfaction.

Variable	B	SE B	β
Social Presence	0.60***	0.07	0.64
R ²	0.40		

N=112

*** $p < 0.001$

The third research question asks what variables affect students' perceptions of Social Presence. Many potentially influential variables, such as demographic information, were not tested in the study due to the need to protect students' anonymity. However, it was possible to observe students' previous online experience, as well as the campus at which each student was enrolled. Students in the online sections were asked if this was their first online course, if they

had taken two online courses including the current one, or if they had taken more than two online courses including the current one. Table 3 shows the results of an OLS regression test. The higher the number of online courses students had taken, the more positive their perceptions of Social Presence (0.82, $p < 0.05$) in this online course. Very little of the variance in the scores is explained by the number of online courses, but the effect is statistically significant. To examine the differences between the students at different campuses, an independent samples t-test was performed to compare the means of the Social Presence scores of the two groups. The results are shown in Table 4. This test included all urban students and all rural students, as opposed to the regression test above which included only students in the online courses. A significant difference was found ($t = 2.81$, $p < 0.01$), indicating that students on the urban campus had significantly higher Social Presence scores than students on the rural campus.

Finally, the fourth question addresses whether students' perceptions of social presence have an effect on learning outcomes. An OLS regression test was performed to test the effects of the Social Presence factor on students' scores on their individual papers. No significant effect was found.

Table 3. OLS Regression Analysis for the Effect of Number of Online Courses on Social Presence.

Independent Variable	Social Presence		
	<i>B</i>	<i>SE B</i>	β
Number of online courses	0.92*	0.39	0.26
R^2	0.07		

N=80

* $p < 0.05$

Table 4. Independent Samples T-Test of the Social Presence Scores of Rural vs. Urban Students.

Students' location	Group Mean	SD	t
Urban (N=71)	26.50	2.82	2.81**
Rural (N=41)	24.68	3.73	

N=112

** $p < 0.01$

C. Discussion.

The findings from this study show that an online class can be designed that facilitates students' perceptions of social presence at rates similar to those in a face-to-face class. In fact, the social presence scores for both types of courses are relatively high, with a mean of 5.17 overall. This number is between "agree" and "strongly agree" on the six-point scale. These findings are relevant to those of a study of graduate students in an online conference, Gunawardena and Zittle (1997). The authors apparently summed all the items in the social presence scale to obtain an overall mean of 3.5. This number is halfway between "uncertain" and "agree" on a five-point Likert scale (strongly disagree to strongly agree). The reliability for the

social presence scale in their study is a Cronbach's alpha of 0.88 Gunawardena and Zittle (1997). Similarly, Swan and Richardson's (2003) study of returning students in a variety of online courses combined several survey items regarding social presence, and obtained a mean of 4.4. This figure lies between "somewhat agree" and "agree" on their six-point Likert scale. The findings of the current study also are similar to those comparing students' attitudes in face-to-face class versus a hybrid (online with three face-to-face sessions), in which the students in the hybrid class felt it was more personal than previous face-to-face courses they had taken (Mama, 2001). These results may be seen as encouraging evidence for those teaching online courses who have been concerned that the human contact and sense of a learning community might be hindered by a computer-mediated environment.

The measure Learning Satisfaction was also relatively high in the current study, with an overall mean of 5.3 (between "agree" and "strongly agree"). The study suggests that students' perceptions of social presence in an online course positively influence their satisfaction with their learning in the course. This is similar to the study by Gunawardena and Zittle (1997), in which social presence scores strongly predicted student satisfaction scores. Their study obtained a mean for summed items in the satisfaction scale of 3.3, just over "uncertain," and the reliability for the satisfaction scale was a Cronbach's alpha of 0.87. Their students' social presence scores explained 60% of the variance in satisfaction scores, whereas in the current study social presence scores explained 40% of the variance. A study by Swan (2002) indicated that online class discussions appeared to students to be more egalitarian than those in face-to-face classes. Richardson and Swan's study (2003) also found that social presence was positively correlated with students' perceived learning.

The third research question asks what variables affect students' perceptions of social presence in the course. It was found that for students in the online class, taking more online classes positively influenced their Social Presence scores. It may be that students with more experience in online courses have developed specialized learning skills and thus are better able to utilize the computer-mediated learning environment. Perhaps they understand the necessity for students to contribute to the learning community, since there is no "sage on the stage" (King, 1993, p. 1). In addition, they may have a better perspective on how people "project their identities" (Gunawardena, 1995, p. 151) in an online course.

In the next part of the third research question, Social Presence scores were compared by campus. The results showed that students on the urban campus had significantly more positive perceptions of social presence in their courses, whether they were face-to-face or online. This is a perplexing finding which clearly requires further research.

The final research question found that students' perceptions of social presence had no significant effect on learning outcomes. This stands in contrast to studies by Picciano (2002), Schutte (1997), and Rodriguez, Plax and Kearney (1996). An explanation may be that the course is a senior seminar in which students generally earned high grades on their papers. Perhaps there is not enough variation in their grades to adequately test this question. Further research should be done with students in introductory courses, where a greater range of grades is found.

The current study sought to examine students' perceptions of social presence for traditionally-aged undergraduates in a course offered both online and face-to-face. In this study, the social presence scale, as originated by Gunawardena and Zittle (1997) and modified by Richardson and Swan (2003), is clarified and extended. Neither of the prior studies used factor analysis to understand the themes underlying the items in the instrument. Gunawardena and Zittle (1997) summed all the items in the scale; Richardson and Swan (2003) combined the

answers to several survey items, although it is not apparent which individual items were used. The use of factor analysis in the current study should help other researchers utilize the scale to study social presence in computer-mediated learning environments. Further, the study finds results for traditionally-aged undergraduate students that are similar to those with graduate students and returning students, in that Social Presence is a predictor of Learner Satisfaction.

D. Limitations.

The generalizability of these findings is affected by several limitations. The sample is somewhat small and was not randomly selected. This is consistent, however, with much of the research in the scholarship of teaching and learning. The statistical tests used were appropriate for the sample size and a Levene's test suggested that the variance of the scores was approximately equal for the groups under comparison. Another potential limitation of these results is that the measures use self-report and thus respondents may be providing socially desirable answers. Assuring students of their anonymity was used to minimize this possibility, but it is granted that the amount of social desirability is unknown.

V. Conclusion.

A strength of the study is the use of the social presence scale originated by Gunawardena and Zittle (1997) and modified by Richardson and Swan (2003). The previous two studies examined the perceptions of students in computer-mediated environments, the former in a graduate conference and the latter in several courses taken predominantly by students not of traditional college age. Applying the scale to undergraduate education, both online and face-to-face, extends the knowledge developed in the previous studies. Performing a factor analysis to identify the two factors underlying the scale, Social Presence and Learner Satisfaction, hones the usefulness of the scale for future researchers.

The study has implications for educators teaching online courses. Information provided by the study may be used to support pedagogy that increases social presence. Many authors see the benefit provided to students who interact with faculty and each other in a learning community. Each learning community, then, has the opportunity to increase students' collaborative skills in future social environments. As more courses are offered entirely online, it clearly is important to establish standards for excellence in computer-mediated education.

References

- Benson, A. (2003). "Dimensions of quality in online degree programs." *The American Journal of Distance Education*, 17(3), 145-159.
- Bonk, C. J., and Cunningham, D. J. (1998). "Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools." In C. J. Bonk, and K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (pp. 25-50). Mahwah, NJ: Erlbaum.
- Brueggemann, W.G. (2002). *The Practice of Macro Social Work*. 2nd Ed. Belmont, CA: Brooks/Cole.

Carlson, S. (2004). "Online-Education Survey Finds Unexpectedly High Enrollment Growth." *Chronicle of Higher Education*, 5(14), 11-26-04. Retrieved 7-31-05.

<http://chronicle.com/weekly/v51/i14/14a03001.htm>

Chickering, A.W. and Gamson, Z.F. (1987). "Seven principles for good practice in undergraduate education." *AAHE Bulletin*, 39(7), 3-7.

Christophel, D. (1990). "The relationships among teacher immediacy behaviors, student motivation, and learning." *Communication Education*, 39, 323-340.

Flinn, J. (1995) "Use of electronic mail to promote computer literacy in social work undergraduates." *Journal of Teaching in Social Work*, 12(1/2), 73-83.

Folaron, G. (1995). "Enhancing learning with e-mail." *Journal of Teaching in Social Work*, 12(1/2), 3-18.

Fung, Y. Y. H. (2004). "Collaborative online learning: Interaction patterns and limiting factors" *Open Learning*, 19(2), 135-149.

Gorham, J. (1988.) "The relationship between verbal teacher immediacy behaviors and student learning." *Communication Education*, 37, 40-53.

Green, K. (2001). "eCommerce comes slowly to the campus." *The Campus Computing Project* (October). Retrieved 10-10-03. <http://www.campuscomputing.net/pdf/2001-CCP.pdf>

Gunawardena, C. N. (1995). "Social presence theory and implications for interaction and collaborative learning in computer conferences." *International Journal of Educational Telecommunications*, 1(2/3), 147-166.

Gunawardena, C. N. and Zittle, F.J. (1997). "Social presence as a predictor of satisfaction within a computer-mediated conferencing environment." *The American Journal of Distance Education*, 11(3), 8-26.

Indiana University. (2003). IU *Online and Distance Education*. Retrieved 10-12-03. <http://www.iuonline.iu.edu/de/degrees.html>

Jones, S.G. (1995). *Cybersociety*. Newbury Park, CA: Sage.

Kearney, P., Plax, T., and Wendt-Wasco, N. (1985). "Teacher immediacy for affective learning in divergent college classes." *Communication Quarterly*, 3(1), 61-74.

King, Alison (1993). "From Sage on the Stage to Guide on the Side," *College Teaching*, 31(1).

Kuh, G.D. (2002). *The national survey of student engagement: Conceptual framework and overview of psychometric properties*. Retrieved 10-5-03.

http://222.indiana.edu/~nsse/html/2002_NSSE_report/html/conceptual_1.htm

Mama, R.S. (2001). "The classroom or the Internet? Student experiences and research results of supervision in social work class." Unpublished paper. Retrieved 10-9-03.
<http://www.cosw.sc.edu/conf/tech/2001/abstracts/PDFs/Mama.pdf>.

Picciano, A.G. (2002). "Beyond student perceptions: Issues of interaction, presence and performance in an online course." *Journal of Asynchronous Learning Networks*, 6(1).

Reid, E. (1995). "Virtual worlds: Culture and imagination." In S.G. Jones (Ed.), *Cybersociety: Computer-mediated Communication and Community* (pp.164-183). Thousand Oaks, CA: Sage.

Richardson, J. and Swan, K. (2003). "Examining social presence in online courses in relation to students' perceived learning and satisfaction." *Journal of Asynchronous Learning Networks*, 7(1). Retrieved 7-11-03. <http://www.aln.org/publications/jaln/v7n1/pdf/v7n1richardson.pdf>.

Rodriguez, J.L., Plax, T.G. and Kearney, P. (1996). "Clarifying the relationship between teacher nonverbal immediacy and student cognitive learning: Affective learning as the central causal mediator." *Communication Education*, 45, 293-305.

Rourke, L., Anderson, T., Garrison, D.R., and Archer, W. (1999). "Assessing social presence in asynchronous text-based computer conferencing." *Journal of Distance Education*, 14(2), 50-71.

Rovai, A.A.P. (2002). "A preliminary look at the structural differences of higher education classroom communities in traditional and ALN courses." *Journal of Asynchronous Learning Networks*, 6(1), 41-56. Retrieved 8-1-05. http://www.sloan-c.org/publications/jaln/v6n1/v6n1_rovai.asp

Schoech and Helton. (2002). "Qualitative and quantitative analysis of a course taught via classroom and internet chatroom." *Qualitative Social Work*, 1(1), 111-124.

Schutte, J.G. (1997) "Virtual teaching in higher education: The new intellectual superhighway or just another traffic jam?" Retrieved 8-1-05. <http://www.csun.edu/sociology/virexp.htm>

Short, J., Williams, E., and Christie, B. (1976). *The Social Psychology of Telecommunications*. London: John Wiley.

Swan, K. (2002). "Building learning communities in online courses: The importance of interaction." *Education, Communication and Information*, 2(1), 23-49.

Tu, C. and McIsaac, M. (2002). "The relationship of social presence and interaction in online classes." *The American Journal of Distance Education*, 16(3), 131-150.

Twigg, C.A. (2003). "Improving learning and reducing costs: New models for online learning." *Educause Review*, 38(5), 28-38. Retrieved 8-1-05.
<http://www.educause.edu/ir/library/pdf/erm0352.pdf>

Vygotsky, L.S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.

Wegerif, R. (1998). "The social dimension of asynchronous learning networks." *Journal of Asynchronous Learning Networks*, 2(1), 34-49.

Wiener, M., and Mehrabian, A. (1968). *Language within Language: Immediacy, a Channel in Verbal Communication*. New York: Appleton-Century-Crofts.

Wise, A., Chang, J., Duffy, T. and del Valle, R. (2004). "The effects of teacher social presence on student satisfaction, engagement, and learning." *Journal of Educational Computing Research*, 31(3), 247-271.

Different Approaches to Teaching the Mechanics of American Psychological Association Style

Timothy M. Franz¹ and Tam M. Spitzer²

Abstract: Students have to learn two distinctly different tasks when writing research papers: a) creating and organizing prose, and b) formatting a manuscript according to the nuances and mechanics of a pre-determined format, such as Modern Language Association (MLA) or American Psychological Association (APA) guidelines. Two studies examined different approaches for teaching the details of APA style, including: a template, a checklist, and a combined approach that used both the template and checklist. The results demonstrated that while each technique individually helped students to learn APA style, using both together appeared to provide the most help to students.

Keywords: APA format, APA style, writing mechanics, teaching writing.

I. Introduction.

Writing research papers “serves as an important socialization experience” for psychology students (Madigan, Johnson, and Linton, 1995, p. 428) and is one of the more challenging tasks they face. According to Ault (1991), it “is a complicated task for undergraduates because they are creating and organizing the prose while trying to follow format conventions” (p. 45). Thus, students simultaneously have to learn two different tasks: a) creating and organizing prose, and b) formatting a manuscript according to predetermined guidelines, such as those in the *MLA Handbook for Writers of Research Papers* (Gibaldi, 2003) or the *Publication Manual of the American Psychological Association* (2001).

While there are many resources available to help students create and organize prose, relatively few are available that focus solely on teaching the mechanics and nuances of a specific formatting style. However, students often express frustration with formatting a manuscript according to the predetermined guidelines, which can become the focus of their efforts and undermine the quality of their prose. Our goal was to investigate ways to simplify teaching the format conventions – specifically, the formatting guidelines required by APA (i.e., APA style) – so that students could instead focus on their prose.

Although there are many methods that can be used to teach APA style in the classroom, the present study compared the effectiveness of two. The first method used was a template. By providing students with a template, they can have a paper that can serve as a model when writing future papers. Models are useful when learning writing (Braaksma, Rijlaarsdam, and van den Bergh, 2002) because they serve as an information source and provide practice. The second method was a checklist (Stahl, 1987). This can be helpful because it reminds students about necessary format conventions in the *Publication Manual* (2001) and helps them assess whether they have matched the conventions.

¹tfranz@sjfc.edu, 585/385-8170.

²tspitzer@sjfc.edu, both at Department of Psychology, St. John Fisher College, 3690 East Avenue, Rochester, NY 14618.

The purpose of this paper is to report on two studies testing the effectiveness of these two different techniques to teach APA style. We accomplished this by assessing the students' knowledge of APA style at the beginning of the course, varying our teaching techniques in our sections of a similar course (a psychology laboratory course taught within a content area), and then again assessing the students' knowledge of APA style. In Study 1, the course taught by the first author used the template, while the one taught by the second author used the checklist. To help to control for potential course differences, we conducted a second study. In Study 2, the course taught by the first author again used the template, while the course taught by the second author used both the checklist and template.

II. Study 1.

A. Method.

Design. The study utilized a 2 (Teaching technique: Template versus checklist) $\times$ 2 (Time of testing: Pretest-posttest) mixed-factorial quasi-experimental design.

Participants. Participants included 43 students in two different laboratory classes (1 male and 42 females; $n = 21$ and $n = 22$ in the template and checklist classes, respectively). All participants were told that their participation was voluntary – it did not impact their grade in any way – and required their consent to participate. Because of tardiness and attendance, only 33 of the students completed both the pretest and posttest ($n = 16$ and $n = 17$ in the template and checklist classes, respectively). Students were juniors or seniors, had completed approximately 18-30 credits in Psychology, and received some past elementary instruction in APA style, which included at least one research course where students wrote one paper in APA style and potentially another 200-level psychology course where a literature review paper was required.

Procedure. As a pretest, the authors assessed students' knowledge of APA style during the second week, prior to any additional instruction on APA style. Students met in a computer laboratory and, after consenting to participate, received copies of *PsycInfo* abstracts of an article (Duncan, 1976) and book (Eiser and Stroebe, 1972). Duncan's article examined participants' attributions of an ambiguous action by a black versus a white target. The book by Eiser and Stroebe presented a theory of social judgment that attempts to explain, in part, attributions. They were also provided with the basics of a research study (modeled after Sagar and Schofield, 1980), which examined the attributions of children about ambiguous actions of black versus white targets. They used these bulleted details to create the method in their paper as if it were their study. As can be seen, these three papers were chosen for their relevance to the content of a social psychology laboratory class and a developmental laboratory class.

To maintain anonymity and keep instructors blind while allowing a method of comparing across the two classes?, students used pseudonyms on all materials. Students were instructed to type a Title page, an Introduction using the two abstracts, a Method section, and a Reference page. Specifically, we provided students with the following instruction:

Please write a short paper, in APA style, that includes the following sections: 1. A title page; 2. A short one to two paragraph introduction; 3. A method section; and 4. A reference page. Please skip all other sections. For each section you write, use complete APA format to the best of your ability.... You have 35 minutes to complete this. Please use your time wisely. Two references are provided to use in

your introductory paragraph. Please use some of the content when writing your paragraph, and cite and reference them appropriately.

Because of time constraints, other sections (Results, Discussion, etc.) were excluded. Students had no access to the *Publication Manual* (2001), and had approximately 35 minutes to write the papers.

In the week following the pretest, each instructor taught APA style. The first author used the template while the second author used the checklist. The semester then progressed as planned. Students wrote three research papers and instructors provided feedback on the first two, including comments on the proper use of APA style. At the end of the semester, a posttest session was conducted using the same materials and procedures as the pretest; these papers were the posttest data. After this session, we debriefed students and answered questions.

Template. During the third week of class students received a copy of an eight-page manuscript³ written about and typed in APA style (Downing and Franz, 2002). As a homework assignment, each student typed the manuscript verbatim, retaining the formatting and returning it to the instructor five days later. This technique had three goals. First, students read a summary of the format conventions contained in the *Publication Manual* (2001) while re-typing the paper. Second, when done, students had a paper and electronic document that could serve as a template for subsequent papers. Third, it generated a question and answer session about APA style.

Checklist. A checklist was distributed to students in the third week of classes (Spitzer, 2002). The checklist included general items about APA style. The checklist also included corresponding page numbers indicating where key information was located in the *Publication Manual* (2001). The checklist had two goals. First, students would know, in advance, the items on which they would be graded. Second, the checklist would encourage students to use the *Publication Manual* (2001) by referring them to specific pages to find information.

Scoring. To keep our scoring unbiased, we (the authors/course instructors) had a laboratory assistant code each paper as to its course and pretest/posttest statuses, record the codes on a piece of paper, place that paper with the codes in a sealed envelope, and then completely randomize the order in which we read the papers. Because of the pseudonyms and codes, we were completely blind to pretest/posttest and class. To measure the students' knowledge of APA style, we then both scored each paper for adherence to APA style (ignoring the quality of the prose). Each section (Title Page, Introduction, Method, References) was scored on a 1-7 scale, where 1 represented no adherence to APA style at all and 7 represented perfect adherence to APA style. A Pearson correlation between the two sets of ratings demonstrated they were reliable ($r = .96, p < .001$), thus demonstrating inter-rater reliability. These correlations did not differ substantially by section, technique, or pretest-posttest. We then discussed any disagreements and agreed upon one score for each section of the document. Analyses of internal consistency of these four rating scores (Title page, Introduction, Method, and References) demonstrated they were sufficiently reliable, average $\alpha = .86$. Thus, the four rating scores were averaged to create one overall evaluation score for each manuscript.

³ We recognized in our research that eight pages was too long, and a shorter version would be more useful. The Downing and Franz (2002) manuscript has been shortened to six pages since collecting the data presented in these studies.

B. Results and Discussion.

Because of the nature of the quasi-experimental design, we examined pre-existing differences in knowledge of APA style between the classes (see Table 1). A t-test of the pretest scores revealed no significant differences, $t(31) = 1.7, p = 0.10$.

A mixed-factorial ANOVA on the evaluation scores revealed a significant main effect for time of testing, $F(1, 31) = 230.0, p < 0.001, \omega^2 = 0.67$, demonstrating that the scores were significantly higher on the posttest than the pretest. This main effect was qualified by a significant teaching technique by time of testing interaction, $F(1, 31) = 12.3, p < 0.001, \omega^2 = 0.07$. Thus, students using either technique improved considerably, although the skill with APA style improved more with the template than with the checklist.

Table 1. Study 1: Means and standard deviations for pretest and posttest average evaluation scores as a function of teaching technique.

APA Style Teaching Technique	Pretest		Posttest	
	Mean	SD	Mean	SD
Template	2.92 _a	0.94	5.94 _b	0.59
Checklist	3.62 _a	1.34	5.50 _c	0.97
Overall	3.28 _a	1.20	5.71 _b	0.83

Note: Different subscripts represent means that are significantly different at $p < 0.05$.

Study 1 demonstrated that the template and checklist were both effective methods of teaching APA style. Second, the template appeared to be more effective than the checklist, although this effect size was small. Although several students complained about “how boring” the template copying assignment was, some also later commented about the usefulness of the electronic template when typing papers.

However, conversations during debriefing revealed diffusion across the classes – two students in the template class reported “borrowing” checklists from peers. Both commented on the usefulness of combining both approaches. In addition, the results from Study 1 could possibly be due to instructor/feedback differences rather than teaching approach. As a result, a second study was conducted to compare the effectiveness of a “combined” approach, using the template and checklist together, to the template alone. In addition, the instructor-teaching approach combination varied from that in Study 1. Specifically, the second author, who used the checklist in Study 1, used the combined approach in Study 2. The first author, who used the template in Study 1, continued to use the template in Study 2. Thus, if the differences in Study 1 were due solely to instructors, then the class using the template approach should again be more effective in Study 2 than the class using the combined approach. On the other hand, if the teaching approach rather than instructor caused the difference, the class using the combined approach should be at least as effective (and potentially more effective) than the class using the template.

III. Study 2.

A. Method.

Design. The study utilized a 2 (Teaching technique: Template versus combined template and checklist) $\times$ 2 (Time of testing: Pretest-posttest) mixed-factorial quasi-experimental design.

Participants. Participants included 35 students in two different laboratory classes (8 males and 27 females; $n = 15$ and $n = 20$ in the template and combined classes, respectively). Because of tardiness and attendance, only 31 students completed both the pretest and posttest ($n = 12$ and $n = 19$ participants in the template and combined classes, respectively).

Procedure. The procedures were the same as Study 1, except that the second author (who used the checklist in Study 1) used both the template and checklist in Study 2.

Scoring. We scored each paper as we did in Study 1. A Pearson correlation demonstrated that the evaluation scores were again reliable ($r = 0.97$, $p < 0.001$), and these correlations again did not differ substantially by section, technique, or pretest-posttest. Analyses of the internal consistency demonstrated that the scores were sufficiently reliable, average $\alpha = 0.79$, and thus, the four rating scores were again averaged.

B. Results and Discussion.

We first examined the pretest scores (see Table 2) to determine whether there were pre-existing differences in knowledge of APA style between the two classes. A t-test revealed no significant differences, $t(29) = 0.76$, $p = 0.46$.

A mixed-factorial ANOVA revealed a significant main effect for time of testing, $F(1, 29) = 145.7$, $p < 0.001$, $\omega^2 = 0.51$. This main effect was qualified by a significant teaching technique by time of testing interaction, $F(1, 29) = 12.2$, $p < 0.001$, $\omega^2 = 0.05$, showing that students learned APA style better when taught using both the template and checklist combined than with the template alone.

Table 2. Study 2: Means and standard deviations for pretest and posttest average evaluation scores as a function of teaching technique.

APA Style Teaching Technique	Pretest		Posttest	
	Mean	SD	Mean	SD
Template	3.90 _a	1.04	5.18 _b	1.15
Template and Checklist	3.61 _a	1.00	5.96 _c	0.67
Overall	3.72 _a	1.01	5.65 _b	0.94

Note: Different subscripts represent means that are significantly different at $p < 0.05$.

IV. General Discussion.

The results of the two studies indicate that students improved their skills with APA style whether instructors used a template, a checklist, or both together. The results from Study 1 and 2, taken together, also make it likely that the effects are due to teaching technique rather than

instructor. The teaching approach and instructor combinations were varied in Studies 1 and 2. Thus, if the effects in Study 1 were solely due to the instructor, we should have seen the same pattern of results due to instructor in Study 2. However, the results showed the opposite, suggesting that the results are due to teaching approach rather than instructor.

As with most research, the reader should exercise caution when interpreting the results because of several limitations. First, the evaluation focused on APA style while ignoring “creating and organizing prose” (Ault, 1991, p. 45). This was because the main goal of the study was to make it easier for students to understand APA style so they would focus on and improve their prose rather than dwell on formatting issues. Unfortunately, we did not test this second question in this study and leave that for future research. Second, we used quasi-experimental designs, which cannot absolutely rule out differences due to instructors. However, it is more parsimonious to attribute the results of both studies together to the teaching approach. Third, because of time constraints students did not use the *Publication Manual* (2001) while writing their papers. Although this is a different from what instructors often encourage, these pretest-posttest differences would only be expected to improve if students did use the manual. One remaining concern is that the checklist actually encourages students to use the manual, so the effect of the checklist may actually be underrepresented in the evaluation scores. Fourth, a certain portion of the pretest-posttest effect is certainly due to instructor feedback. However, if the results were solely due to feedback, we would not expect any differences due to technique(s) used.

Finally, the present research focused only on APA style because that is the predominant format used in the psychology courses that served as the sample and it is the style used by the authors. However, there are many other formats that students can use (e.g., Modern Language Association/MLA and Council of Biological Editors/CBE). While the paper and checklist used in this study are not directly transferable to these other styles, they could easily be modified (or similar ones could be created) that would be likely to yield comparable results in the classroom.

In conclusion, teaching approaches like the template and checklist appear to help students learn formatting conventions especially when they are used together. There are other teaching approaches that could be used (e.g., Addison, 2000; Ault, 1991; Madigan et al., 1995; Peden, 1994; Rileigh, 1998; Smith and Eggleston, 2001), but many of these emphasize prose *and* format while the template and checklist approaches used in this study emphasize format only. Although well-written and organized prose is essential to writing papers, our goal was to separate formatting conventions from prose and make the difficult task of following the requirements easier for students, allowing instructors to spend more time helping students improve prose and organization.

Author Notes

An earlier version of Study 1 was presented at the 2003 Teaching of Undergraduate Psychology/Midwestern Psychological Association Joint Conference, Chicago, IL. We thank the students in our laboratory classes for participating in the study, and Valerie Cole, John Mavromatis, Laura Phelan, and Dawn Rager for commenting on earlier versions of this paper.

References

- Addison, William. E. (2000). "Student Research Proposals in the Experimental Psychology Course. In Mark. E. Ware, & David. E. Johnson (Eds.)," ***Handbook of Demonstrations and Activities in the Teaching of Psychology***, 2nd Ed., Vol. 1, pp. 226-227. Mahwah, NJ: Erlbaum.
- Ault, Ruth L. (1991). "What Goes Where? An Activity to Teach the Organization of Journal Articles." ***Teaching of Psychology***, Vol. 18, pp. 45-46.
- Braaksma, Martine A. H., Rijlaarsdam, Gert, and van den Bergh, Huub (2002). "Observational Learning and the Effects of Model-Observer Similarity." ***Journal of Educational Psychology***, Vol. 94, pp. 405-415.
- Downing, Leslie L., and Franz, Timothy (2002). "American Psychological Association (APA) Format Requirements for Research Papers in Psychology Courses." Available on-line at <http://keep2.sjfc.edu/faculty/tspitzer/>.
- Duncan, Birt L. (1976). "Differential Social Perception and Attribution of Intergroup Violence: Testing the Lower Limits of Stereotyping Blacks." ***Journal of Personality and Social Psychology***, Vol. 34, pp. 590-598.
- Eiser, J. R., and Stroebe, W. (1972). ***Categorization and Social Judgment***. New York: Academic Press.
- Gibaldi, Joseph. (2003). ***MLA Handbook for Writers of Research Papers*** (6th ed.). New York: Modern Language Association.
- Madigan, Robert., Johnson, Susan., and Linton, Patricia. (1995). "The Language of Psychology: APA Style as Epistemology." ***American Psychologist***, Vol. 50, pp. 428-436.
- Peden, Blaine. F. (1994). "Do Inexperienced and Experienced Writers Differentially Evaluate Ault's (1991) 'What Goes Where' Technique?" ***Teaching of Psychology***, Vol. 21, pp. 38-40.
- Publication Manual of the American Psychological Association*** (5th ed.). (2001). Washington, D.C.: American Psychological Association.
- Rileigh, Kathryn. K. (1998). "Teaching Communication Skills in Psychology." ***Teaching of Psychology***, Vol. 25, pp. 279-282.
- Sagar, Andrew. H., and Schofield, Janet. W. (1980). "Racial and Behavioral Cues in Black and White Children's Perceptions of Ambiguously Aggressive Acts." ***Journal of Personality and Social Psychology***, Vol. 39, pp. 590-598.
- Spitzer, Tam M. (2002). "American Psychological Association Checklist." Available on-line at: <http://keep2.sjfc.edu/faculty/tspitzer/>.

Smith, Gabie. E., and Eggleston, Tami. J. (2001). "Comprehending APA Style Through Manuscript Analysis." ***Teaching of Psychology***, Vol. 28, pp. 108–110.

Stahl, Jeanne. M. (1987). "Facilitating Use of the *APA Publication Manual* in the Scientific Writing Course," In V. P. Makosky, L.G. Whittemore, & A. M. Rogers (Eds), ***Activities handbook for the teaching of psychology***, Vol. 2, pp. 311-314. Washington, D.C.: American Psychological Association.

Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation

Tony N. Docan¹

Abstract: This study examined how particular grading systems motivate students. Since competency-based grading and point systems are most prevalent (Hendrickson and Gable, 1999), the current study is modeled around these systems. The grading systems used for this study were divided into two categories and defined as those students who earned their grades (“earners”) and those who maintained their grades (“maintainers”). The earners started the semester with 0 points and added points with each graded assignment, whereas the maintainers were given the maximum number of points available for the course at the beginning of the semester and then subtracted points from this overall total as they lost points on a graded assignment. The earners received positive incentives (i.e., the addition of points), whereas the maintainers received negative incentives (i.e., the subtraction of points). It was hypothesized that students who received negative incentives would exhibit higher levels of motivation than those who received positive incentives. Quantitative and qualitative methodologies were used to test the hypothesis. Although the quantitative results of this study marginally support the hypothesis, the qualitative results illuminate how the different incentives motivated students differently. The maintainers were driven by satisfaction (i.e., saw grading practices as fair; liked starting with an A), unfamiliarity (i.e., had to learn a new grading system), stress, and punishment (i.e., the threat of losing points). On the other hand, the earners were motivated by familiarity (in that they were used to the grading system used) and rewards. Implications of this study are also discussed.

Keywords: student motivation, grades, incentives.

I. Introduction.

Imagine trying to increase your students’ motivation to learn and earn good grades. You might try using uplifting narratives, engaging activities, or innovative technology. You might hold individual conferences with students to discuss their progress and help them set goals. You might even try to bribe students in the classroom with participation points or candy. But, what if there is still a lack of motivation on behalf of the students? This is not uncommon. Some students at nearly all levels of education seem unmotivated to learn and earn good grades. These students have a tendency to be apathetic and disinterested when learning about course concepts, perspectives, theories, and ideas. Unfortunately, there is no single consistent format that teachers can use to motivate their students. In an attempt to discover a way in which teachers can

¹ Teaching Associate and Ph.D. Candidate, University of Washington, tdocan@u.washington.edu, (206) 706-5469.

motivate their students, this study compares the motivating power of two different systems used to grade students. A variety of research will be reviewed to set the stage for additional investigation into this critical issue faced by students and teachers.

A. Review of Literature.

The two main bodies of literature discussed here include (1) grades as motivators and (2) grading systems. The first section, which examines grades as motivators, discusses grades as intrinsic and extrinsic motivators (Deci, 1975; Reeve, 2001; White, 1959), numerous functions of grades (Covington and Mueller, 2001; Leonard, 1968), achievement goal theory (i.e., Kaplan, Middleton, Urdan, and Midgley, 2002; Urdan, 1997; Midgely et al., 1998), and the positive and negative consequences of grades (i.e., Cleary, 1990; Condry, 1977, 1987; Deci and Ryan, 1987; Mandrell, 1997; Ryan and Connell, 1989; Skinner and Belmont, 1993). This section sets the stage for the more specific topic of grading systems, which discusses a variety of approaches used to evaluate students (Gallagher, 1998; Hendrickson and Gable, 1999; Venn, 2000) and studies that have investigated how motivation plays a specific role in how grades are presented to students (Bressette, 2002; Cullen et al., 1975).

B. Grades as Motivators.

Motivation generally is divided into two separate categories: intrinsic and extrinsic. White (1959) illuminates that intrinsic motivation involves “individuals [that] are motivated by an innate, pervasive need to seek out challenging tasks that provide feelings of general competence and mastery” (p. 18). Intrinsic motivation involves “the innate propensity to engage one’s interests” and it “emerges spontaneously from organismic psychological needs, personal curiosities, and innate strivings for growth” (Reeve, 2001, p. 119). On the other hand, extrinsic motivation involves engaging in an activity to receive rewards (Deci, 1975). An extrinsic motivator can also be defined as an “environmentally created reason to initiate or persist in an action” which arises from environmental incentives and consequences (Reeve, 2001, p. 119). Extrinsic motivation functions as a means to an end in that the means is the behavior and the end is some consequence. For example, a student may study for hours for an exam simply to receive a good grade. In this case, the student is motivated by an incentive and a consequence. The incentive is receiving a good grade. Incentives always precede behavior, and excite or inhibit the initiation of a behavior. Thus, students often are motivated by the incentive of receiving a good grade before the assignment is due. Furthermore, students also may be motivated by consequences, such as doing poorly on an exam. Consequences always follow behavior and often times increase or decrease the persistence of behavior.

In addition to increasing the likelihood of certain behaviors, rewards such as grades also function to communicate about a student’s progress and competence. At the same time, grades are so powerful that they judge a student’s overall success or failure in school (Covington and Mueller, 2001; Leonard, 1968). Furthermore, Reeve (2001) points out that most people find grades to be positive reinforcers in that they increase the probability that the behavior that produced the consequences likely will reoccur in the future. For example, if a student received an A on an assignment, she or he would be more likely to study or prepare for future assignments.

With regard to motivation, grades also have been classified as being quasi-needs, which are defined as “ephemeral, situationally induced wants that create tense energy to engage in

behavior capable of reducing the built-up tension” (Reeve, 2001, p. 149). Reeve (2001) also points out that because there is often a sense of urgency about grades, they possibly can overwhelm other needs. For example, a student may claim that he or she “has to pass the test” in order to keep a scholarship or remain on the basketball team. Furthermore, since grades function as quasi-needs, there often is considerable emotional response on the part of students.

Although grades may motivate some students to learn and study course concepts, grades also have limitations in terms of motivation. A wide variety of research demonstrates that extrinsic reinforcers, such as grades, work to decrease intrinsic motivation and interfere with the process and quality of learning (Condry, 1977, 1987; Deci and Ryan, 1987). Unfortunately, grades may distract from the learning process and focus attention on the final result—that of getting a grade. Additional research demonstrates that learners who are motivated extrinsically are less likely to experience positive emotions such as enjoyment (Harter, 1978; Ryan and Connell, 1989; Skinner and Belmont, 1993), and are more likely to use a negative emotional tone, such as displaying frustration in the classroom (Garbarino, 1975). Even “good” grades can create unmotivated students. Cleary (1990) discovered that students who were rewarded with good grades became apprehensive writers. Moreover, Benware and Deci (1984) discovered that learners motivated extrinsically are often passive information processors. This passive approach ultimately may discount what a student learns, while simultaneously favor the importance of getting good grades.

Understanding the premise of achievement goal theory (AGT) provides further illumination. AGT suggests that a student’s behaviors related to both achievement and motivation can be understood by examining the reasons they adopt while engaged in academic work (Ames, 1992; Dweck and Legget, 1988; Urdan, 1997). In essence, one of two messages is demonstrated in the classroom dependant upon whether the environment promotes performance or mastery goals (Ames, 1992; Pintrich, 2000). Mastery goals primarily focus on engaging in achievement behavior through developing competence (Kaplan, Middleton, Urdan, and Midgley, 2002). Midgley et al. (1998) point out that an environment with mastery goals conveys that learning is important, every student is valued, effort is important, and that success comes through hard work and learning. On the other hand, performance goals primarily focus on engaging in achievement behavior to demonstrate superiority over other students. Success is demonstrated through extrinsic rewards, demonstrating ability, and doing better than other students (Midgely et al., 1998).

Environments stressing performance goals can be problematic. For instance, performance based climates are associated with decreased intrinsic motivation (Xiang and Lee, 2002; Parish and Treasure, 2003; Solomon, 1996; Treasure and Roberts, 2001). Kumar, Gheen, and Kaplan (2002) argue that performance goals can potentially lead to academic struggle. Earlier research also demonstrates the consequences of rewards and learning: rewards often focus attention on learning factual information rather than conceptual information (Benware and Deci, 1984; Boggiano et al., 1993; Flink, Boggiano, and Barrett, 1990), limit one’s thinking and problem solving skills (McGraw and McCullers, 1979), and undermine creativity (Amabile, 1985; Amabile, Hennessey, and Grossman, 1986). Furthermore, research conducted by Condry and associates (1977, 1978) has demonstrated that learners’ curiosity, interest, and mastery of a subject remain more prevalent when rewards are not involved. Interestingly, Midgley (2002) points out that the promotion of mastery goals over the school years decreases. The learning process and quality of learning are at risk when grades are used as a motivating force. However, there is another side to this story.

On the positive side, extrinsic rewards do carry advantages. Reeve (2001) mentions that “rewards can make an otherwise uninteresting task seem suddenly worth pursuing” (p. 130). For instance students already may be unmotivated because of the nature of the subject. Students that are required to take certain classes often bring an unmotivated and negative attitude into the classroom. However, teachers who can facilitate engaging discussions and involve their students in activities may be able to motivate them to learn. Furthermore, if the teacher gave the students participation points for active participation, students likely would see the task ahead of them as worth pursuing. Extrinsic motivators have been used in a variety of instances to increase socially important, yet uninteresting tasks: motivating young children to do their homework (Miller and Kelley, 1994), teaching nearsighted children to wear contact lenses (Mathews et al., 1992), and getting children to participate in recycling (Austria et al., 1993; Brothers, 1994).

Because of the many disadvantages of using grades, attempts have been made to teach without them; however, students were less motivated to study when no grades were used (Mandrell, 1997). Additional research has illuminated that using grades not only enhances students’ motivation, but allows for better differentiation among students (Bressettee, 2002). In a study of 3,400 college students, Stallings and Leslie (1970) discovered that grades serve a motivational function in that when students were asked if grades provided them the motivation to complete assigned coursework, more than 67% responded positively. Grades are not the only motivator that students receive; however, they do play a role in motivation.

Although some academics argue that grades should be abandoned, scholars have asserted that arguments against using grades are empirically unfounded (Ebel, 1974). Despite the disadvantages of extrinsic motivators, grades continue to be used for fostering motivation in classrooms. Perhaps presenting the grading system for students in a different manner can compensate for the disadvantages. In this study, such a possibility was investigated.

C. Grading Systems.

Venn (2000) discusses a variety of different systems used to grade students. For example, teachers often use competency-based grading and point systems. Competency-based or criterion-based grading requires students to attain certain skills and students are graded on achieving appropriate competency in particular skill areas. Point systems allow students to earn points (in full or partial) for completing coursework such as tests, quizzes, papers, etc. Final grades are then determined by the students’ point totals at the end of the semester. Feldman et al. (1998) mention that teachers often use point systems because it allows teachers to keep a detailed account of student work and progress, allows students to keep on track, and allows teachers to unambiguously assign final grades. Competency-based grading and point systems are the most frequently used grading systems (Hendrickson and Gable, 1999).

Other, more alternative forms of grading exist as well. Venn (2000) points out that some teachers use “multiple grading”, which means that students can earn more than one grade (i.e., one grade for performance and one for effort). Some teachers grade strictly on effort and put performance aside. Others simplify the letter grade system (i.e., A, B, C, D, and F) for a pass/fail system. Criteria are generally established for what counts as passing and failing. Another grading system is referred to as contract grading. This involves the teacher and student essentially making and signing a contract that indicates the work the student will complete in a certain amount of time. Often times contracting for a higher grade requires more work. Other teachers have used portfolio grading, which is a collection of original student work. Portfolios can be

used as the only system of grading or in addition to any type of grading system. Lastly, qualitative grading moves away from letter grade assignment and relies on narrative feedback from the teacher. The teacher might write a letter about student's performance, participation, improvement, attitude, etc. Qualitative grading can be used as a sole means of grading or with other grading systems.

Although a variety of grading systems are used, teachers from nearly all angles of education in the U.S. still experience difficulty with issues that arise around grading. A variety of projects have attempted to discover how to use and present grades in the classroom. Bressette (2002) discovered that the use of a plus/minus grading system serves as an "excellent motivator for students to continue strong efforts on all assignments and examinations right up to the last day of classes" and that "if the benefit of receiving a high grade (i.e., a B+ compared to a B) is not a strong motivator, the fear of moving to a lower grade (i.e., a B- compared to a B) from lack of effort might be an even stronger motivator" (p. 38).

Bressette's (2002) research investigated the motivating factors of using a plus/minus grading system; however, prior research has examined the effects of positive and negative incentives with regard to grading. Cullen et al. (1975) used a sample of 233 students from 14 high schools to discover how positive and negative incentives motivate students to complete assignments. The students were "either offered points (ranging from 2 to 12) on their final grade of the term for completing an assignment or threatened with loss of points (ranging from 1 to 7) for not completing an assignment" (Cullen et al., 1975, p. 277). These researchers used different point values for the classes: 0, +2, +3, +5, +6, +8, +10, +12, -1, -2, -3, -4, -5, and -7. Their results attest that "grades used as either positive or negative incentive secures greater assignment completion than when no incentive is offered" (p. 278). In essence, when either type of incentive was offered, the students were more motivated to complete their work. They also discovered that grades used as a negative incentive are more powerful than when they are used as a positive incentive. For example, having points deducted (e.g. -3) was more powerful than having points added (e.g. +3). Furthermore, the greater the negative incentive, the more likely students were to complete the assignment. For example, a student would be more motivated to complete the assignment if there was a threat of losing 5 points over losing 2 points. Interestingly, Cullen et al. (1975) mention that reasons as to why the correlation between incentive (positive or negative) and the completion of assignment was higher with negative incentives than with positive incentives has yet to be determined.

Cullen et al.'s (1975) study opens the door for further research. Although their study provides insightful results, it only used the positive and negative incentives for the final assignment of the class—and not throughout the entire semester. A semester long analysis may provide more fruitful results. Questions also remain as to why students were more motivated when a negative stimuli was presented. Asking the students themselves about their motivation and what was affecting it may provide further answers.

The current study aims to better understand the relationship between student motivation and how grades are presented (i.e., by using positive and negative incentives). This study aims to investigate two grading systems and the motivation they produce. One system will provide students with negative incentives, while the other system will provide students with positive incentives. Because grades clearly do not motivate all students, this study seeks to discover if these problems can be alleviated by exploring two different ways teachers can present grades. Based on Cullen et al.'s (1975) study, one hypothesis guides this study: Students who receive

negative incentives will exhibit higher levels of motivation than those who receive positive incentives.

II. Methods.

A. Procedures.

Two grading systems were designed for this study. These two systems were developed with the intention that one system would surface as more conducive to student motivation. Student motivation levels (dependent variable) were measured (at three points in the semester; see “Variables and Scales” section) in response to the two different types of grading systems. Since competency-based grading and point systems are most prevalent (Hendrickson and Gable, 1999), the current study is modeled around these systems.

There were two grading systems used for this study. The students who earned their grades (“earners”) received positive incentives and those who maintained their grades (“maintainers”) received negative incentives. The earners started the semester with 0 points and added points with each graded assignment, whereas the “maintainers” were given the maximum number of points available for the course at the beginning of the semester and then subtracted points from this overall total as they lost points on a graded assignment. The earners received positive incentives (i.e., the addition of points), whereas the maintainers received negative incentives (i.e., the subtraction of points).

Students in each class were given a prepared sheet of paper to record their grade for each assignment. An example of the grade sheet for the earners is provided in Table 1. An example of the grade sheet for the maintainers is provided in Table 2.

Students were asked to complete a Student Motivation questionnaire (see Appendix I) during the first week of the semester and at end of the semester. Students were also asked to complete an open ended questionnaire regarding their motivation. Participation was voluntary and extra credit was provided.

Table 1. Grade sheet for earners.

Assignments	Assignment Value	Your Score
Introductory Speech	3	+2.50
Group Impromptu	5	+4.00
Narrative Speech	10	+8.00
Total Points	100	14.50/100

Table 2. Grade sheet for maintainers.

Assignments	Assignment Value	Points Lost	Running Total
			100 points
Introductory Speech	3	-.5	99.5
Group Impromptu	5	-1.0	98.5
Narrative Speech	10	-2.0	96.5
Total Points	100		/100

B. Participants.

Six public speaking courses at a large southwestern university were selected to participate in this study. Three of the classes were designated as the “earners” and the other three were designated as the “maintainers.” A total of 101 students participated in the study; earners ($n = 49$) and maintainers ($n = 52$). Forty-eight percent ($n = 48$) of the participants were male and 52% ($n = 52$) were female. The average age of participants was 20.4 with the youngest participant being 18 and the oldest being 41. Thirty-nine percent ($n = 39$) of participants were freshman, 36% ($n = 36$) were sophomores, 16% ($n = 16$) were juniors, 9% ($n = 9$) were seniors, and 1% ($n = 1$) were graduate students.

The six public speaking classes were taught by three instructors. Each instructor taught one class of earners and one class of maintainers. Specific guidelines were set before the instructors began teaching their courses to ensure that each course was taught similarly and that students were treated similarly as well. For instance, assignments (i.e., speeches, papers, quizzes), in-class activities, and classroom policies (i.e., attendance, rules for late work, plagiarism, etc.) were designed in a similar manner (some of which was already a department requirement). Each instructor was given a journal and asked to record any deviations from the guidelines set out before the semester. At the end of the study, it was determined that there were no significant deviations.

C. Variables and Scales.

The hypothesis examined in this study looked at the grading system (independent variable) and students' levels of motivation (dependent variable). Levels of motivation were measured by administering a questionnaire consisting of 16 questions that derived from the Student Motivation Scale (Beatty and Payne, 1985; Christophel, 1990). This scale was distributed at the beginning and end of the semester. This scale was used because it has acceptable reliability and validity (Beatty and Payne, 1985; Beatty et al. 1986; Christophel, 1990; Richmond, 1990). For instance, Christophel's (1990) reliability coefficient ranged from .95 to .96 and Richmond (1990) reported alpha co-efficient of .94. The Student Motivation Scale has not only been used over the years, but expanded and improved as well. The original version, which was created by Beatty et al. (1980), has been expanded as literature measuring motivational states in students has increased. For instance, Beatty, Forst, and Stewart (1986) added bipolar adjectives to the scale, as did Richmond (1990) and Christophel (1990). Other scholars have used the scale more recently as well (i.e., Corrigan, 2004; Guzley, Avanzino, and

Bor, 2001). Previous research shows that each of the questions included in the scale are indicators of motivation (Beatty and Payne, 1985; Christophel, 1990).

An open ended questionnaire was also utilized to gain a greater understanding of the variables under examination (see Appendix II). The open-ended questionnaire was intended to get a first hand account of how the students felt about and reacted to their particular grading system. The self-reporting nature of the survey was specifically used to allow the students to express their own thoughts, feelings, and impressions about the grading systems, instead of forcing the students to rate how much of a particular feeling they may have experienced. The open ended questionnaire was distributed to the students half way through the semester.

After the open-ended scale was originally constructed, a pilot test and numerous brainstorming sessions were conducted with volunteers to revise the questionnaire, clarify questions, and suggest additional questions. Volunteers were also asked to complete the questionnaire and make suggestions regarding the clarity of questions, grammar, and wording. Their ideas and suggestions for revision were incorporated into the study and helped make the questionnaire easier to understand. After the students completed the surveys, their answers were entered into a computer and common themes for each question were identified and clustered into categories.

III. Results.

A. Quantitative Findings.

The hypothesis stated that students who receive negative incentives would exhibit higher levels of motivation than those who receive positive incentives. A one-tailed independent samples t-test showed that there was a slight difference in levels of motivation in earners ($M = 2.91$, $SD = 1.393$) and maintainers ($M = 2.49$, $SD = 1.197$)^{2 3}. Furthermore, the t-test [$t(99) = 1.620$, $p = 0.054$] shows a moderate significance. The hypothesis is subsequently marginally supported by the quantitative data.

B. Qualitative Findings.

Student motivation with regard to receiving positive or negative incentives was discovered by numerous themes that emerged from the data. These themes include (1) student satisfaction with regard to the particular grading system; (2) student dissatisfaction; (3) grades as stressors; (4) and motivation when compared to other classes. The majority of the themes demonstrate differences between the earners and maintainers; however, there are some instances in which similarities between the two groups emerge.

Student satisfaction with grading systems. One main theme that emerged from the data concerns student satisfaction with the grading systems. Deci et al. (2001) point out that satisfaction is positively related to motivation; thus, the themes of “student satisfaction with grading systems” and “student dissatisfaction with grading systems” will provide an understanding of student motivation levels. In essence, when students are more satisfied (i.e., they feel content, are in approval, or have a liking towards something in the course), they will be inclined to be more motivated.

² Note that lower Ms indicate higher levels of motivation.

³ Note that averages were used.

Mostly similarities between the earners and maintainers emerged within this theme. The students were asked how they felt about having to earn or maintain points in their class. Over half of the respondents in each grading system responded that they thought the system was “good,” “okay,” “fair,” and “they liked it.” One question that addresses satisfaction asked the students about the strengths of the grading procedures used in their particular class. First, nearly 40% (N = 41) of the students responded that knowing their grade and progress in the class were strengths. These students likely reported this because they were provided a grade sheet to keep track of their points. However, more maintainers (N = 25) than earners (N = 16) remarked that knowing where they stood in the class was a strength. This may be the case because the maintainers self-reported that their grading system was unfamiliar and the earners consistently remarked that their grading system was similar to their other classes. These feelings of unfamiliarity with the grading system may have prompted the maintainers to direct additional attention towards keeping track of their own grades and to report that they understood what their grade was in the class on a consistent basis. Furthermore, this unfamiliar system of having points deducted, or receiving a negative incentive likely caused the stress reported by the maintainers, which ultimately may have motivated the students to pay more attention to their grade and track their progress more consistently.

In addition to consistently knowing their grade in the class, another strength emerged for the maintainers. One third of the maintainers self-reported that a strength of the grading procedures used in the class was that everyone started with an A, or the maximum amount of points. Some of the students’ comments included “The fact that you start with an A in the class gives you a more positive outlook,” “It’s more positive than having to work up to an A,” “It’s better than earning them,” and “It lets you focus more on the work than on what grade you currently have or need to get.” Factors such as satisfaction with the overall grading system, starting with an A, and committing extra time to understanding an unfamiliar system likely contributed to the motivational levels of the students under this system. However, levels of student dissatisfaction likely played a role in motivational levels as well.

Student dissatisfaction with grading systems. In addition to student satisfaction with grading systems, signs of dissatisfaction emerged as well. When a student is dissatisfied (i.e., they feel discontent, are in disapproval, or have a disliking towards something in the course) they will be inclined to be less motivated. Both similarities and differences between the earners and maintainers emerged. Students were asked to list the weaknesses of the grading procedures used in their class. Some students in both the maintainer (N = 15) and earner (N = 13) groups reported that there were no weaknesses. However, other respondents did list areas of dissatisfaction. Over half of the earners (N = 29) and 30% of the maintainers (N = 16) mentioned that they did not like specific aspects of the course. For instance, students wrote that there was “too much busy work,” that the pop quizzes “don’t help with learning,” and that points should not be deducted for being absent.

Each student was also asked to describe their teacher’s grading procedures. Fourteen percent (N = 8) of the earners and 6% (N = 3) of the maintainers described the grading practices as unfair and harsh. Although this is not a huge difference in responding with dissatisfaction, these findings may assist in discovering the overall motivating factors with regards to these grading systems in the end. Additional dissatisfaction illuminates that the particular grading systems were difficult to understand for certain students. Only one earner reported that the grading system takes time to get used to. However, over 10% of the maintainers reported that their grading system was hard to understand—likely because this system is rarely used. For

instance, some of the students commented that there is a “lack of familiarity” and it is “a little hard to understand.” Again, the maintaining system emerges as more difficult to understand, which will be further discussed later. Furthermore, nearly one-fourth of the maintainers were dissatisfied that there was no way to get lost points back, whereas 5% of the earners mentioned that it was difficult to “catch up” or “earn more points.” Since Deci et al. (2001) point out that satisfaction is positively related to motivation, it can be surmised that factors such as satisfaction with the overall grading system (i.e., starting with an A, familiarity with a particular system, etc.) and dissatisfaction (i.e., unfamiliarity with a system, etc.) contribute to the motivational levels of the students.

Grades as stressors. Throughout this analysis, grades seem to cause stress for the majority of students in one way or another. Stress and motivation are related (Lazarus, 1966). Stress is a motivationally related response to particular environments or conditions. Interestingly, when one experiences stress (e.x., one’s possible reaction to a new system of rules), the nervous system releases epinephrine (or adrenaline), which leads to increased heart rate, blood pressure, and respiration rate. Furthermore, when one experiences stress, cortisol (a hormone) is released, and when cortisol levels are high, problem solving is significantly impaired (Kirschbalm et al., 1996). Thus, understanding grades as stressors will help further explain students’ motivation levels.

Although similar results between the two groups were found when asking about grades, differences emerged when asking about particular grading systems. Nearly half of the sample (N = 45) responded that some type of stressful emotion runs through their body. Some of the answers included “tension,” “fear,” “paranoia,” “anxiety,” “discomfort,” “anger,” and “anxious.” Very few students remarked that positive feelings result when thinking about grades.

When the students were asked how they felt about having to earn or maintain their grade, the results demonstrate that the maintainers reported more stress and negativity with regard to their specific grading system. Numerous respondents stated that maintaining points made them feel “nervous” and “pressured.” Furthermore, additional comments illuminate why this stress may arise: “I can’t do anything to make it higher,” “There’s no way you can go but down,” and “You can easily give up when you feel that you lose points.” Although the earners and maintainers are being graded in the same ways, the maintainers pick up on the negative slant of their grading system and eventually feel more stress. It appears that stress may contribute to a student’s level of motivation.

Motivation when compared to other classes. Students were also asked to compare and contrast their motivation to earn/maintain their grade in the class in which they received the positive or negative incentives with their other classes. Slight differences between the two groups were discovered. Slightly more maintainers (N = 19) than earners (N = 16) reported that they were more motivated to maintain their grade in the class under investigation than in their other classes. The same amount of earners and maintainers reported that their motivation was the same as other classes. Lastly, more earners than maintainers (20% versus 13%) reported that they were less motivated in the class under investigation in the study compared to their other classes (though this could relate to course content). Overall, more maintainers reported being more motivated when compared to other classes.

Numerous themes have emerged as important components of this analysis: satisfaction with the overall grading system, dissatisfaction, levels of stress, and motivation compared to other classes. These findings point out that students were motivated in different ways.

IV. Discussion.

A. Interpretation.

In order to improve some of the motivational and grading system problems presented to teachers, an attempt was made to better understand student motivation by investigating the ways teachers assign grades. Although scholars (Cullen et al., 1975) have examined the impact of positive and negative incentives on students for single assignments, this project consisted of a semester-long study that examined the impact of positive and negative incentives on student motivation. Furthermore, this study utilized both quantitative and qualitative methods to gain a rich understanding of this issue.

The hypothesis in the current study stated that upon completion of the semester, students who received negative incentives would exhibit higher levels of motivation than those who received positive incentives. The results of the quantitative questionnaire were marginally supported in that those who received negative incentives were slightly more motivated than those who received positive incentives. The difference between the two groups corresponds with Cullen et al.'s (1975) findings in that those who received negative incentives were overall, slightly more motivated. However, the results of the qualitative questionnaire demonstrate more descriptive results in that students were motivated in different ways. A discussion of the similarities and differences between the two groups will further illuminate this finding.

Similarities between earners and maintainers. There are numerous points of similarity between the two groups: (1) grades in general do create stress and (2) knowing one's progress and grade in a class is a strength. First, the fact that grades caused stress for students in this study is notable. Although motivation and stress are related (Lazarus, 1966), a more significant issue exists: stress is likely one of the major confounding variables that contributes to the interference with learning. For instance, Kirschbalm et al. (1996) point out that stress can significantly impair problem solving. This is problematic when considering that problem solving is one of the fundamental skills that students at nearly all levels of education need to develop.

A second area of commonality between the two groups of students demonstrates that knowing their progress and grades in the class gave them a sense of satisfaction and motivation. As mentioned earlier, the students were given grade sheets at the beginning of the semester and required to track their grades. The students in both groups were more satisfied and motivated because of this requirement. On a pragmatic level, this study demonstrates that teachers can potentially increase their students' levels of satisfaction with the course by handing out pre-prepared grade sheets at the beginning of the semester. However, this can potentially create more stress on a student, because of the additional focus on grades. Although there were two areas of commonality between the earners and maintainers with regards to student motivation, a variety of differences between the two groups emerged.

Differences between earners and maintainers. Four points of discussion are noteworthy: (1) satisfaction with the system; (2) unfamiliarity as motivation; (3) negativity as motivation; and (4) attention focused on grades.

The maintainers appeared to be more satisfied with their system. Deci et al. (2001) point out that satisfaction is positively related to motivation; thus a discussion of students' satisfaction is noteworthy. One reason why the maintainers saw their system as more satisfying is because the teacher's grading practices were more often described as "good" and "fair." Interestingly, the students were graded the same way in every class; however, the grades were presented in

different ways. In addition, any stress involved in maintaining grades seems to coincide with strong opinions that the instructors' grading procedures are still good and fair. The maintainers were further satisfied and motivated by starting out with an A. A large group of the maintainers reported that a strength of the grading procedures used in the class was that everyone started with an A, or the maximum amount of points. In addition, maintainers reported that they were more motivated in this class than in their other classes.

The maintainers consistently self-reported that their grading system was unfamiliar. The feelings of unfamiliarity with the grading system may have prompted the maintainers to direct additional attention towards keeping track of their own grades, and resultantly report that they understood what their grade was in the class on a consistent basis. Furthermore, this unfamiliar system of having points deducted, or receiving a negative incentive likely caused the stress reported by the maintainers, which ultimately may have motivated the students to pay more attention to their grade and track their progress more consistently than the earners. In addition, these results correspond with Bressette's (2002) assertion that "if the benefit of receiving a high grade (i.e., a B+ compared to a B) is not a strong motivator, the fear of moving to a lower grade (i.e., a B- compared to a B) from lack of effort might be an even stronger motivator" (p. 38). Thus, the maintainers could have been motivated by the fact that their grade could further drop. On the other hand, the earners were to some degree motivated by familiarity in that their grading system was much more common, yet still worked as a motivating factor in the course.

Overall, the earners and maintainers expressed some kind of dissatisfaction with their grading systems. Nearly a quarter of the maintainers were dissatisfied that there was no way to get lost points back, whereas only 5% of the earners mentioned that it is difficult to "catch up" or "earn more points." This dissatisfaction may work in a manner that motivates the maintainers. Because the negative incentive bothers the student and makes her or him feel behind, she or he is more likely to reverse this pattern and attempt to maintain points in the future.

Another point of divergence rests in the fact that the maintaining system may also have focused students' attention away from areas of the course that were considered unpleasant. Over half of the earners (N = 29) and only 30% of the maintainers (N = 16) mentioned that they did not like specific aspects of the course. Interestingly, the maintainers' attention may have been more focused on the grading system at hand, and not at finding problems with other areas of the course. This facet of the study demonstrates how the maintaining system works to focus students' attention on grades and points, and away from other areas, such as "having too much work" or believing that "too many quizzes are given." This system also potentially kept the maintainers from focusing on learning. Grades and a new system of being graded were emphasized so much, that the actual task at hand—that of learning—may have been disrupted. It is likely that the maintaining system de-emphasized mastery goals or engaging in achievement behavior through developing competence (Kaplan, Middleton, Urdan, and Midgley, 2002). Grades (and subsequently performance goals) were stressed over goals of a mastery goal-oriented environment: learning is important, every student is valued, effort is important, and that success comes through hard work and learning. This is problematic in that performance goals can potentially lead to academic struggle (Kumar, Gheen, and Kaplan, 2002). However, grades were also stressed for the earners, but not to the same degree (i.e., the earners better understood their grading system, therefore there was less time spent explaining how the grading system worked).

Although the quantitative results of this study are in slight alignment with Cullen et al.'s (1975) study, the qualitative results illuminate how the two types of incentives motivated the

students differently. Both earners and maintainers were motivated by stress and understanding their progress in the class. However, the maintainers were motivated by satisfaction (i.e., saw grading practices as fair; liked starting with an A), unfamiliarity (i.e., had to learn a new grading system). They were also motivated by the fear of punishment (i.e., the threat of losing points) and stress. On the other hand, the earners were to some degree driven by familiarity in that they were used to the grading system used.

B. Implications.

Eiszler (1983) asserted that “the evidence that grades derive their meaning, in part, from the context in which they are assigned, implies, however, that all grading systems are not equally useful in this regard” (p. 19). This assertion was clearly illuminated in this study. The maintainers in this study had to deal with a negative style of being judged, and were often driven by this negativity. This negativity may ultimately change the way these students handle and perceive the subject matter under investigation. In addition, most of the students were extrinsically motivated by grades in this study.

If the maintaining system is widely adapted, students likely will adjust to the system over time, as the earners have; however, after spending time adapted to a grading system, students’ motivation in wanting and working for good grades may decrease as it did for the earners. Thus, a longitudinal study examining how grading systems are adapted to by students and how this adaptation may alter levels of motivation is recommended. However, this study did allow for a rich examination of how grading systems motivate students in particular ways.

Although this study discovered that the two grading systems motivated students in different ways, the way in which they were motivated was problematic in that mastery goals were left on the backburner and extrinsic rewards were stressed. A wide variety of research demonstrates that extrinsic rewards, such as grades, work to decrease intrinsic motivation and interfere with the process and quality of learning (Condry, 1977, 1987; Deci and Ryan, 1987; Kumar, Gheen, and Kaplan, 2002; Parish and Treasure, 2003; Solomon, 1996; Treasure and Roberts, 2001; Xiang and Lee, 2002). Unfortunately, grades may distract from the learning process and focus attention on the final result—that of getting a grade. Additional research demonstrates that learners who are motivated extrinsically are less likely to experience positive emotions, such as enjoyment (Harter, 1978; Ryan and Connell, 1989; Skinner and Belmont, 1993). Thus, new ways to motivate students should be sought.

V. Conclusions.

A. Limitations.

While this study did provide noteworthy findings for the study of student motivation and grading, there are limitations. One limitation of the current study is the sample size. A larger sample size may have yielded statistically more significant results for the hypothesis that those who received negative incentives would be more motivated than those who received positive incentives. A second limitation of the study had to do with the fact that the students were examined in a controlled environment. They knew that they were in a research project examining grades. There may have been too much emphasis placed on grades throughout the semester (i.e., students had to fill out surveys, students asked their teacher about the project, students were

required to keep track of their grade with a grade sheet, etc.). Finally, this study only used public speaking classes. Including courses in different subject matters could have had an impact on how students view the course itself, therefore impacting how motivated the students are and how they perceive their teachers. However, a strength to using students in public speaking classes is that it is a required course for all students, which alleviates the concern that students self-selected this course and consequently would begin with a high degree of motivation.

B. Future Research.

Future research should examine alternative ways in which students can be motivated. As mentioned, future studies should use a larger sample, as well as different types of courses. With a larger sample, it would be appropriate to retest the hypothesis to examine if the findings are significant. Including students from different types of courses, such as math, science, English, and electives might have an impact on the level of motivation for the student. The level of motivation the student feels at the beginning of the semester towards a particular course may impact that student's level of motivation throughout the course.

Scholars should also consider comparing how different grading systems (i.e., contract grading, point systems, qualitative grading) motivate students both intrinsically and extrinsically. Additionally, since it is evident that grades work as extrinsic motivators, research needs to look more at how intrinsic motivators can enhance students' learning experiences, by not only inspiring and motivating the students, but also by increasing students' interest in the subject. Some questions that still need to be addressed might include: how can both mastery and performance goals be optimally implemented while using both intrinsic and extrinsic rewards? How can new systems of grading be invented to successfully motivate students? How is teacher-student interaction influenced by the ways teachers assign grades? By addressing some of the suggestions for future research, we can hopefully discover ways to enhance the learning experience for students in the classroom.

Author Notes

I would like to thank Karen Foss for her helpful feedback on this paper.

Appendix I

Questionnaire

#_____

Demographics:

Sex: Male Female Age: _____ Major: _____

Year in School: Freshman Sophomore Junior Senior Other_____

The following questionnaire is designed to assess your motivation in the course. Please circle the number toward either word which best represents your feelings.

Example: If you feel very empowered, you would circle the number 1 since it is the number closest to the word that best represents your feelings.

	Empowered	1	2	3	4	5	6	7	Unempowered
1. Motivated		1	2	3	4	5	6	7	Unmotivated
2. Interested		1	2	3	4	5	6	7	Uninterested
3. Involved		1	2	3	4	5	6	7	Uninvolved
4. Not stimulated		1	2	3	4	5	6	7	Stimulated
5. Want to study		1	2	3	4	5	6	7	Don't want to study
6. Inspired		1	2	3	4	5	6	7	Uninspired
7. Unchallenged		1	2	3	4	5	6	7	Challenged
8. Uninvigorated		1	2	3	4	5	6	7	Invigorated
9. Unenthused		1	2	3	4	5	6	7	Enthused
10. Excited		1	2	3	4	5	6	7	Not excited
11. Aroused		1	2	3	4	5	6	7	Not aroused
12. Not fascinated		1	2	3	4	5	6	7	Fascinated
13. Dreading it		1	2	3	4	5	6	7	Looking forward to it
14. Important		1	2	3	4	5	6	7	Unimportant
15. Useful		1	2	3	4	5	6	7	Useless
16. Helpful		1	2	3	4	5	6	7	Harmful

Appendix II

Questionnaire

#_____

Demographics:

Sex: Male Female Age: _____ Major: _____

Year in School: Freshman Sophomore Junior Senior Other _____

Directions: Please answer the following questions being as descriptive as possible.

Did having to earn/maintain your grade make a difference in your motivation in this course? If so, in what ways? If not, explain why.

Compare and contrast your motivation to earn/maintain your grade in this class with other classes you are taking.

When someone says the word *grade* or *grading*, what words, phrases or feelings, immediately pop into your mind?

How do you feel about having to earn/maintain points in this class?

What are the strengths of the grading procedures used in this class?

What are the weaknesses of the grading procedures used in this class?

References

- Amabile, T. M. (1985). Motivation and creativity: Effect of motivational orientation on creative writers. *Journal of Personality and Social Psychology*, 48, 393-399.
- Amabile, T. M., Hennessey, B. A., and Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology*, 50, 14-23.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261-271.
- Austira, J., Hatfield, D. B., Grindle, A. C., and Baily, J. S. (1993). Increasing recycling in office environments: The effects of specific, informative cues. *Journal of Applied Behavior Analysis*, 26, 247-253.
- Beatty, M. J., Behnke, R. R., and Froelich, D. L. (1980). Effects of achievements incentive and presentation rate on listening comprehension. *Quarterly Journal of Speech*, 66, 193-200.
- Beatty, M. J., Forst, E. C., and Stewart, R. A. (1986). Communication apprehension and motivation as predictors of public speaking duration. *Communication Education*, 35, 143-146.
- Beatty, M. J., and Payne, S. K. (1985). Is construct differentiation loquacity?: A motivational perspective. *Human Communication Research*, 11, 605-612.
- Benware, C., and Deci, E. L. (1984). The quality of learning with an active versus passive motivational set. *American Educational Research Journal*, 21, 755-765.
- Bressette, A. (2002). Arguments for plus/minus grading: A case study. *Educational Research Quarterly*, 25, 29-41.
- Brothers, K. J. (1994). Office paper recycling: A function of container proximity. *Journal of Applied Behavior Analysis*, 27, 153-160.
- Christophel, D. M. (1990). The relationship among teacher immediacy behaviors. Student motivation, and learning. *Communication Education*, 39, 323-340.
- Cleary, L. M. (1990). The fragile inclination to write: Praise and criticism in the classroom. *English Journal*, 79, 22-28.
- Condry, J. (1977). Enemies of exploration: Self-invited versus other-initiated learning. *Journal of Personality and Social Psychology*, 35, 459-477.
- Condry, J. (1987). Enhancing motivation: A social development perspective. *Advances in Motivation and Achievement: Enhancing Motivation*, 5, 23-49.

Covington, M. V., and Leonard, K. J. (2001). Intrinsic versus extrinsic motivation: An approach/avoidance reformulation. *Educational Psychology Review*, 13, 157-176.

Corrigan, M. W. (2004). *An empirical measurement of interpersonal community engagement: Implications to youth communication behaviors and the instructional setting*. Unpublished dissertation, West Virginia University.

Cullen, F. T., Cullen, J. B., Hayhow, V. L., and Plouffe, J. T. (1975). The effects of the use of grades as an incentive. *The Journal of Educational Research*, 68, 27-279.

Deci, E. L (1975). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

Deci, E. L., and Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 39-80). New York: Academic Press.

Deci, E. L., Ryan, R. M., Gagne, M., Leone, D. R., Usunov, J., Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former eastern block country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin*, 27, 930-942.

Dweck, C. and Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.

Ebel, R. L. (1974). Shall we get rid of grades? *Measurement in Education*, 5, 1-6.

Eiszler, C. (1983). The meaning of college grades in three grading systems. *Educational Research Quarterly*, 8, 12-19.

Feldman, A., Alibrandi, M., and Kropf, A. (1998). Grading with points: The determination of report card grades by high school science teachers. *School Science and Mathematics*, 98, 140-148.

Flink, C., Boggiano, A. K., and Barrett, M. (1990). Controlling teaching strategies: Undermining children's self-determination and performance. *Journal of Personality and Social Psychology*, 59, 916-924.

Gallagher, J. D. (1998). *Classroom assessment for teachers*. Columbus, OH: Merrill.

Garbarino, J. (1975). The impact of anticipated reward upon cross-aged tutoring. *Journal of Personality and Social Psychology*, 32, 421-428.

Guzley, R. M., Avanzino, S., and Bor, A. (2001). Simulated computer-mediated/video-interactive distance learning: A test of motivation, interaction satisfaction, delivery, learning, and perceived effectiveness. *Journal of Computer Mediated Communication*, 6 (3).

Harter, S. (1978). Pleasure derived from optimal challenge and the effects of extrinsic rewards of children's difficulty level choices. *Child Development*, 49, 788-799.

Hendrickson, J. M., and Gable, R. A. (1999). Can everyone make the grade? Some thoughts on student grading and contemporary classrooms. *High School Journal*, 82, 248-254.

Hufton, N. R., Elliot, J.G., and Illushin, L. (2003). Teachers' beliefs about student motivation: Similarities and differences across cultures. *Comparative Education*, 39, 267-289.

Kaplan, A., Middleton, M., Urdan, T. and Midgley, C. (2002). Achievement goals and goal structures. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 21-53). Mahwah, NJ: Erlbaum.

Kirschbaum, C., Wolf, O. T., May, M., Wippich, W., and Hellhammer, D. H. (1996). Stress and treatment-induced elevations of control levels associated with impaired declarative memory in healthy adults. *Life Sciences*, 58, 1475-1483.

Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin.

Kumar, R., Gheen, M. H., and Kaplan, A. (2002). Goal structures in the learning environment and students' disaffection from learning and schooling. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 143-174). Mahwah, NJ: Erlbaum.

Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.

Leonard, G. B. (1968). *Education and ecstasy*. Delacorte, New York.

Mathews, J. R., Hodson, G. D., Crist, W. B., and LaRoche, G. R. (1992). Teaching young children to use contact lenses. *Journal of Applied Behavior Analysis*, 25, 229-235.

McGraw, K. O., and McCullers, J. C. (1979). Evidence of detrimental effects of extrinsic incentives on breaking a mental set. *Journal of Experimental Social Psychology*, 15, 285-294.

Midgley, C. (2002). *Goals, goal structures, and patterns of adaptive learning*. Mahwah, NJ: Erlbaum.

Midgley, C. et al. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology*, 23, 113-131.

Miller, D. L., and Kelley, M. L. (1994). The use of goal setting and contingency contracting for improving children's homework performance. *Journal of Applied Behavior Analysis*, 27, 73-84.

Parish, L. E., and Trasure, D. C. (2003). Physical activity in physical education: Influence of the motivational climate and perceived ability. *Research Quarterly for Exercise and Sport*, 74, 173-182.

Pintrich, P. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544-555.

Reeve, J. (2001). Understanding motivation and human emotion (3rd ed.). Orlando, FL: Harcourt.

Richmond, M. J. (1990). Communication in the classroom: Power and motivation. *Communication Education*, 39, 181-195.

Ryan, R. M., and Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749-761.

Skinner, E. A., and Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85, 571-581.

Solmon, M. A. (1996). Impact of motivational climate on students' behaviors and perceptions of a physical education setting. *Journal of Educational Psychology*, 88, 731-738.

Stallings, W. M., and Leslie, E. K. (1970). Student attitudes toward grades and grading. *Improving College and University Teaching*, 18, 66-68.

Treasure, D. C., and Roberts, G. C. (2001). Students' perceptions of the motivational climate, achievement beliefs, and satisfaction in physical education. *Research Quarterly for Exercise and Sport*, 72, 165-175.

Urdan, T. (1997). Achievement goal theory: Past results, future directions. In M. Maehr and P. R. Pintrich (eds.), *Advances in motivation and achievement* (Vol. 10, pp. 99-141). Greenwich, CT: JAI Press.

Venn, J. (2000). Assessing students with special needs (2nd ed.). Columbus: Merrill Education.

Xiang, P., and Lee, A. (2002). Achievement goals, perceived motivational climate, and students' self-reported mastery behaviors. *Research Quarterly for Exercise and Sport*, 73, 58-65.

Negotiating Roles and Meaning While Learning Mathematics in Interactive Technology-Rich Environments

Jack Bookman¹ and David Malone²

Abstract: The authors examined how undergraduate students negotiated roles and developed a shared understanding of mathematics while working together on computer-based modules. The subjects were videotaped while working on these modules and their computer output was simultaneously collected on a separate videotape. Examination of the data tentatively suggests ways pairs of students approach academic work and ways they interact as they process and organize knowledge. This study builds on an exploratory study the authors conducted that generated a set of research questions addressing the nature of learning in interactive technological environments.

Key Words: Mathematics education, Computer algebra systems, Learning styles, Collaborative learning, Interactive learning.

I. Introduction.

In the early 1900's, Edison predicted that motion pictures would make books obsolete and in the 1950's, many mistakenly believed that educational television would revolutionize schooling (Reiser, 2001). Personal computers and the Internet, like TV in the 1950's, have exploded on the educational scene in the last decade with hyperbolic promises and predictions about how they will affect the way we teach and learn. Similarly, collaborative approaches to learning have been endorsed as a means of ensuring deeper and more authentic learning. As teachers many of us are eager to embrace these educational innovations that are touted as holding great promise of energizing our classrooms. Often, however, teachers adopt curricular and instructional changes without carefully evaluating the efficacy and consequences of these new approaches.

This study builds on an exploratory study the authors conducted that generated a set of research questions addressing the nature of learning in interactive technological environments (Bookman and Malone, 2003). In that study, the authors formulated three categories of research questions: (1) what is the role of the university instructor in interactive technology-rich environments? (2) What types of behavior and thinking processes are university students engaged in as they work together in front of the computer? and (3) what opportunities and obstacles are raised by the technology itself? Our objective in the study reported in this paper is to analyze the interactions and social relationships between students as they worked together on computer-based math modules. *The primary question of interest is: What patterns of behavior and social relationships emerge when students learn mathematics collaboratively in technology rich, socially interactive learning situations, and what impact might these patterns of behavior have on the opportunity to learn in these settings?* In the case of these two particular innovations

¹ Mathematics Department of Duke University bookman@math.duke.edu.

² Program in Education at Duke University dmalone@duke.edu.

mentioned above (collaboration and computer enhanced learning) there has been, over the past decade, a significant amount of research. For example, Barron (2003) and others (Rogoff, Turkkanis, and Bartlett, 2001; Roschelle, 1992) have investigated the nature of collaborative learning and conceptual change (see:

<http://ctl.sri.com/publications/downloads/ConvergentConceptual.pdf>

and http://www.leaonline.com/doi/abs/10.1207/S15327809JLS1203_1?journalCode=jls

Likewise, Inkpen (1997) and others (Shechtman, Roschelle, Haertel, Knudsen, and Tatar 2005), have examined the shared use of the computer as a pedagogical tool (see:

http://www.futureofchildren.org/information2826/information_show.htm?doc_id=69809

and <http://www.cs.sfu.ca/people/Faculty/inkpen/publications.html>.

However, much of this work on the collaborative use of technology as a pedagogical approach has focused on younger K-8 students. Less empirical research has been published concerning the way undergraduate students learn college-level mathematics interactively using computers and Internet based problem solving environments. Thus, the focus of our study was to examine how *undergraduate* students go about negotiating roles and developing a shared understanding of mathematics while working together on computers solving Internet based problem modules.

The study reported in this paper utilized a grounded theory approach (Glaser and Strauss, 1967) to generate several conceptual categories based on observations of students' work. Glaser and Strauss (1967) describe *grounded theory* as "the discovery of theory from data systematically obtained from social research." Grounded theory is more of an inductive approach as contrasted with "theory generated by logical deduction from *a priori* assumptions." The first step in this process is to examine the data with the purpose of establishing categories and/or constructs while minimizing the effect of preconceptions, preexisting theories or prejudices. For these reasons, this methodology is particularly appropriate for the kind of exploratory research described in this paper.

Our interpretation of the "grounded" approach to theory development is that it occurs in roughly four stages:

1. Start with the data and observations and see what research questions and categories emerge.
2. Refine and reformulate specific categories with the goal of generating hypotheses.
3. Design and implement new and more focused data collections whose purpose is to eliminate and refine hypotheses and move to an emerging theory. By emerging theory we mean the accumulation of evidence that supports certain hypotheses and the organization of those hypotheses into a coherent framework that may explain the phenomenon being studied. Formulate theories that can be empirically investigated.
4. During this four stage process the specific research tools get refined (e.g. coding schemes get developed, tested and revised) so that eventually empirical studies can be designed and implemented to test the theory.

Our previous work examining collaborative student learning in web-based environments addressed the *first* stage of this process. In this paper, we describe our work on the *second* stage, refining and reformulating specific categories. Our objective was to analyze the interactions and social relationships between students as they worked together on computer-based math modules. At this stage of this research agenda, any analysis or conclusions must be tentative, preliminary, and subject to revision and further data collection. Our purpose in this paper is not to propose a

comprehensive theory but, instead, to identify and clarify issues and suggest hypotheses that we believe will lead researchers into the third stage of the development of a grounded theory of learning mathematics in technology rich environment. We hope that the small steps we take in this paper will contribute to moving this endeavor forward.

II. Background and Literature Review.

Mathematics educators largely agree that students in mathematics classes should: investigate meaningful contextualized applications of math; utilize technology to solve problems; work together cooperatively; and engage in collaborative discourse to use language to communicate mathematical ideas (Zemelman, Daniels and Hyde, 1998; National Research Council, 1991). Many of the projects that grew out of the calculus reform movement of the 1990's also reflect these goals (McCallum, 2000).

Duke University's Connected Curriculum Project (CCP) was developed in this educational and historical context. "The Connected Curriculum Project is a coordinated effort to create interactive learning environments for a wide range of mathematics and mathematically-based applications. Our materials combine the flexibility and connectivity of the Web with the power of computer algebra systems. These materials may be used by groups of learners as an integrated part of a course or by individuals as independent projects or supplements to classroom discussions." (<http://www.math.duke.edu/education/ccp/aboutccp.html>). The research discussed in this study examines students working on CCP modules.

The uses of computers and technology in support of learning mathematics have been well documented in the research literature. As early as the 1970s, researchers such as Papert (1980) were studying the ways computers might foster greater understanding of mathematics. By the early 1980s, Kelman and associates (1983) had completed extensive studies describing the potential role of computer technology in mathematics education. There was great hope that computers would make possible new approaches to teaching and learning mathematics.

Expectations have been high that computers and technology would have significant effects on instructional practices and learning outcomes. However, as educational historians such as Cuban (1986, 2001) have indicated, the expectations for technology have typically been significantly greater than the actual outcomes. Pea (1987) provided a historical perspective on the transformational roles played by computers and advanced technology in mathematics education. Pea indicated that although the computer has the potential to serve as a mediational tool for promoting dialogue and communication on mathematical problem solving, computers have rarely been used to facilitate this function explicitly. (p. 105)

More than fifteen years have passed since Pea and Cuban first questioned how well educators have incorporated computers and technology into their instructional practices. Today, very few critics would question that significant progress has been made in discovering meaningful ways that computers can foster the learning (Aleven, Stahl, Schworm, Fischer, and Wallace, 2003; Farrell, 1996). For example, Becker and Riel (2000) reported that when computers are integrated into constructivist instructional approaches, computers could become effective tools for improving students' learning. Ellis (2000) argued that "technology is changing the way calculus is taught and learned, as well as the topics presented and the interactions in and out of the classroom" (p.67). Dubinsky, Matthews, and Schwingendorf (2001) indicated that the thoughtful use of technology can be very beneficial to student learning. Goos, Galbraith, Renshaw, and Geiger (2003) found that technology can serve as a "discourse tool" which is

useful in mediating class discussions and changing the ways teachers and students interact with each other and with learning tasks.

Along with the research on technology and mathematics, an equally rich research literature exists that examines the use of interactive social contexts and cooperative learning in mathematics education (Dubinsky, Mathews, and Reynolds, 1997). Much of the research in this area is based on the foundational work of cognitive psychologist Lev Vygotsky (1978) who maintained that the development of higher level thinking in mathematics is rooted in social interactions. More current theorists such as Noddings (1990) and Schoenfeld (1985) have contributed to this understanding of mathematics as a social activity. These researchers, as well as others, have demonstrated how dialogue and structured social interaction among mathematics learners can be helpful in fostering mathematical thinking and conceptual change.

For example, in their edited book, *Cooperative Learning in Undergraduate Mathematics*, Rogers, Reynolds, Davidson, and Thomas (2001) examined issues surrounding the use of dialogue and student-to-student collaboration in college mathematics classrooms. They concluded that "mathematics problems are particularly well suited for group discussions because they have solutions that can be logically demonstrated." (p. 3). These researchers pointed out that a meta-analysis (Springer, Stanne, and Donovan, 1999) of studies involving college mathematics students indicated that cooperative learning has significant positive effects on achievement and attitudes among undergraduates learning mathematics.

Van Zee (2000) used audio-tapes and video-tapes to examine the nature of student-to-student discourse in a science and mathematics education seminar. Van Zee interpreted the dialogue among what she termed "collaborative sense making" students in the seminar to determine instances of inquiry learning, student questioning, and collaborative sense making. For instance, Van Zee examined the specific questions students asked each other about a particular issue having to do with the phases of the moon. The framework for analysis of the students' conversations in the classroom "was based on a negotiation metaphor" (p. 119) that identified instances of students helping other students make their meanings clear. Van Zee concluded that both "students and teachers can build principled knowledge through joint talk and action." (p. 137).

In recent years the phrase "social negotiation of meaning" has appeared more frequently in the research literature. Woolfolk (1995) defined "social negotiation" as an "aspect of the learning process that relies on collaboration with others and respect for different perspectives." (p. 482). Woolfolk indicated that when the American Psychological Association Task Force on Psychology in Education published its twelve "Learner Centered Psychological Principles" that the ninth principle stated that "learning is facilitated by social interaction and communication with others in flexible, diverse, and adaptive instructional settings," (p. 480). Woolfolk noted that the notion that students develop higher mental processes through collaborative discourse and the social negotiation of meaning is rooted in the work of Vygotsky and is an underlying principle of constructivist approaches to teaching.

Alexander and Murphy (1998) noted that, "Learning is as much a socially shared undertaking as it is an individually constructed enterprise. One of the most powerful observations that has emerged in the psychological literature in the past several years ... is the recognition that learning is continuously and markedly shaped by the social context in which it occurs" (p. 41). They quote Resnick (1991) who argued that:

Recent theories of *situated cognition* are challenging views that the social and the cognitive can be studied independently, arguing that the social context in which cognitive

activity takes place is an integral part of that activity, not just the surrounding context for it ... every cognitive act must be viewed as a specific response to a specific set of circumstances. Only by understanding the circumstances and the participants' construal of the situation can a valid interpretation of the cognitive activity be made (p. 4).

As we will argue in this paper, these insights are relevant to developing a framework for understanding learning in the collaborative interactive technology-rich environments explored in this study.

This emphasis on social interaction has given rise to a significant body of research that examines various aspects of cooperative and collaborative learning in mathematics. For many years, organizations such as the Network for Cooperative Learning in Higher Education have disseminated research on both cooperative and active-learning (http://www.csudh.edu/SOE/cl_network/default.htm). As Cooper and Robinson (1998) pointed out, the evidence for the effectiveness of cooperative learning in science and math instruction is strong: "Perhaps the most compelling evidence regarding the power of small-group instruction in SMET (Science, Mathematics, Engineering, and Technology) disciplines comes from a recent evaluation of over 340 NSF project directors. They were asked to evaluate which of 13 possible innovations in undergraduate teaching were central to effective teaching. Students working in teams was ranked highest of the thirteen."

Our review of the research on collaborative learning in mathematics also revealed a number of researchers specifically investigating the interaction of technology and collaborative learning. For example, Roschelle, J., Pea, R., Hoadley, C., Gordin, D., and Means, B. (2001) investigated the ways computers can be used to improve learning in the classroom in light of four fundamental characteristics of effective learning: active engagement, participation in groups, frequent interaction and feedback, and connections to real-world contexts. The researchers indicated that although some critics maintain that the computer fosters asocial behavior, the use of computers to facilitate educational collaboration is increasing dramatically. Roschelle et al noted that: "Reports from researchers and teachers suggest that students who participate in computer-connected learning networks show increased motivation, a deeper understanding of concepts, and an increased willingness to tackle difficult questions."

Another sign of rising interest in the interaction of collaborative and technology is a recent issue of the journal *Educational Psychologist* (Volume 40, Number 4, Fall 2005) which was focused entirely on ways computers can be used as metacognitive tools for enhancing learning. Included in this journal is a study by White and Fredericksen on the development of self-regulatory skills among fifth graders working collaboratively with computers. The researchers used videos of students working together in classrooms, as well as interviews with students and teachers, as a basis of their analysis. They concluded that the collaborative use of technology is not only highly engaging, but leads to the development of metacognitive knowledge and skills necessary for collaborative inquiry and reflective learning.

Other researchers have also examined the use of instructional approaches that effectively combine both technology and collaborative learning. For example, Edelson, Pea, and Gomez (1996) argued that "math and science reforms of the 1960's that were most successful were not just those that emphasized the active nature of the learning through manipulatives and hands-on inquiry, but also those that provided opportunities for students to talk while they were engaged in learning, interacting about what they were learning, what they believed, and what they had difficulty understanding." (p. 152). These researchers developed the Learning Through Collaborative Visualization Project (CoVis) for high school science classrooms. The CoVis

Project utilized computers to engage students in open-ended scientific investigations; students worked collaboratively with other students, with teachers, and with scientists. The researchers indicated that: “social interactions enhance the learning that students achieve through the transformative process of communication” (p.162). The researchers suggested that technology has the potential to enhance social interaction and serve as a mediational tool. The researchers concluded that teachers must begin “to take advantage of these sorts of new technologies to provide their students with opportunities for active learning and meaningful social interaction about scientific subjects” (p.162).

Despite this compelling evidence, our review of the research on collaborative learning of mathematics in technology rich environments yielded little research that focused more specifically on the styles of interacting or learning styles that paired undergraduates establish as they work together cooperatively. Although Inkpen, McGrenere, Booth, and Klawe (1997) examined interaction styles in educational computer environments, their focus was chiefly on “computer mouse interactions.” And, while Ross and Lukow (2004) investigated the predictive value of individual learning styles for integrating technology into the curriculum, they focused primarily on the attitudes and learning styles of individual children.

Terms such as learning styles and cognitive styles have long been used by educators and psychologists to describe the different ways that individual learners approach tasks-- their preferences and approaches to doing academic work, as well as their preferred ways of processing and organizing information. Dunn, Dunn, and Price (1984) developed an instrument to measure the learning styles of students, including students’ preference for visual versus auditory instruction, working alone versus with others, and psychological inclinations such as working reflectively versus impulsively. While the idea that students bring established learning styles to individual learning tasks seems widely accepted, few studies have examined the interaction of learning styles and collaborative learning. A question of interest in this study was: Do pairs of students who work together to solve a mathematics learning task establish a “collaborative learning style”?

In reviewing the literature on the uses of computers in mathematics education, we discovered that the phrase “mediational tool” was used by researchers to communicate the notion of the computer as intermediary between the mathematical concepts and the learner. At times the term mediational tool has been used explicitly by researchers such as Pea (1987), and by Goos, Galbraith, Renshaw, and Geiger (2003) who indicated, “little consideration has been given to the pedagogical implications of technology as a mediator of mathematics learning” (p.1).

Often, however, this notion of the computer acting as a “go-between” or a mediator is implied. These references in the literature to the computer as a mediational tool describe the potential of the computer to mediate the process of learning by bridging the gap between the learner’s current understanding and the new concepts being taught. For the purposes of this study, we defined mediation of learning to mean the process of promoting learning by providing to the learner a tool to assist in making connections between new concepts and existing schema.

With this in mind, the primary objective of our research was to examine the behaviors, interactions, and conversations between students who were using computers to learn mathematics. Two questions of interest to us were: To what degree does the computer “mediate” or foster conversations and social interactions having to do with learning mathematics. And, what patterns of behaviors and interactions emerge?

III. Methodological Issues.

As mentioned above, the study reported in this paper builds on an earlier, preliminary analysis of videotapes of students using Connected Curriculum Project (CCP) modules (Bookman and Malone, in press). This is consistent with our view of how grounded theory is developed. The first paper focused on identifying categories and questions generated by the data; in this second paper, we reexamine the data focusing on a particular category (in this case, interactions and social relationships between students as they worked together on computer-based math modules). The subjects and the data collected were, therefore, the same as in our earlier study.

The subjects were college students at a highly selective research university taking a mathematics course (at a level beyond calculus). The subjects were each paid \$25 for volunteering to be videotaped for the purposes of this study. The research procedures and consent forms were approved by the University's Institutional Review Board. Their participation in the study consisted of working through one of the CCP modules with a partner. The students working together were videotaped and their computer output was simultaneously collected on a separate videotape. Each session was 1-2 hours in length and data were collected from a total of 10 pairs of students. The students were familiar with the format of the modules and *MAPLE* (the computer algebra system) having used them for several weeks in their mathematics course work. For all but one pair of the students, the particular module used in the study was a requirement for a course in which they were enrolled. The subjects had been in class together and, in most of the cases, had previously been lab partners with the person they had been paired with in this study.

The data were gathered in an office (rather than a computer lab) so that the videotaping could be done more effectively. For most of each session, one of the investigators was present in the room, serving the same role that the instructor would serve in the computer lab. Pencil, paper, and a computer with *MAPLE* were on the table, as well as a video camera to record their work and a scan converter connected to a VCR and television to record their computer output. Videotapes of ten pairs of students were collected. We chose vignettes from five of the ten pairs of subjects whose behaviors most clearly illustrated or typified the categories generated. Because the cost of transcribing all these conversations was prohibitive, we identified vignettes in these tapes that seemed particularly interesting and transcribed those vignettes, leaving out extraneous verbiage that did not convey any added meaning. In our second study, we revisited these tapes focusing on the social interactions between the subjects. Because of this particular focus, we transcribed these vignettes more exactly and in greater detail, also adding in descriptions of the nonverbal behavior that was observable.

A unique aspect of the current study was that the subjects were videotaped working together and, simultaneously, their computer output was recorded on videotape. Using these simultaneous video recordings as a method of investigating student learning in computer labs has not, to our knowledge, been reported in the research literature. This methodology provided the researchers with an opportunity to closely examine and document student behavior. Viewing both tapes simultaneously was necessary because it is not possible to understand the students' dialogue and interactions without seeing both what the subjects were seeing and what they were working on. Examples of these tapes can be seen at the links below (Note that camera1.rm is paired with computer1.rm):

<http://www.math.duke.edu/~bookman/Camera1.rm>

<http://www.math.duke.edu/~bookman/Computer1.rm>

<http://www.math.duke.edu/~bookman/Camera2.rm>

<http://www.math.duke.edu/~bookman/Computer2.rm>

Using this methodology, we focused on one of our research categories, the role of social interaction and collaborative discourse in computer-based mathematics instruction. We began by re-watching the tapes, paying particular attention to social interactions. We catalogued the social interactions and behaviors and then reorganized by clumping and condensing these behaviors in order to determine “which phenomena share sufficient similarities that they can be considered instances of the same concept” (Gall, Borg, and Gall 1996, p. 564). This iterative and recursive process required frequent reformulation of the categories which required frequent re-examination of the tapes. Our goal was to extract categories from the data that were coherent, self-contained, sufficiently general, and recognizable and we believe that the process of cataloguing, clumping, condensing and reexamination of the tapes allowed us to make significant progress towards that goal. This is consistent with Romberg’s (1992) method of clinical observations where “the details of what one observes shift from predetermined categories to new categories, depending upon initial observations” (p. 49). It is also consistent with the principle of grounded theory that one generates conceptual categories from evidence and that the categories that “emerged from the data are constantly being selectively reformulated by them. The categories, therefore, will fit the data, be understood both to sociologists and to laymen who are knowledgeable in the area, and make the theory usable for theoretical advance as well as for practical application” (Glaser and Strauss, 1967, p. 249).

To place our analysis of the data in a context, we describe below the three CCP modules on which our subjects worked. In one module, The Equiangular Spiral, <http://www.math.duke.edu/education/ccp/materials/mvcalc/equiang/index.html>, students examine properties of the chambered nautilus to learn about equiangular spirals in general. The lab also provides an opportunity for students to review polar coordinates. The students are given a picture of a chambered nautilus, superimposed on a polar grid, and asked to show that the radius is an exponential function of the angle theta. They also develop the mathematical basis for why these spirals are called equiangular. In another module, Rotation Matrices, <http://www.math.duke.edu/education/ccp/materials/linalg/rotation/index.html>, the students learn how to use matrices to represent rotations in the plane, and rotations in space about one of the axes. They learn about the relationship between multiple rotations and matrix multiplication and about determinants and inverses of rotation matrices. In the last module, Linear Correlation and Regression, (<http://www.math.duke.edu/education/modules2/materials/test/test/>) students examine scatter plots to learn about correlation and lines of best fit. They also examine the difference between correlation and causation.

IV. Analysis of the Data.

Our observations of students working on these three CCP modules provided evidence that suggests different ways students interact while learning math in a technology rich environment. For example, one of the most apparent and recurring observations was that the students focused their attention almost entirely on the computer that served as both a mediator and object of their communication. They conversed with each other by pointing to objects on the computer screen and did so while rarely looking at each other. In addition to these more easily recognized behaviors, an analysis of students’ work revealed more complex interactions that will be described below. Two thematic categories emerged: (1) establishment of roles and (2) social

negotiation of meaning. We use the term “social negotiation of meaning” as described above by Woolfolk (1995) and we examined three roles – who controlled the mouse and keyboard, who made decisions concerning the direction and pace of their work, and who served as checker or verifier. These roles were not assigned but were established by the subjects as they proceeded through the assigned modules. Below we describe several vignettes that are exemplars of the categories that emerged from the data.

A. Establishment of roles.

Observations of students working collaboratively in front of the computer revealed that some students explicitly decided who would be responsible for what task, while others arrived at these decisions less consciously and without discussion. For example, the following vignette illustrates how a pair of students verbalized and established who would control the keyboard and mouse. The names used here are pseudonyms; the real names of the subjects were not used.

Alex: Here, why don't you type dude? *(Looks at Neil while speaking to him)*

Neil: Are you sure? *(Looks back at Alex and raised his eyebrows questioning Alex's decision. At this point, Alex stands up and begins switching seats)*

Alex: Yeah, yeah. *(mumbles something inaudible)(Neil begins to take the chair in front of the computer)*

Neil: I thought you wanted to type. *(Sitting and readjusting the keyboard)*

Alex: You're better with commands. *(getting seated)*

Investigator: So what's the deal? Does he usually...*(Neil begins to shake his head in disagreement)*

Alex: Uh, he uh, he did it before because he knew MAPLE *(Alex looks at the investigator in the room while addressing him)* and, I kind of took the last couple.

Neil: We take turns. *(says this while still looking at the screen)*

Alex: Yeah, it's his turn anyway.

In another vignette, Hope and Amit explicitly discussed role assignments. This conversation occurred just after they sat down to begin working.

Hope: Here. Do you want to use the keyboard or mouse or do you care? *(it looks as if she might be pushing the keyboard or mouse closer to him as she asks her question.)*

Amit: You can have both of them and be happy. *(They both laugh at this)*

Two minutes later Amit takes the mouse while Hope is writing at the board. He passes it back to her as she sits down.

Hope: No, go ahead, it doesn't matter.

Amit: *Pushes the keyboard towards her.* No, go ahead.

This interaction was the first of many times that Amit grabbed the mouse or keyboard when Hope was away from her position, but he seemed to relinquish it when she returned. From his knowledge of her, Amit sensed (correctly) that Hope wanted to control the mouse and keyboard and was just being polite in offering it to him. This vignette illustrates a more subtle way of establishing the control of the keyboard and mouse than in the case of Neil and Alex.

Whereas Neil and Alex comfortably and naturally made this decision, there was more tension in how Hope and Amit decided on their role assignments.

In some cases, students who had worked together prior to the videotaped session had already established roles in advance. For example, Kevin, a math major, and Carl, an electrical engineering major, had been lab partners for most of the semester prior to the day when their work was videotaped. They had been working on the module for several minutes when the investigator asked:

Investigator: So you guys have a routine down yet, working together?

Kevin: *(Shrugs and turns to the investigator)* We take turns typing, although Carl types more.

Carl: Yeah, I get along better with the computer; he gets along better with the math. *(Carl finally looks up to acknowledge the investigator, who he is talking to.)*

Sometimes, as in the case of Andy and Larry, no discussion of establishing roles took place. Andy just sat down and took charge of the keyboard and the mouse. These vignettes describe only four instances, on a continuum from explicit to unspoken, about how decisions were made concerning who controlled keyboard and the mouse. These data don't provide an explanation of how these roles were formed; a future study that includes follow-up interviews with the subjects might provide some insight on this question. As we will discuss later, we believe that these observations are consistent with other behaviors of these pairs and might lead to some categorization of the different ways pairs of students work together.

After the pairs of students established who controlled the keyboard and who controlled the mouse, they began to work on the module. At certain critical points in the problem solving process, the students had to establish roles having to do with making decisions about how next to proceed. At these transition moments, these decisions were sometimes jointly made and sometimes made by one individual.

For example, the following vignette illustrates how Carl and Kevin made a decision about who would control navigation. In this situation, Kevin needed to assert himself in order to get Carl to slow down so he could get his question answered:

Investigator: Do you know what standard deviation means?

Carl: Yes.

Kevin: I kind of know intuitively what it means, *(he the looks to the investigator)* is there a good definition?

Investigator: Yeah, there is. You can click on the link.

Kevin: Do you want to click on that? *(Kevin points to the link with the eraser end of his pencil. However, Carl ignores Kevin and continues typing in the answer with which he has been working. At this point, Kevin takes control of the mouse.)*

Kevin: I am just going to click on that and see what standard deviation means.

Carl: Okay. *(Carl sits back in his chair and yawns while he waits for Kevin to read the definition of standard deviation. Kevin finishes reading the link on standard deviation and closes the window but is unsure how to use the computer to retrieve the module, so he relinquishes control of the mouse again to allow Carl to re-open the window.)*

This exchange was typical of their division of labor as when Carl was explaining why he was at the keyboard and he said, "I get along better with the computer; he gets along better with the math."

In another case, even though Jim did all the typing, his partner Mary directed the decisions for the computing process.

Mary: See, this (*points to paper*) and this (*points to a new spot*) are what you want to see. Natural log of r , so if you type it in like this it should work. (*pointing to the piece of paper that they had been given by the investigator*).

Jim: Take the zip and rewrite l and r ?

Mary: Yeah.

Jim: Okay. Oh I see.

This was one of many examples we observed where, in some of the pairs, the person controlling the keyboard and the mouse was not the person in control of some of the decisions about how next to proceed.

In almost complete contrast, Larry made numerous suggestions to his partner Andy who usually ignored him.

Larry: I am surprised that you can't just get it [the computer] to find that for you. (*both Andy and Larry look at each other and then Andy turns back to work on the problem*).

Just set up a function now that iterates from like 1 to... (*Andy turns and looks at Larry again. Andy is smiling and silently laughing*)

Andy: No. (*Andy continues smiling, but he does not ever look away from the problem, and his language is very curt.*)

Larry: 10. (*Larry is still looking at Andy*)

Andy: No. (*Andy continues not to acknowledge Larry*)

Larry: by 0.001

Andy: No. (*looks at Larry this time when he responds, but continues to use the curt tone*).

Larry: And return the one.

Andy: No, we are not doing that. (*Shakes his head no as he turns back to the computer again*)

These vignettes demonstrate some of the ways in which decisions were made and how roles were established, varying from shared to unilateral. In the case of Andy and Larry, Andy controlled the mouse and keyboard and also directed the decision making process. This left Larry no role to play and feeling like an outsider or observer of the learning process. In the case of Kevin and Carl, although Carl controlled the keyboard, Kevin insisted on making critical decisions when he felt he needed to. If the person controlling the mouse and keyboard was not the key decision maker, a pattern of advice and consent by the keyboarder often emerged. As seen above, this was the case with Mary and Jim. What we saw in our data was that, in the well functioning pairs, the person not at the keyboard had an equal or greater share of the decision making.

In some cases, the subjects also established roles concerning who would take primary responsibility for mathematical thinking. For example, in the case of Carl and Kevin, Kevin

assumed primary responsibility for that role, as was clear when Carl said, “I get along better with the computer; he gets along better with the math.” In fact, throughout the module, Kevin (a math major) almost always did the pencil and paper algebraic computation and other mathematical thinking. Carl (an electrical engineering major) worked equally hard on the technical (e.g., syntax) aspects of the problem. Although there was a clear division of labor for who was leading in a particular task, they each took responsibility to understand what the other was contributing. As we have pointed out this is consistent with their collaborative style.

A variation of this sharing of the mathematical thinking is seen in the work of Mary and Jim (Bookman and Malone, in press). The following vignette can be viewed by using the following links to Realplayer files: <http://www.math.duke.edu/~bookman/Camera1.rm> and <http://www.math.duke.edu/~bookman/Computer1.rm>. After a few minutes of trying to remember how to get Maple to compute derivatives (they needed to find the derivative of $x=r_0 e^{k\theta} \cos \theta$ with respect to θ), Mary gave up and said, “We can just do it by hand.” She began to do the calculation on paper, but Jim said, “I’m trying to remember how Maple works.” After about a minute, Mary completed the calculation by hand and then said:

Mary: Okay.

Jim: Shut up. (*said in friendly and jocular manner*).

Mary: (*laughs*) Here, it’s just the product rule.

Jim: Yeah. It would be nice if Maple will do it for us.

Mary: It will.

Jim: Yeah. I want it to do it.

A minute later, working together, they got MAPLE to do the calculation.

Mary: You see, it’s exactly what I did.

Jim: Yeah, but your way is stupid.

Mary: But it was quicker.

The instructions then asked them to divide $dy/d\theta$ by $dx/d\theta$ to get a formula for dy/dx . Although this computation would have been quite difficult to do by hand, they were now (because they had figured out the correct syntax) able to use *Maple* to do this computation in a couple of seconds. The instructions then directed them to evaluate an even more complicated expression that reduced to $1/k$. Jim said, “Wow. I want to work this out on paper. I don’t believe that.” Here, although their roles are reversed, with Jim advocating use of pencil and paper, they shared the responsibility for making the mathematical decisions. One might think that the person in control of the keyboard and mouse controlled the pace and direction of their work as well as their mathematical thinking. But our analysis of the data, as illustrated above, indicated that these responsibilities were shared more frequently than we had originally anticipated.

Another role, sometimes taken on by the student not at the keyboard, was that of checker/verifier. In two of the five cases, the pair of students worked closely together where the non-keyboard person monitored their work, acting as the checker/verifier. For example, while Mary and Jim were trying to figure out the best fit line for the data points, they had the following conversation:

Jim: What do you think the formula is? It is going from 18-70.

Mary: What is?

Jim: The data points. It looks like it is doubling for every gap of 2.

Mary: No, not doubling. It is more like multiplying by 1.5.

Jim: Should we say 1.5 then?

Mary: Yeah, I think somewhere between $\frac{4}{3}$ to $1\frac{1}{2}$ relationship.

Jim types this in.

Notice that Mary, who was not at keyboard, was instrumental in checking and verifying that the work is correct. Again, this is consistent with what we saw in pairs that worked well together.

In another example, Alex and Neil demonstrated similar behavior of establishing roles of checker and verifier.

Neil: All right. What do the pictures say to us about data with correlation coefficients near +1 or -1? *(Neil reads the problem out loud, while both he and Alex read the problem off the computer monitor.)*

Alex: They fit. *(He speaks while still looking at the computer screen).*

Neil: All of the... *(moves his fingers in almost a snapping motion as he tries to say points)*

Alex: Yeah, the tightness of fit. *(Neil types this into their answer sheet)*

Alex: Of 1 is perfectly linear *(Neil types this in as Alex says it, at the same time shrugging his shoulders and raising the corner of his mouth to the statement that most are 0.99), most of them are like .99 or something. (Neil nods his head left to right, but his facial expression seems to be frustration that he has typed something into the worksheet incorrectly)*

Neil: Is this fine? *(Turns wrists palms up, like a mini-shrugging motion to question Alex).* No this right here.

Alex: That's fine. *(nods his head in agreement)*

Neil: Approaching? *(Turns his one hand upward in a questioning gesture and nods his head while raising his shoulders)*

Alex: Yeah that's good.

Neil: What correlation...*(he and Alex both are intent on reading the screen)* Oh wait, correlation of 0. *(begins typing again)*

Alex: There's no relationship between the lines. *(rolls his eyes up, as if he is thinking about what he is saying. Neil types it in, still looking at screen)*

Alex's short, quick comments indicated to Neil that he was in agreement and that they could proceed. We noticed that each of the pairs of students developed its own style for checking work. The checking and verifying by the non-keyboarder appeared to help these students focus on the learning situation.

Conversely, when the student not at the keyboard was not the checker, we saw, in the case of Andy and Larry that their work proceeded badly. Andy and Larry had trouble establishing roles and determining how their work would be checked and verified. Failure to establish these roles often led to a breakdown in the learning process. For example, Andy's refusal to listen to Larry's suggestions resulted in Larry being less focused on their work and Andy going off in wrong directions. In fact, for several minutes, Larry, who later turned out to be right, suggested to Andy that they must be on the wrong track. Several times Larry said things

like, “That can’t be right.” and was totally ignored. These difficulties communicating with each other clearly impeded their efforts to solve the math problems presented in the module.

An alternative hypothesis is that the lack of understanding determined what we observed in the collaborative styles. We feel, though, that this data provides some tentative evidence for the opposite view – that collaborative style affected the student’s progress through the assignments. As will be seen in a vignette to be discussed later, Larry did understand that something was wrong but couldn’t get Andy to listen to him. And we observed that Mary and Jim’s cooperation helped them overcome difficulties in their understanding.

B. Negotiating Meaning.

The establishing of well defined roles is closely related to issues concerning the negotiation of meaning, which was the second thematic category that emerged from the analysis of data. We turn next to examining situations in which students had to negotiate meaning and understanding through dialogue, writing, and non-verbal communication. By “negotiating meaning” we refer to the process of collaborative discourse in which students take turns putting their understanding of a problem into their own words by challenging and building on the ideas their partner has expressed. We include non-verbal communication because analysis of the data indicated that students frequently pointed to the computer screen and made other gestures in an effort to communicate their understandings.

For example, the following vignette illustrates how a pair of students collaborated to construct a shared understanding of linear regression. In the linear regression module, the subjects were asked, “Given scatter plots of Test 2 scores versus Test 1 scores and Test 2 scores versus Test 3 scores, if a student scored an 82 on Test 1, what do you predict he or she would score on Test 2? If the student scored an 82 on Test 2, what would you predict for his or her score on Test 3? Which prediction do you expect to be more accurate? Why?” As they were trying to construct a written response to this question (the CCP modules typically require students to express their understanding in writing), the following dialogue occurred between Carl and Kevin:

Carl: So...(They are beginning to answer a problem about predicting a students’ test scores from the data).

Kevin: Test 2 and Test 3 you can’t fit because they are not related. (Carl continues to type out their joint written response) Therefore, any prediction for Test 2 would be more accurate than Test 3.

Carl: Based on Test 1.

Kevin: Yeah. (Both are looking at the screen and then they read part of the problem)

Carl: Just say you can only make a judgment on positive and negative association if there is some sort of linear association.

Kevin: If you can more easily fit a line though the data and it is easier to make predictions and find a relationship.

Carl: Do we need to explain here?

Kevin: In the second scatter plot, there is no line that is going to fit nicely, therefore it is hard to come up with a relationship and make predictions. I guess it wouldn’t have to be a line, but since we are talking about linear stuff, focus on the line. (Makes a suggestion)
An informed guess.

Carl: Form a guess. *(repeats Kevin's words back and continues to type out their written response)*

Kevin: *(makes another suggestion as to wording that could be used in the written explanation used)* I would say, in other words, it is hard to guess a student's test score.

Kevin offered a tentative provisional understanding that Carl responded to. They went back and forth until they were comfortable that they shared an understanding of this particular concept. This form of discourse provides students the opportunity to articulate their understanding and to seek agreement on meaning. Vygotsky (1978) maintained that peers can work together to co-construct knowledge as they provide cognitive scaffolding for one another. Kevin and Carl appear to be providing cognitive scaffolding for each other as they actively negotiate their written solution of the problem.

Another example of social negotiation of meaning is seen in the following dialogue between Mary and Jim (a dialogue we used earlier in this paper to illustrate a different point). The two students were trying to find a functional model for some data points that appear to be growing exponentially.

Jim: What do you think the formula is? It is going from 18-70.

Mary: What is?

Jim: The data points. It looks like it is doubling for every gap of 2.

Mary: No, not doubling. It is more like multiplying by 1.5.

Jim: Should we say 1.5 then?

Mary: Yeah, I think somewhere between $\frac{4}{3}$ to $1\frac{1}{2}$ relationship. *(Jim types this in and as he types he appears to be thinking deeply about it).*

Jim: That's not what it does. This does not look correct.

Jim and Mary then reconsidered their response and jointly tried to make sense of the problem. It is not until later, and after the investigator intervened with some advice, that Mary and Jim solved the problem. However, this short dialogue exemplifies the type of exchanges that were quite typical among the pairs of students who appeared to be working together effectively. These dialogues are typified by a bantering quality, with short sentences and polite interruptions. We refer to this kind of dialogue as "cognitive bantering." These types of discussions, of course, are typical of discussions in other, less technology intensive cooperative learning environments. We expect, however that technology (perhaps particularly in the area of mathematics) may be able to increase the likelihood of these opportunities for collaborative and active learning because the screen is a physical object that focuses their attention, and the objects on the screen can be changed quickly to respond to inputs from the users. For example, as in the case of Jim and Mary above, the computer algebra system provides immediate feedback to their hypothesis that 1.5 is the correct parameter.

As discussed earlier we analyzed videotapes of five pairs of students working on three different CCP modules. The five pairs were chosen from the ten pairs of students videotaped because highly the behaviors of these 5 pairs of subjects most clearly illustrated or typified the categories generated. Of the three CCP modules, the linear regression module produced particularly highly interactive conversations. This module included the java applet Guessing Correlations, <http://www.stat.uiuc.edu/~stat100/java/GCApplet/GCAppletFrame.html>. The applet shows students four scatter plots and gives them four correlation coefficients; their task is

to match the correlation coefficients with the scatter plots. If their matching is correct, they get a “point” and can continue to see how many matchings they can get correct. This particular applet seemed to capture the students' interest and appeared to increase the intensity of the discussion. The following is an excerpt of Neil and Alex working on this applet:

Alex: The other one is pretty weak though *(Neil points to the screen)*. It looks like, uh, yeah it looks like, no, it's not that bad though. Maybe, uh, maybe use 0.3.

Neil: Point 3? *(Raises his eye brows while asking this in a unbelieving tone)*

Alex: I mean. Yeah, I said point 3. *(Neil looks over at him like he cannot believe what his partner is saying; he raises his eyebrows and scrunches his nose)*

Neil: It's not point 3; it's not nearly that bad. *(Neil looks over at Alex and points to the graph on the screen)*

Alex: No it's better than that, that's why I'm saying it's point 3. *(Alex points to a lower figure on the monitor)*

Neil: Oh that's a negative? What are you talking about, that's 0.7. *(and points to the figure that he had originally pointed to again, and turns toward Alex waiting until he replies)*

Alex: That's point seven *(motions with pen to scatter plot on screen)*, that's not point 7. *(motions with pen to scatter plot on screen)*

Neil: No that's .96 *(Neil points to the graph)*

Alex: Oh really. Here go down. *(Pointing toward the scrollbar, indicating that he wants Neil to scroll downward)* So, what do you want to say then?

Neil: Huh?

Alex: You want to guess like .7?

Neil: I would say like, I would say .75. Or .95 and .75.

Alex: All right that sounds good. *(Shakes his head in agreement)*

Neil: Oh, and they are both positive, right?

This transcript of the conversation does not fully reflect the high level of engagement and the intensity of Neil and Alex's interaction. The Java applet (perhaps because of the game-like nature of the applet or the immediate feedback it provided in response to the students' predictions) appeared to engage the students deeply in a conversation about the mathematical problem. The applet uses variable and positive reinforcement methods from behaviorism yet also provides an opportunity for students to support each other as they develop and test their understanding. This combination of Skinnerian ideas of reinforcement together with Vygotskyian notions of social learning seemed to have a powerful effect on students. The students' natural and seamless dialogue is another example of what we refer to as cognitive bantering as each student took turns offering a provisional answer and then waited for the response of the partner. In this particular case, the immediate and visual feedback provided by the applet appeared to make this productive dialogue more likely to occur than it would have with a pencil and paper task. The computer appeared to serve as a mediational tool for fostering the students' thinking aloud activities.

This same scatter plot applet also appeared to bring about a clear change in the ways other pairs of students interacted when they began work on the applet. For example, Carl and Kevin appeared to have different goals and priorities throughout most of their work session. As we discussed earlier, Carl preferred being responsible for understanding the software and Kevin

preferred being responsible for the mathematical understanding. However, when Carl and Kevin began working on the scatter plot applet, their dialogue became more focused on the mathematical problem as seen below:

Carl: I would say this one is going to be 0.94.

Kevin: There are two that are positive. That one (*points with his pen*) is a closer correlation than that one (*uses his pen to point again*) so, yeah A would be 0.94 and D would be 0.47. The other ones are going to be harder to predict because they are both negative.

Carl: I think this one is better than that one. What do you think?

Kevin: Yeah. (*Steven takes his pencil to the screen to try to make slope predictions*) What I am looking at here is that they are almost evenly distributed on either side of the lines. (*uses his finger to point to this on the screen*) What do you think?

Carl: I like this more because ...(*Carl proceeds to give his reasoning*)

For this activity, they seemed to equally share the responsibility for understanding the mathematical concepts and for the proper use of the computer.

In analyzing the videotapes we noted that during these moments when the pairs of students were engaged with each other's thinking and with the mathematical problem they were attempting to solve, the students seemed "in-synch." Aspects of being "in-synch" include active listening, asking each other questions, and feeling comfortable challenging each other in a constructive way; these are also behaviors often associated with meaningful learning. These moments in which students appeared to be "in-synch" also seemed to be the moments when the most learning was taking place.

On the other hand, the fact that students worked together using a computer to solve interesting math problems did not always result in "in-synch" collaborative learning. Not all of the dialogues were productive. Even when working on the scatter plot applet, our most dysfunctional pair, Andy and Larry, had difficulty working effectively with each other. Andy rarely took Larry seriously and this lack of respect seemed to contribute to unproductive dialogues even when Larry was offering ideas that would have helped.

Larry: I guess just Test 1, Test 2.

Andy: Nope, no, we are not plotting tests against each other, we want to plot...(*squints his eyes, like he is thinking...*)

Larry: Yeah we were, weren't we? (*Still looking at the monitor*)

Andy: No. (*shakes his head left to right in a short motion and carries tone of annoyance in his voice*)

Larry: I thought we were plotting the data in Test 1 against the data in Test 2.

Andy: No, um, we're plotting Test 1, Test 2 (*points to the screen*) so we want to do those against, just like, (*uses hand gesture, turning palm upright*) the one so that each number represents a 1.

Larry: Oh wait, you mean we are just putting the plots of both of them on the same graph and not actually plotting them against each other?

Andy: Yes. I think that's the idea. (*Mumbles, and nods his head at the same time*)

Larry: Okay.

Andy: *(mutters)* I don't know if this is going to work. *(Mutters under his breath as he enters the numbers and smiles)* This is going to be good though. Ready? *(mumbles to himself)* We just need to try this plot.

Even though Andy and Larry are engaged in a dialogue, it would be difficult to characterize the dialogue as “collaborative discourse” in the sense that this term is typically used. Andy's inability to actively listen to Larry impeded the collaborative learning process. Very little “meaning” is being negotiated because the two students do not establish a shared understanding of what the problem is asking them to do. Andy seemed convinced that his approach was correct (even though he was completely off track) so the qualities of provisionalism and negotiation that we saw in the other pairs of students were not present in this case.

In each of the videotapes we observed instances of students negotiating the meaning of the mathematical problems they are confronted with. In most cases, we observed that the students' efforts to construct a shared understanding reached a high level of engagement and thought. This was particularly true when the CCP modules contained a feedback loop and required the students to actively make predictions and hypotheses.

V. Discussion.

Our analysis of the five pairs of students working on three CCP modules provided insights into the behaviors of undergraduate students learning math in a collaborative technology rich environment. In this study, we've focused on how students negotiate roles and meaning while learning in these environments. Below we summarize the analysis of our observations and offer some tentative conclusions. Some observations are evident from the sample of vignettes discussed above, while others are based on the many hours of videotape that could not be summarized in a few vignettes.

One observation that is repeatedly supported by the videotape data is that the computer plays a significant role in the collaborative learning process. In this study, the computer served the dual role as a mediator between the two students as well as the object of their communication. The students rarely looked at each other while conversing and working together; their eyes were almost always focused on the computer screen. The students pointed to the screen to demonstrate ideas or to make a point. We concluded that the medium of the computer appeared to be more of a “player” in the learning process than a textbook might be.

We also concluded that working in pairs in front of a single computer necessitates that students establish certain roles, such as control of the mouse and the keyboard. Our observations indicate that these roles may not always be discussed explicitly, even when pairs of students work together effectively. Contrary to what many might think, the student in control of the keyboard and mouse did not necessarily control the direction of the learning and mathematical work. The student not burdened with the keyboard and mouse often took on more responsibility for the mathematical thinking, such as assuming the role of verifier. Furthermore, as was seen in the case of Amit and Hope, where control of the mouse changed hands in subtle ways, these roles can be fluid and interchangeable. The only pair of students we observed that experienced significant difficulty in negotiating meaning and developing an understanding of the problem (Andy and Larry) was the pair in which one partner took all the responsibilities and acted in a unilateral fashion.

Much of what we observed confirmed or was consistent with the work of other researchers who have examined the role of technology in interactive learning environments. For example, our observations corroborated the claims of Van Zee (2000), discussed earlier in this paper, that social negotiation of meaning and collaborative sense making appear to help students build conceptual understanding. Analysis of our data supports Alexander and Murphy's (1998) assertion that "learning is as much a socially shared undertaking as it is an individually constructed enterprise" (p. 41). This study provided support for Edelson's et al. (1996) claim that technology can serve as a mediational tool to enhance social interaction and learning. In addition, our findings are consistent with those of Goos et al. (2003), who indicated that technology can facilitate collaborative inquiry through eliciting conversation and discussion among students.

In addition to supporting the findings of prior research, this study also provides evidence for a hypothesis that extends the work of other researchers. Our observations lead us to hypothesize that when pairs of student are placed in a collaborative learning situation, pairs of students often establish recognizable ways of interacting and learning together. We will call this pattern of behaviors a "collaborative learning style." An individual learning style is a preference and approach to doing academic work and a preferred way of processing and organizing knowledge. In contrast, a collaborative learning style refers to the way a pair of students approaches academic work and the ways they interact as they process and organize knowledge. When we used the term collaborative learning style we do not necessarily mean a fixed entity that is immutable and uncontrollable. Collaborative learning styles could change as partners change, as tasks change and other conditions change; collaborative learning styles are probably situational states as opposed to characteristic traits. But, at this stage, these are still open questions and our purpose here is to suggest that these collaborative learning styles may exist and are an object worthy of further study. The impact that these collaborative learning styles have on student behavior in academic situations could have implications for developing a broader theory of how students learn.

This hypothesis was developed in a way consistent with the notion of grounded theory methodology - starting with the data and observations, then seeing what research questions and categories emerged, refining and reformulating those categories, and generating hypotheses. As we focused our observations on the social aspects of student learning in collaborative technology-rich learning environments and as we catalogued, clumped, condensed and reexamined the data, this hypothesis of the existence of collaborative learning styles emerged.

Our observations have led us to tentatively hypothesize the existence of three distinct collaborative learning styles: (1) in-sync or congruent, (2) parallel, and (3) orthogonal. In-sync pairs have shared goals and many of the characteristics we list below that are typical of productive partners (Neil and Alex and Mary and Jim would represent pairs of students with an in-sync learning style). A parallel collaborative learning style is manifested by compatible but different goals, division of labor, and mutual respect (as in the case of Carl and Kevin). Pairs that have an orthogonal collaborative learning style display a lack of mutual respect, differing goals, and the absence of shared responsibility (as in the case of Andy and Larry). These collaborative learning styles may not be fixed. For example, even though Carl and Kevin exhibited a parallel collaborative learning style during most of their work session, when they were using the Guessing Correlations applet they exhibited an in-sync style.

Our observations lead us to conclude that students that exhibit in-sync collaborative learning styles are more likely to become deeply engaged in mathematical problem solving in technologically rich environments. Pairs of students we classify as in-sync tended:

- (a) to feel comfortable interrupting and challenging each other. There existed a shared understanding that challenges were productive and appropriate. We call this back and forth, give and take conversation, “cognitive bantering.”
- (b) to share humor and exhibit an intellectual playfulness.
- (c) to respect their partners.
- (d) to feel comfortable thinking aloud.
- (e) to actively listen to each other in order to understand their partner’s point, often rephrasing their partner’s ideas in their own words.
- (f) to communicate nonverbally (such as pointing and facial expressions) in order to build and demonstrate shared understanding.
- (g) to make predictions and hypotheses.
- (h) to offer provisional ideas as opposed to definitive responses.

The results of this research have theoretical and practical implications for teaching and learning mathematics. Much of the variation we observed in the ways that students went about solving mathematical problems, establishing roles, and negotiating meaning can be explained by examining collaborative learning styles. These collaborative learning styles, which may go unrecognized by instructors, may determine to some extent the success students experience as they engage in interactive learning activities in technology rich environments. In order to better understand under what conditions the use of technology and socially interactive, inquiry-based approaches to learning mathematics lead to student understanding, this concept of collaborative learning styles needs to be further examined. An understanding of the role of collaborative learning styles may have important implications for classroom practice. In particular, teachers need to develop an awareness of what kinds of instructional materials (e.g., java applets like “Guessing Correlations”) are more likely to foster in-sync collaborative learning styles.

VI. Limitations.

As we’ve stated in our initial analysis of these data (Bookman and Malone, 2003), “In interpreting these data, it is important to realize that these students were talented students doing mathematics at a level beyond calculus and using specific software in a laboratory setting. It is not our purpose here to generalize these results to a larger population, but to use these observations to suggest areas for future study. It is also important to note that each entering class of students brings more familiarity, more comfort, and more sophistication with using educational technology. It is not clear which problems faced by the subjects in this study will likely be problems for students several years from now.”

Another shortcoming is the lack of triangulation. The only source of data was the videotapes and the observations made during the videotaping by the investigators. Pre and post interviewing of the subjects and collecting other sources of data, such as students’ written work, would have been helpful in documenting and cross-checking conclusions. Since these data were collected outside the classroom, issues concerning the classroom environment – the pedagogical, affective and physical environment – were not addressed. Neither were gender and cultural differences addressed. These are all certainly important areas for future study.

VII. Future research.

Many of the questions and issues raised by the current and previous study on interactive technology rich learning are relevant to active learning environments in general. Throughout our work on this research project, we found ourselves asking whether a particular instance of behavior was unique to a computer-based learning environment or more relevant to all active learning situations. A next step in this line of research would be to investigate the differences between students working in an active learning environment using only pencil, paper and hand-held calculators and students working in a technology-rich active learning environment.

Because we realize that the limited number of subjects and observations limits our ability to generate these results, another next step would be to develop and test a coding scheme for analyzing the kind of videotapes we've collected of students working together, where the coding scheme and categories grow from the ground up as a theory emerges. This would allow for a more efficient and reliable collection of data so that larger samples could be studied resulting in more replicable and more generalizable research. Verifying the existence of collaborative learning styles, categorizing them, and placing these styles into a larger theory of collaborative learning is a potentially rich area for future research. Researchers will then need to investigate the interaction among collaborative learning styles, the tools available to the student (e.g. computer algebra systems) and prior mathematical experiences and knowledge and how these factors impact learning and affect achievement.

Author Notes

This research was funded by National Science Foundation grant # DUE-9752421. The authors would like to acknowledge the support for this project from the Carnegie Academy for the Scholarship of Teaching and Learning, a program of The Carnegie Foundation for the Advancement of Teaching and the authors would also like to acknowledge the help of Elizabeth Hackney contributed to this paper as an undergraduate research assistant.

References

- Aleven, V., Stahl, E., Schworm, S., Fischer, F., and Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73, 277-320.
- Alexander, P.A. and Murphy, P.K. (1998). The research base for APA's learner-centered psychological principles. In N. Lambert and B. McCombs (Eds.), *How students learn* (pp. 25 - 60). Washington, DC: American Psychological Association.
- Archer, J. (1998). The link to higher scores. *Education Week on the Web*. Retrieved March 15, 2003 from <http://www.edweek.org/sreports/tc98/ets/ets-n.htm>.
- Azmitia, M. and Crowley, K. (2002). Reflections on the social context of scientific understanding. In Crowley, K.; Schunn, C. D.; and Okada, T. (Eds.), *Designing for Science: Implications from Professional, Instructional and Everyday Science*, Mahwah, NJ: Lawrence Erlbaum Associates.

Barron, Brigid (2003). When Smart Groups Fail. *Journal of the Learning Sciences* 12:3, 307-359.

Becker, H and Riel, M (2000). Teacher professional engagement and constructivist-compatible computer use. teaching, learning, and computing: 1998 National Survey. Report #7. Irvine, CA: Center for Research on Information Technology and Organizations.

Bookman, J. and Malone, D. (2003). The Nature of Learning in Interactive Technological Environments - A Proposal for a Research Agenda Based on Grounded Theory, edited by edited by Selden, A., Dubinsky, E., Harel, G., and Hitt, F., Research in Collegiate Mathematics Education V., Providence, RI: American Mathematical Society.

Cooper, J. and Robinson, P. (1998). Small group instruction in science, mathematics, and engineering, and technology: A discipline status report and a teaching agenda for the future. *Journal of College Science Teaching*, 27 (6), 383-8.

Cordes, C. (1998, January 16). As educators rush to embrace technology: A coterie of skeptics seeks to be heard. The Chronicle of Higher Education, p.A25.

Cuban, L. (1986). Teachers and machines: The Classroom Use of Technology Since 1920. New York: Teacher College Press.

Cuban, L. (2001). Oversold and underused: Computers in the classroom. Cambridge: Harvard University Press.

Dubinsky, E., Mathews, D., and Reynolds, B. (Eds.). (1997). Readings in Cooperative Learning for Undergraduate Mathematics. Washington, D.C.: The Mathematical Association of America.

Dubinsky, E., and Schwingendorf, K. (2001). Calculus, Concepts, Computers, and Cooperative Learning (C4L): The Purdue Calculus Reform Project-Constructivist Learning Theory in Practice. Retrieved March 11, 2003 from <http://www.math.purdue.edu/~ccc/>.

Dunn, R., Dunn, K. and Price, G. (1984). Learning Style Inventory. Lawrence, KS: Price Systems.

Edelson, D., Pea, R., and Gomez, L. (1996). Constructivism in the Collaboratory. In B.G. Wilson (Ed.), Constructivist learning environments: Case studies in instructional designs. (pp. 151-164). Englewood Cliffs, NJ: Educational Technology Publications.

Ellis, W. (2000). Technology and calculus. In S. Ganter (Ed.), Calculus renewal: Issues for undergraduate mathematics education in the next decade. (pp. 7-18). Reston, V.A.: Monograph 4 of the National Council of Teachers of Mathematics.

Farrell, A. (1996). Roles and behaviors in technology-integrated pre-calculus classrooms. *Journal of Mathematical Behavior*, 15, 35-53.

Forman, E., and Ansell, E. (2002). Orchestrating the multiple voices and inscriptions of a mathematics classroom. In A. Sfard, and K. McClain (Guest editors). Perspectives on the role of designed artifacts in mathematics learning. *The Journal of the Learning Sciences*, 11 (2and3), 251-274.

Gall, M.D., Borg, W.R., and Gall, J.P. (1996). Educational Research: An Introduction. (Sixth Ed.). White Plains, N.Y.: Longman.

Glaser, B.G. and Strauss, A.L. (1967). The discovery of grounded theory. Chicago: Aldine Publishing.

Goos, A., Galbraith, P., Renshaw, P., and Geiger, V. (2003). Perspectives on technology mediated learning in secondary school mathematics classrooms. *The Journal of Mathematical Behavior*, 22, 73-89.

Inkpen, K., McGrenere, J., Booth, K.S., and Klawe, M. (1997). Turn-Taking Protocols for Mouse-Driven Collaborative Environments. Proceedings of Graphics Interface '97, Kelowna, BC, May 1997, pp. 138-145.

http://www.cs.sfu.ca/people/Faculty/inkpen/Papers/GI97/gi_97.html

Kelman, P., Bardige, A., Choate, J., Hanify, G., Richards, J., Roberts, N., Walters, J., and Tornrose, M. (1983). Computers in teaching mathematics. Reading, MA: Addison-Wesley.

McCallum, W. (2000). The goals of the calculus course. In S. Ganter (Ed.) Calculus renewal: Issues for undergraduate mathematics education in the next decade (pp. 11-22). New York, NY: Plenum Press.

National Research Council. (1991). Moving beyond myths. Washington, DC: National Academy Press.

Noddings, N. (1990). Constructivism in mathematics education. In R. Davis, C. Maher., and N. Noddings (Eds.), Constructivist views on the teaching and learning of mathematics (pp. 7-18). Reston, V.A.: Monograph 4 of the National Council of Teachers of Mathematics.

Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. New York: Basic Books.

Pea, R.D. (1987). Cognitive technologies for mathematics education. In A. Schoenfeld (Ed.), Cognitive science and mathematics education (pp. 89-122). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reiser, R.A. (2001). A history of instructional design and technology part one: A history of instructional media. *Educational Technology Research and Development*, 49, 53-64.

Resnick, L.B. (1991). Shared cognition. In L.B. Resnick, J.M. Levine, and S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp.1-20). Washington, DC: American Psychological Association.

Rogers, E., Reynolds, B., Davidson, N., and Thomas, A. (2001). *Cooperative learning in undergraduate mathematics: Issues that matter and strategies that work*. The Mathematical Association of America.

Rogoff, Barbara, Turkkanis, C.G. and Bartlett, L. (2001). *Learning Together: Children and Adults in a School Community* New York: Oxford University Press.

Romberg, T.A. (1992). Perspectives on scholarship and research methods. In D.A.Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp.49-64). New York: Macmillan.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, 2(3), 235-276.

<http://ctl.sri.com/publications/downloads/ConvergentConceptual.pdf>

Roschelle, J., Pea, R., Hoadley, C., Gordin, D., and Means, B. (2001); "Changing how and what children learn in school with collaborative cognitive technologies," in M. Shields (Ed.), *The Future of Children (Issue on Children and Computer Technology, Los Altos, CA)*, David and Lucille Packard Foundation, 10(2), 76-101.

Roschelle, J., Tatar, D., and Kaput, J. (2005). Getting to scale with innovations that deeply restructure how students come to know mathematics. Unpublished draft manuscript at

Ross, C. M., and Lukow, J. 2004. "Are learning styles a good predictor for integrating instructional technology into a leisure studies curriculum?" *Journal of the Scholarship of Teaching and Learning*, 42, 1, 42-53.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. New York: Academic Press.

Shechtman, N., Roschelle, J., Haertel, G., Knudsen, J., and Tatar, D. (2005). Measuring student learning gains in conceptual mathematics when scaling a technological intervention for middle school mathematics. Paper presented at the American Educational Research Association Annual Meeting, Montreal, Canada.

http://ctl.sri.com/publications/downloads/SimCalc_Assessment_AERA_2005.pdf

Springer, L., Stanne, M., and Donovan, S. (1999). Effects of small group learning on undergraduates in science mathematics engineering and technology: A meta-analysis. *Review of Educational Research*, 69, 21-51.

White, B. and Frederiksen, J. ((2005). A theoretical framework and approach for fostering metacognitive development. *Educational Psychologist* Volume 40, Number 4, Fall 2005, 211-224.

Woolfolk, A. (1995). *Educational psychology*. Boston: Allyn and Bacon.

Bookman, J. and Malone, D.

Van Zee, E.H. (2000). Analysis of a student generated inquiry discussion. ***International Journal of Science Education***, 22, 115-142.

Vygotsky, L.S. (1978). *Mind in society: The development of mental process*. Cambridge, MA: Harvard University Press.

Zemelman, S., Daniels, H., and Hyde, A. (1998). *Best practice: New standards for teaching and learning in America's schools*. Portsmouth, NH: Heinemann.

Reciprocal Teaching of Lecture Comprehension Skills in College Students

Norman R. Spivey¹ and Andrea Cuthbert²

Abstract: This study explored the effects of a reciprocal teaching intervention designed to enhance the lecture comprehension skills of college students. Forty low-verbal ability students and 40 high-verbal ability students (as measured by SAT scores) were chosen for the study and randomly assigned to experimental or control groups. The experimental groups received the reciprocal teaching intervention, whereas the control groups did not. The instructor modeled four listening comprehension activities: summarizing, self-questioning, clarifying, and predicting, and guided the students in performing the comprehension activities. Six lectures were presented to all subjects in both experimental and control groups. Lectures were followed by comprehension tests. Results showed that low-verbal ability subjects receiving the reciprocal teaching method significantly increased their lecture comprehension. These significant increases were maintained over time.

I. Introduction.

Skilled reading, attentive listening, effective studying, and other scholastic activities are complex processes that involve skill and ability. Students who do not possess these skills are at a disadvantage in any learning situation.

Listening comprehension is one of these important skills, and it must be learned if students are going to be successful in school (Senechal and LeFevre, 2002). It is an important skill that affects people's daily lives, and proficiency in listening, and listening comprehension is imperative (Petress, 1999). Students in school settings who receive the majority of their information through lectures and discussion may benefit from research identifying effective strategies that improve listening comprehension. According to Hoover and Gough (1990), listening comprehension is even necessary for reading success. They suggest that impairments in listening comprehension can also limit reading. A context where listening comprehension is of vital importance is in the college classroom. Comprehension of lectures is paramount if college students are going to be successful in the classroom. When listening comprehension skills are highly developed, they occur naturally. When they are poorly developed, they may not occur at all (Aarnoutse and van den Bos, 1998). Unfortunately, in today's colleges and universities, there are many students who lack listening comprehension skills. If such students are going to meet the expectation that they comprehend complex lecture information, they will have to be taught the necessary listening comprehension skills (Block and Pressley, 2000).

Despite the importance of listening in the classroom, the ability to comprehend what is heard has been given little attention in language arts programs. Some researchers, however, have attempted to study various instructional strategies for listening and listening comprehension.

¹Professor of Psychology, spiveyn@bethelcollege.edu.

²Division Assistant, spiveyn@bethelcollege.edu, both at Bethel College, 1001 W. McKinley Ave., Mishawaka, IN 46545.

Cunningham, Cunningham, and Arthur (1981) studied listening instruction and concluded that the use of a Directed Listening Activity would aid in listening and listening comprehension. In a similar study using directed instruction to enhance listening comprehension, Donahue and Pidek (1993) tested an oral paraphrasing strategy. Results show increases in the listening comprehension skills of students with language/learning disabilities. Some researchers (Funk and Funk, 1989; Mandlebaum and Wilson, 1989; Winkle, 1991) studied listening instruction and concluded that an important factor in listening training is integration. Listening cannot be taught effectively as an isolated subject; listening activities should be included in all areas of the curriculum. Other researchers (Riggenbach, 1990; White, 1990) believe that instruction in listening must include metacognitive skills, such as monitoring listening, analyzing language data, and being conscious of language production skills. One such strategy that includes comprehension and comprehension monitoring is reciprocal teaching (Palincsar and Brown, 1984; Palincsar, 1987; Palincsar, 1991). This is a method of cooperative teaching where teacher and pupils take turns in leading a discussion about a text listened to by students. Also, students are taught four strategies—questioning, clarifying, summarizing, and predicting. According to Palincsar and Brown (1984), these four strategies serve the two important functions of comprehension fostering and comprehension monitoring. Palincsar and Brown have found this method particularly effective in the introduction of new skills and concepts. As students gain in mastery they also receive plenty of opportunity for self-directed practice with feedback from peers and teachers. This type of instruction has been successful in a variety of contexts. Because adult learners spend much of their class time listening, and some adult learners can show deficiencies in their listening comprehension, this type of intervention may be particularly effective in not only improving students' comprehension of class lectures, but also improving students' understanding of peer presentations, improving students' understanding of the content in class discussions, and improving students' basic academic conversations. Reciprocal teaching has been hailed as one of the most prominent strategy-instruction programs developed in the last few decades (Dole, Duffy, Roehler, and Pearson, 1991; Glaser, 1990; Pearson and Dole, 1987; Siegler, 1991; Stanovich and Cunningham, 1991).

Most studies of reciprocal teaching were conducted with children (e.g., Aarnutse, Brand-Gruwel, and Oduber, 1997; Aarnoutse, van den Bos, and Brand-Gruwel, 1998; Coley, DePinto, Craig, and Gardner, 1993; Johnson-Glenberg, 2000; and Marks et al., 1993). Investigations with college students are limited (e.g., Hart and Speece, 1998). Hart and Speece (1998) tested the effects of reciprocal teaching in fostering reading comprehension skills in community college students. Their results showed the reciprocal teaching group performed significantly better than the comparison group on tests of reading comprehension.

The purpose of this study was to explore the effects of a reciprocal teaching intervention designed to enhance the lecture comprehension skills of college students. There are four hypotheses for this study: (a) College students with lower verbal ability who receive reciprocal teaching of comprehension fostering skills will show significantly more improvement in lecture comprehension from pretest to posttest compared to students with lower verbal ability who receive no training; (b) College students with lower verbal ability who are exposed to reciprocal teaching will maintain their improvement in lecture comprehension from posttest to delayed posttest compared to students with lower verbal ability who receive no training; (c) College students with lower verbal ability who receive reciprocal teaching of comprehension fostering skills will show higher levels of improvement in lecture comprehension from pretest to posttest compared with higher verbal ability students who receive reciprocal teaching or no training. The

lower verbal ability students should have more room to improve, because these individuals tend to be deficient in strategy usage; (d) College students with lower verbal ability who receive reciprocal teaching of comprehension monitoring skills will maintain their improvement in lecture comprehension from posttest to delayed posttest compared with students with higher verbal ability who receive reciprocal teaching or no training. Because lower verbal ability students tend to be deficient in strategy usage, they should have more room to improve.

II. Methods.

A. Participants.

Participants used in the study were 80 college students, 34 males and 46 females, enrolled in two sections of introductory psychology classes (typical enrollment for each section is approximately 150 students). The two sections were similar in makeup. Both sections had approximately 85% traditional students and 15% nontraditional students. Ninety-five percent of the students in both sections were freshmen.

The investigation included students with varying verbal abilities in order to test the effectiveness of treatment conditions as a function of ability level. The 80 participants (40 low-verbal ability – LVA and 40 high-verbal ability – HVA) were randomly assigned to one of two conditions (reciprocal teaching or control). Twenty low-verbal ability students were randomly assigned to the experimental condition (LVA-E); the remaining 20 low-verbal ability students were randomly assigned to the control condition (LVA-C). Similarly, 20 of the high-verbal ability students were randomly assigned to the experimental condition (HVA-E); the remaining 20 high-verbal ability students were randomly assigned to the control condition (HVA-C). The control condition in this study was similar to that used by King (1989, 1990). The low-verbal ability experimental (LVA-E) group received all assessment tests plus the reciprocal teaching intervention. The low-verbal ability control (LVA-C) group received all assessment tests but did not receive any intervention. The high-verbal ability experimental (HVA-E) group received all assessment tests plus the reciprocal teaching intervention. The high-verbal ability control (HVA-C) group received all assessment tests but did not receive any intervention.

In order to obtain participants who differed considerably in verbal ability level, only those students whose SAT-Verbal scores fell below 420 and above 580 were asked to participate. All participants were asked to sign a form giving permission for their SAT scores to be released from the Admissions office. Ninety students with SAT verbal scores above 580 were identified and 70 students with SAT verbal scores below 420 were identified from existing records. The high-verbal ability and low-verbal ability cut-off scores used for selection in this study were the same as the high-verbal ability and low-verbal ability scores used by Baker (1985) in her study examining the differences in the standards used by college students to evaluate their comprehension. The participants were paid \$10 to participate and received extra credit points in their psychology class. To help in preventing attrition, the students were told that they would not be paid nor would they receive the extra credit unless they attended all of the sessions. All students who were chosen for the present study attended all sessions. Thirty-five students chose not to volunteer for the experiment, and were allowed to write a brief review of a research article, and received the same extra credit as the students who participated in the experiment. Forty students were randomly selected from the ninety identified students with SAT verbal

scores above 580, and forty students were randomly selected from the 70 identified students with SAT verbal scores below 420.

B. Materials.

Several instruments were used to assess the effectiveness of the lecture comprehension skills intervention. The measures were as follows:

Listening Training Passages. Fifteen listening passages, averaging 800 words in length, were available in all the conditions. In the reciprocal teaching condition, students listened to each paragraph of the passage, and the instructor encouraged the dialogue necessary to instruct the students in the prediction, clarification, self-questioning, and summarizing activities. In the control condition, students listened to the passage and discussed them.

The passages were chosen from the following current college psychology texts: *Theories of Human Learning* (Lefrancois, 2000); *Social Psychology* (Aronson, Wilson, and Akert, 1997); *Life-span Development* (Santrock, 1999); and *Abnormal Psychology and Modern Life* (Carson, Butcher, and Mineka, 1996). The passages were expository and represented a wide range of topics. The passages were selected after ensuring that they conformed to college level according to the Fry Readability Formula (Fry, 1977).

Pretest, Practice, Posttest, and Delayed Posttest Lectures. The students in all four groups (experimental and control) were asked to listen to and watch six videotaped course lectures from a General Psychology course. The lectures took place in the regular classroom under normal class conditions. The lectures were conducted by an associate professor with 14 years teaching experience in introductory psychology classes. The lectures were video-taped to allow each section of the introductory psychology classes to observe and listen to the exact same lecture. The duration for each lecture was one-half hour. Each lecture was followed by a comprehension test to gather lecture comprehension data. The first lecture was followed by a comprehension pretest to gather baseline data. The next three lectures, given during the intervention phase, were followed by comprehension tests to gather ongoing data on the reciprocal teaching versus control conditions. The fifth lecture was followed by a comprehension test to gather lecture comprehension data for posttest. The sixth lecture was followed by a comprehension test to gather lecture comprehension data for the delayed posttest. The criteria for selection of these lecture presentations were that they were unrelated to material covered in other classes to avoid effects of prior knowledge at testing, and that they represented a pure lecture format (no discussion groups, activity groups, movies, etc.).

The lectures followed the order of topics in *Essentials of Psychology: Exploration and Application* (Coon, 2000), which was the textbook used for the course. However, the material covered in the lectures was supplemental and did not come directly from Coon's text (to avoid effects of prior reading of material in the text). The lectures covered the following regular course topics: the control of pain, observational learning, creativity, birth order, mental health, and motivational cycle. Each lecture followed a logical outline form with main topics supported by details and examples. The goal of the lecture was to organize the information and present it in a step-by-step fashion, allowing for easy assimilation into existing knowledge schemes. To insure that listening was the only method used by the students to take in the information, no terms were written on the board.

Pretest, Practice, Posttest and Delayed Posttest Lecture Comprehension Tests. Following each lecture, a written test was administered to evaluate participants' comprehension of the content of the lecture. The written tests covered only material presented in the lecture and

consisted of 14 questions, 10 multiple-choice (five text explicit and five text implicit) and four short-answer essay (two text explicit and two text implicit). To aid in classifying the questions as text explicit or text implicit, Pearson and Johnson's (1978) taxonomy was used. The multiple-choice questions were each worth one point and the short-answer questions were each worth two and a half points, making the entire test worth 20 points. The entire Pretest Lecture Comprehension Test is found in Appendix A.

Two independent raters (both taught written communications classes) were asked to use Pearson and Johnson's (1978) taxonomy to classify the 10 multiple-choice questions and the four essay questions for each lecture. Using Pearson and Johnson's taxonomy, the raters classified each question as either text implicit or text explicit. Using the lecture as an indicator of content covered, the two independent raters, neither knowing the identity of the student nor the conditions of the experiment, scored the multiple-choice questions and the open-ended questions on each of the six tests. A scoring key of the expected answers for each multiple-choice question and each essay question was prepared for the two raters. The key contained the answers to each question, and the amount of credit allotted to each question. For each essay question, each of the participants' answers were compared to the ideal answer in the scoring key and a given number of points were assigned in terms of the adequacy of the answer. To clarify the coding scheme used for the essay questions, the following is an example of an essay question used in the present study. Two participants' answers are also provided, one receiving full credit and one receiving partial credit.

Question: Kagan noted two contrasting styles of thinking exhibited by children. Name and describe each style.

Answer (full credit, 2.5 points): Impulsive thinkers: tend to look at problems in a global way and offer quick responses. Reflective thinkers: concentrate on the details of the problem and are more deliberate (slower) in offering responses.

Answer (partial credit, 1.0 points): Impulsive: Come up with first answer they think of. Responsive: Respond with an answer.

After all the questions on the test had been graded, the points for the multiple-choice question were added together with the points for the essay questions, and a composite score was computed. The scores on the multiple-choice questions and essay questions were analyzed separately, and were found to be significantly correlated, $r(79) = 0.61$, $p < 0.05$. Therefore, a composite score was used in the present study. The internal consistency of each test was examined using the Kuder-Richardson Formula 20 (KR-20). The internal consistency of the multiple-choice portion of the six lecture comprehension tests ranged from 0.70 to 0.77. Interrater reliabilities on the tests were also examined. Agreement between the raters was 85 percent. On the lecture comprehension tests, the range of actual scores was: HVA-E, 15-20; LVA-E, 12-18; HVA-C, 15-20; LVA-C, 8-16.

C. Procedure.

The steps in the procedure for the control and experimental conditions were as follows:

Low Verbal Ability Control Group. Twenty low verbal ability students (LVA-C) were assigned to four groups; each group consisted of five students who worked with an instructor. The instructor was a senior undergraduate student majoring in Psychology. The instructor was trained by the experimenter to give summaries, lead discussions, and answer questions. Each group listened to a training passage, then the instructor summarized the passage for the students,

led a discussion about the passage, and answered any questions that the students had about the passage. This type of control condition was similar to that used by King (1989, 1990) in assessing comprehension of lecture material. This control condition provided activities for the students which allowed for the control groups to be treated similarly to the experimental groups (they met with an instructor, they discussed the listening passages, they met for the same amount of time each day, etc.). The only difference was that students in the control condition did not receive the reciprocal teaching intervention. These control students also listened to lectures and answered pretest, practice, posttest and delayed posttest lecture comprehension tests.

High Verbal Ability Control Group. Twenty high verbal ability students (HVA-C) received no intervention but listened to the training passages and discussed the information in the passages. The students also listened to lectures and answered pretest, practice, posttest and delayed posttest lecture comprehension tests. Procedures used with Control Group 2 were identical in all respects to procedures used with Control Group 1.

Low Verbal Ability Experimental Group. Twenty low verbal ability students (LVA-E) were assigned to four groups; each group consisted of five students who worked with a reciprocal teaching instructor in the experimental condition. Each group listened to a training passage, then the instructor directed the dialogues and instructed the students in the use of the four strategies. The instructors received three training sessions (Palincsar and Brown, 1984).

High Verbal Ability Experimental Group. Twenty high verbal ability students (HVA-E) were assigned to four groups; each group consisted of five students who worked with a reciprocal teaching instructor. The students received the same tests and intervention as the students in the low verbal ability experimental group, and the procedures used with the high verbal ability experimental group were identical in all respects to procedures used with the low verbal ability experimental group.

For the control groups and experimental groups, there were four phases to the study: (a) pretest, consisting of a lecture and a comprehension test; (b) treatment, consisting of 15 days of training--half an hour per day using the reciprocal teaching format for the experimental group or the discussion format for the control group, plus three lectures and three comprehension tests; (c) posttest, consisting of a lecture and comprehension test at the termination of the training phase; (d) delayed posttest, consisting of a lecture and comprehension test eight weeks after termination of the posttest phase. The students were apprised of their progress on the lecture comprehension tests. They were shown graphs depicting the percentage correct for each lecture comprehension test.

Reciprocal Teaching. A procedure developed by Palincsar and Brown (1984), reciprocal teaching, was used in the present intervention study. A reciprocal teaching instructor, grouped with either five low verbal ability college students (LVA-E), or five high verbal ability college students (HVA-E), first presented an overview of four strategies designed to enhance comprehension and comprehension monitoring: summarizing, questioning, clarifying and predicting. For the remainder of the intervention phase, the instructor and students listened to tape recorded passages and entered into dialogues pertaining to the passages to which they listened. The instructor and students took turns leading the group, and the instructor gradually shifted responsibility to the students as they gained expertise. The main objective of the four comprehension strategies was for students to understand the passages and remember them. All of the activities were embedded in as natural a dialogue as possible, with students and instructor giving feedback to each other within the context of actually listening to the passages. The students were encouraged to summarize the content of what they heard and then ask a question

about the main idea of the passage. Students were also encouraged to ask questions for clarification if anything in the passage was not understandable, as well as to make predictions about future events from the content of the listening passage.

After listening to a training passage, the following dialogue occurred between instructor (I) and student (S) early in the training program of the present study (first week). Part of the passage the students listened to was as follows:

The traditional view of abnormal psychology has been based on the assumption that a fixed set of mental disorders exists, whose obvious manifestations cut across cultures. This psychiatric tradition dates back to Emil Kraepelin, who felt that depression, sociopathic behavior, and especially schizophrenia were universal disorders that appeared in all cultures and societies. Early research supported the belief that these disorders occurred worldwide, had similar processes, and were more similar than dissimilar. Such cultural universality has led to the belief that a disorder such as depression would be similar in origin, process, and manifestation in Asian, Black, Hispanic, or White clients. As a result, no modifications in diagnosis and treatment need to be made. Western concepts of normality and abnormality could be considered universal and equally applicable across cultures (Sue, Sue, and Sue, 1994, p. 9).

I: Can anyone summarize this passage?

S: Abnormal behaviors are the same around the world?

I: Good job. And how were these abnormal behaviors similar?

S: They occurred the same way everywhere?

I: O.K. You're right. But, the disorders were also similar in other ways, like their origin.

S: And their processes?

I: Right! Anything else?

S: I can't remember.

I: O.K. Let me try to do a summary for you. The most important thing about this passage is that according to some researchers, abnormal behavior is the same from one culture to the next. The disorders have the same origin, processes, and manifestation no matter what culture they appear in.

After listening to another training passage, the following dialogue occurred between instructor (I) and student (S) later in the training program of the present study (third week). Part of the passage the students listened to was as follows:

Why do people overuse drugs? The answer to this question is complicated by the number of different kinds of drugs that are used and the number of factors that interact to account for the use of any drug. An explanation of drug abuse must take into account several general observations. First, in the 1960s and 1970s some researchers had hoped that they could identify a cluster of personality traits that could account for addiction to substances. However, simple attempts to find a common pattern of personality traits that underlie addiction have failed. It is highly unlikely that addiction is caused by a single personality type (Sue, Sue, and Sue, 1994, p. 281).

I: Can you summarize the paragraph?

S: Sure. At first researchers thought that drug abuse was connected to the personality. Like a personality trait for drug use. But then they didn't find any personality traits connected to drug use. Drug addiction probably isn't caused by personality traits.

I: Excellent Job. Now predict what you think will come next in the passage.

S: Another explanation for why people abuse drugs. Maybe something to do with peer pressure or other types of pressure like stress.

I: You're right. The third paragraph deals with that!

In the treatment phase, the instructor introduced the passage with a brief discussion. Since the passage was new to the students, the title was called to their attention and they were asked to predict the content of the passage based on the title. The instructor then indicated which group member would teach the first segment of the passage (usually one paragraph); the instructor or one of the five students. The instructor led the first few dialogues to model the appropriate techniques for utilizing the four comprehension strategies. After listening to the segment, the leader (student or instructor) for that segment asked a question like those that might be on a test of the material. The question was generated on the spot based on the material presented in the segment. The instructor or student leader then summarized the segment, discussed and clarified any difficulties, and finally made a prediction about future content. The reciprocal teaching instructor provided the guidance necessary for the students to complete the preceding activities by using a number of techniques: prompting (asked a question to generate dialogue), instructing (corrected student if something was stated incorrectly, gave information when it was needed) and modifying the activity (had student summarize if they were having trouble formulating a question). Throughout the daily reciprocal teaching intervention, segment to segment and paragraph to paragraph, the students were told that these activities were strategies that would help them understand better as they listened, and that they should try to use the strategies every day.

III. Results.

To assess effects of the reciprocal teaching strategy on the two verbal ability groups over time, a 2 (verbal ability) x 2 (treatment group) x 6 (time) repeated measures Analysis of Variance was conducted on the participants' scores on the six lecture comprehension tests.

The main effect of group was significant, $F(1,76) = 35.65, p < 0.01$. The groups receiving reciprocal teaching scored higher ($M=16.57, SD=1.18$) on the lecture comprehension tests as compared to the control group receiving no reciprocal teaching ($M=15.14, SD=1.56$). The main effect of verbal ability was also significant, $F(1,76) = 252.53, p < 0.01$. The higher verbal ability group scored higher ($M=17.64, SD=1.22$) than the lower verbal ability group ($M=14.30, SD=2.50$). These effects were qualified in that the group x verbal ability interaction was significant, $F(1,76) = 32.70, p < 0.01$. Table 1 presents the lecture comprehension scores of the experimental group and the control group as a function of verbal ability. In order to simplify the interpretations of interactions, a test of simple main effects indicated that high verbal ability students outperformed low verbal ability students on the lecture comprehension tests in the experimental condition, $F(1,76) = 4.03, p < 0.05$, as well as in the control condition, $F(1,76) = 38.12, p < 0.01$, however, the differences between the two ability groups were largest in the control condition.

Table 1. Mean and Standard Deviation Scores for Lecture Comprehension as a Function of Group and Verbal Ability.

	High Verbal		Low Verbal	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Reciprocal Teaching Group	17.68	1.98	15.94	1.99
Control Group	17.53	2.05	13.31	2.44

The effect of time was significant, $F(5,380) = 16.97, p < 0.01$. A Newman-Keuls test of multiple comparisons revealed significant differences ($p < 0.05$) between means (Test #1 $M=14.83 < \text{Test #6 } M=15.92 < \text{Test #2 } M=16.03 = \text{Test #3 } M=16.03 < \text{Test #4 } M=16.20 < \text{Test #5 } M=16.56$) indicating that students increased their scores on the lecture comprehension tasks from Test #1 through Test #5, then achieved lower scores on the delayed Test #6, perhaps because there were practice effects. This effect was qualified in that the group x time interaction was significant, $F(5,380) = 9.36, p < 0.01$. Table 2 presents the lecture comprehension scores of the experimental group and the control group as a function of time.

Table 2. Mean and Standard Deviation Scores for Lecture Comprehension as a Function of Group and Time.

		Reciprocal Teaching		Control	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test1	(Pretest)	14.87	1.89	15.07	2.38
Test 2		16.69	1.63	15.60	2.20
Test 3		17.05	1.79	15.35	2.40
Test 4		17.13	2.20	15.35	2.33
Test 5	(Post test)	17.72	2.26	15.85	2.20
Test 6	(Delayed Post test)	16.95	2.14	15.30	1.98

A test of simple main effects indicated that the group the subject belonged to (reciprocal teaching or control) significantly affected test performance over time. Significant increases in comprehension were found in the reciprocal teaching condition, $F(1,380) = 19.33, p < 0.01$. A Newman-Keuls test of multiple comparison revealed that significant differences ($p < 0.05$) between means existed (Test #1 $M=14.87 < \text{Test #2 } M=16.69 < \text{Test #6 } M=16.95 < \text{Test #3 } M=17.05 < \text{Test #4 } M=17.13 < \text{Test #5 } M=17.72$) indicating that students in the reciprocal teaching condition increased their scores on the lecture comprehension tasks from test #1 through Test #5 and then achieved lower scores on the delayed Test #6, implying that the intervention led to a sizable impact, which was only partially maintained over time. No significant differences were found in the control condition, $p > 0.05$.

A group x verbal ability x time interaction was also significant, $F(5,380) = 11.11, p < 0.01$. The means and standard deviations related to lecture comprehension as a function of group (experimental vs. control), verbal ability and time (tests 1 - 6) are found in Table 3.

Table 3. Mean and Standard Deviation Scores for Lecture Comprehension as a Function of Group, Verbal Ability, and Time.

		Reciprocal Teaching Group			
		High Verbal		Low Verbal	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test1	(Pretest)	16.90	2.17	12.90	1.62
Test 2		17.80	1.22	16.45	2.03
Test 3		18.10	1.65	16.20	1.94
Test 4		17.70	2.30	16.40	2.11
Test 5	(Post test)	18.00	2.36	17.40	2.17
Test 6	(Delayed Post test)	17.60	2.20	16.30	2.08

		Control Group			
		High Verbal		Low Verbal	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Test 1	(Pretest)	16.70	2.01	13.45	2.75
Test 2		17.50	2.06	13.70	2.35
Test 3		17.70	2.14	13.00	2.67
Test 4		17.50	2.11	13.20	2.55
Test 5	(Post test)	18.30	1.91	13.40	2.49
Test 6	(Delayed post test)	17.50	2.10	13.10	1.86

A Newman-Keuls test of multiple comparison ($p < 0.05$) indicated that there were significant differences between high verbal ability students and low verbal ability students in both the experimental, and control conditions during the first lecture comprehension test. However, for the experimental condition, tests two through six showed no significant differences between high and low verbal ability students in the experimental condition. This indicated that the low verbal ability students who received the reciprocal teaching intervention improved their scores to the level of the high verbal ability students, and this improvement remained stable from posttest to delayed posttest. The data also indicated that the high verbal ability students who received the reciprocal teaching intervention did not significantly increase their lecture comprehension scores. For the control groups, significant differences between the high verbal ability students and low verbal ability students continued throughout the six testing situations. This indicated that low verbal ability students who did not receive a reciprocal teaching intervention continued to perform poorly on the listening comprehension task, and those high verbal ability students not receiving a reciprocal teaching intervention continued to perform at a high level on all six tests.

IV. Discussion.

A. Lecture Comprehension.

The purpose of this study was to investigate the effects of a reciprocal teaching intervention on lecture comprehension. The findings yielded support for Palincsar and Brown's (1984) research on the effects of reciprocal teaching. Those low verbal ability college students who experienced the reciprocal teaching intervention improved their listening comprehension scores from pretest to posttest to delayed posttest compared to low verbal ability students who received no training. It appears that the experience of the reciprocal teaching intervention increased low verbal ability students' ability to understand information presented orally. This understanding was reflected in their increased comprehension scores. The scores improved rapidly, showing improvement by the end of the first week of treatment. The improved lecture comprehension was also reflected in the stability of scores from posttest to delayed posttest. Those low verbal ability students in the control condition (not receiving a reciprocal teaching intervention) continued to perform poorly from pretest to delayed posttest.

Another possible explanation for the sudden improvement of the low verbal ability students in the reciprocal teaching groups is accountability. The students in the reciprocal teaching groups were held accountable for leading the groups. The students felt it was important to lead the group with competence and to gain the respect of their fellow group members. The students appeared to take this responsibility very seriously. The feelings of accountability may have caused students to work harder in the groups, thus improving their lecture comprehension scores. Reinforcement may also have contributed to the improved scores of the low verbal ability students in the reciprocal teaching groups. The students appeared to respond well to the feedback and reinforcement they received in their groups. They looked forward to interacting with the group and felt that the group had a positive influence on them personally. The reinforcement the students received in the group may have motivated them to improve their performance on the lecture comprehension tests.

Another explanation for the increased lecture comprehension scores of the low verbal ability students in the reciprocal teaching groups may have been the communication that took place between students in all of the groups. It was possible that students talked to each other and compared their experiences. Those in the reciprocal teaching groups received more attention (dialogue within the group, responsibility of leading the group, etc.) than students in the control conditions. This knowledge of group differences (they were receiving special attention) may have caused students in the reciprocal teaching group to try harder to please the instructor. An instructor effect may be another explanation for the improved lecture comprehension scores of the low verbal ability students in the reciprocal teaching group. The reciprocal teaching instructor for the low verbal ability students may have been more effective or enthusiastic than teachers in the other conditions. The instructor for the low verbal ability students was excited about the position and was motivated to do well. The instructor's enthusiasm may have positively affected the low verbal ability students and caused them to perform at a higher level than they would have with a less enthusiastic teacher.

High verbal ability students in both the reciprocal teaching group and the control group continued to perform at high levels from pretest to posttest to delayed posttest. This would indicate that the reciprocal teaching intervention did not significantly affect the lecture comprehension scores of students already high in verbal ability. Lower verbal ability students

had more room to improve and the reciprocal teaching intervention did in fact improve their scores. Higher verbal ability students did not have as much room to improve because they were already scoring at higher levels in their lecture comprehension. The lack of improvement in scores could be due to a ceiling effect. The lack of improvement could also be because the passages needed to be more difficult. The range of scores for the HVA experimental group was 15-20, meaning the lowest score was 15 points out of a possible 20 points, and the highest score was 20 points out of a possible 20 points. This would suggest that there was little room for improvement and if the questions were more difficult, or if the passages were more difficult, the range may have been wider.

B. Implications.

A major implication of this study is that a reciprocal teaching intervention does appear to be successful in fostering listening comprehension abilities in low verbal ability college students. The reciprocal teaching intervention provided a model of what expert listeners do when they are trying to understand and remember information. The intervention allowed students to observe the comprehension fostering activities in which they were expected to engage. The intervention also provided appropriate feedback for the student. The instructor was able to gauge the students' abilities and provide information to increase their level of competency. Through the interactions with the instructor and the other students, low verbal ability students were able to increase their listening comprehension abilities, and to maintain those increases over time.

C. Suggestions for Further Research.

The reciprocal teaching intervention appeared successful in fostering lecture comprehension. However, more research is necessary to identify the specific components responsible for the improvement in lecture comprehension. Also, more research is necessary to understand the effects of a reciprocal teaching intervention on high verbal ability students. The difficulty of the comprehension tasks and testing procedures need to be investigated for high verbal ability students. It is also important to investigate if higher verbal ability students are already using similar comprehension strategies. It would seem important to investigate variations or changes in this type of teaching intervention and assess what effects they would have on a more average or above average population.

Further research should also be conducted investigating the effects of a reciprocal teaching intervention on the comprehension abilities of adults at different ages. Assessing age differences in adults' abilities to comprehend auditory information is important. It is also important to assess the effectiveness of a reciprocal teaching intervention for subjects across the adult life-span. Given the high proportion of older "nontraditional" students enrolling in colleges and universities (Aslanian and Brickell, 1980), research addressing this issue is critical.

It is also important to investigate the effects of a reciprocal teaching intervention on comprehension abilities of students in an average college classroom. If this training procedure is to be practical, it must be tested in a normal, average classroom. Because some students at the college level struggle with their understanding of information, and reciprocal teaching could increase their comprehension abilities, it is necessary to investigate the strengths or weaknesses of this intervention in the college classroom with college teachers. Further research should also be conducted to assess if peers (other college students) could take on the role of teachers in the

reciprocal teaching procedure. The question of whether peer tutoring could effectively promote comprehension skills needs to be investigated.

D. Conclusions.

The message of programs such as reciprocal teaching is that knowledge of the learning process and the conditions that affect it should be a major part of the curriculum in elementary schools, secondary schools, and colleges. Reciprocal teaching is a highly effective method for teaching metacognitive reading and listening skills. As students master these metacognitive reading and listening skills, their reading comprehension and listening comprehension improve. It seems that instructional strategies focusing on these skills should be implemented in the classroom.

References

- Aarnoutse, C., Brand-Gruwel, S., and Oduber, R. (1997). "Improving reading comprehension through listening." *Educational Studies*, 23(2), 209-227.
- Aarnoutse, C., and van den Bos, K. (1998). "Effects of listening comprehension training on listening and reading." *Journal of Special Education*, 32, 115-127.
- Aarnoutse, C. van den Bos, K., and Brand-Gruwel, S. (1998). "Effects of listening comprehension training on listening and reading." *Journal of Special Education*. 32(2), 115-126.
- Afflerbach, P., and Walker, B. (1992). "Main idea instruction: An analysis of three basal Reader series." *Reading Research and Instruction*, 32(1), 11-28.
- Aronson, E., Wilson, T., and Akert, R. (1997). *Social Psychology* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Aslanian, C., and Brickell, H. (1980). *Americans in transition: Life changes as reasons for adult learning*. Princeton, NJ: College Board Pub.
- Belajthy, E. (1986). "The relationship of training in self-generated questioning with passage difficulty and immediate and delayed retention." In J. A. Niles and R.V. Lalik (Eds.), *Solving problems in literacy: Learners, teachers and researchers. Thirty-fifth year-book of the National Reading Conference*. Rochester, NY: National Reading Conference.
- Brozo, W. G., Stahl, N. A., and Gordon, B. (1985). "Training effects of summarization, item writing, and knowledge of information sources on reading test performance." In A. Niles and R. V. Lalik (Eds.), *Issues in literacy: A research perspective* (pp. 48-54) Rochester, NY: National Reading Conference.

Coley, J., DePinto, T., Craig, S., and Gardner, R. (1993). "From college to classroom: Three teachers' accounts of their adaptations of reciprocal teaching." *The Elementary School Journal*, 94, 255-266.

Coon, D. (2000). *Essentials of psychology: Exploration and application* (10th ed.). Belmont, CA: Wadsworth Publishing Company.

Cunningham, J. W., Cunningham, P. M., and Arthur, S. V. (1981). *Middle and secondary school reading*. New York, NY: Longman.

Donahue, M. and Piked, C. (1993). "Listening comprehension and paraphrasing in content area classrooms." *Journal of Childhood Communication Disorders*, 14, 35-42.

Dole, J., Duffy, G., Roehler, L, and Pearson, P. (1991). "Moving from the old to the new: Research on reading comprehension instruction." *Review of Educational Research*, 61(2), 239-264.

Fry, E. (1977). Fry's readability graph: "Clarification, validity, and extension to level 17." *Journal of Reading*, 21, 242-252.

Funk, H., and Funk, G. (1989). "Guidelines for developing listening skills." *Reading Teacher*, 42, 660-663.

Glaser, R. (1990). "The reemergence of learning theory within instructional research." *American Psychologist*, 45, 29-39.

Hart, E., and Speece, D. (1998). "Reciprocal teaching goes to college: Effects for Postsecondary students at risk for academic failure." *Journal of Educational Psychology*, 90(4), 670-681.

Hoover, W., and Gough, P. (1990). "The simple view of reading." *Reading and Writing*, 2, 127-160.

Johnson-Glenberg, M. (2000). "Training reading comprehension in adequate decoders/poor comprehenders: Verbal versus visual strategies." *Journal of Educational Psychology*, 92(4), 772-782.

King, A. (1989). "Effects of self-questioning training on college students' comprehension of lectures." *Contemporary Educational Psychology*, 14, 366-381.

King, A. (1990). "Enhancing peer interaction and learning in the classroom through reciprocal questioning." *American Educational Research Journal*, 27, 664-687.

Klinger, K., Vaughn, S., and Schumm, J. (1998). "Collaborative strategic reading during Social studies in heterogeneous fourth-grade classrooms." *Elementary School Journal*, 99(1), 3-22.

Lefrancois, G. (2000). *Theories of Human Learning* (4th ed.). Belmont, CA: Wadsworth.

Mandlebaum, L., and Wilson, R. (1989). "Teaching listening skills in the special education classroom." *Academic Therapy*, 24, 449-459.

Palincsar, A. S. (1987). Interactive cognition to promote listening comprehension. Paper presented to the annual meeting of the American Educational Research Association, San Francisco, CA.

Palincsar, A. S. (1991). "Scaffolded instruction of listening comprehension with first graders at risk for academic difficulty." In A. McKeough and J. Lupart (Eds.) *Toward the practice of theory-based instruction: Current cognitive theories and their educational promise* (pp. 50-65), NJ: Lawrence Erlbaum Assoc.

Palincsar, A. S., and Brown, A. L. (1984). "Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities." *Cognition and Instruction*, 1, 117-175.

Pearson, P., and Dole, J. (1987). "Explicit comprehension instruction: A review of research and a new conceptualization of instruction." *Elementary School Journal*, 88, 151-165.

Pearson, P. D., and Johnson, D. D. (1978). *Teaching reading comprehension*. New York, NY: Holt, Rinehart and Winston.

Petress, K. (1999). "Listening: A vital skill." *Journal of Instructional Psychology*, 26, 261-262.

Pressley, M. (2002). "Comprehension strategies instruction." In C. Block and M. Pressley (Eds.), *Comprehension Instruction: Research-based best practices* (pp. 11-27), New York: Guilford Press.

Riggenbach, H. (1990). "Discourse analysis and spoken language instruction." *Annual Review of Applied Linguistics*, 11, 152-163.

Santrock, J. (1999). *Life-Span Development*. New York, NY: McGraw-Hill.

Senechal, M., and LeFever, J. (2002). "Parental involvement in the development of children's reading skills: A five-year longitudinal study." *Child Development*, 73, 445-461.

Stanovich, K., and Cunningham, A. (1991). "Reading as constrained reasoning." In R.J. Sternberg and P. A. Frensch (Eds.), *Complex Problem solving: Principles and mechanisms* (pp. 3-60). Hillsdale, NJ: Erlbaum.

Sue, D., Sue, D., and Sue, S. (1994). *Understanding abnormal behavior*. Boston, MA: Houghton Mifflin Company.

White, K. (1990). "Redefining basic literacy through speaking and listening instruction." *Adult learning*, 1, 23-25.

Spivey, N.R. and Cuthbert, A.

Winkle, A. (1991). "Listen my children: Strategies for listening instruction." *Ohio Reading Teacher*, 25, 14-20.

Appendix A.

Pretest Lecture Comprehension Test.

LECTURE TEST #1

I.D. number _____

Please answer the following multiple-choice questions.

1. From two to eight weeks after fertilization, all organs are created and the placenta becomes functional. Which stage of prenatal development is being described here?

- a. germinal
- b. embryonic
- c. fetal
- d. zygote

2. Concern about the health of the fetus would be greatest if the mother had rubella (German measles) during which of the following weeks of pregnancy?

- a. 5th week
- b. 10th week
- c. 15th week
- d. 20th week

3. The end of the fetal period is signaled by:

- a. implantation of the fetus on the uterine wall.
- b. birth.
- c. the differentiation of the blastula.
- d. the start of ossification.

4. Which of the following statements concerning consumption of alcohol by women during pregnancy is true?

- a. Any amount of alcohol is potentially harmful to the developing child.
- b. Consumption during the first eight weeks is relatively harmless due to the small size of the fetus.
- c. Damage occurs only during the embryonic stage.
- d. Consumption of small amounts of alcohol can actually be beneficial to fetal health.

5. If a baby is born with Down Syndrome, the most likely cause is:

- a. disease during pregnancy.
- b. drug use during pregnancy.
- c. a combination of stress and poor nutrition during pregnancy.
- d. maternal age.

6. One of the substages of the fetal period is:
- the differentiation stage.
 - the structural development stage.
 - the cognitive development stage.
 - the critical development stage.
7. Which childbirth technique has been called “birth without violence”?
- A traditional birth.
 - The Leboyer method.
 - The Lamaze method.
 - Natural childbirth.
8. Which is NOT an aspect of the Lamaze method of childbirth?
- The mother is taught a method of breathing and muscular control to minimize pain.
 - The use of a “coach” to give support to the mother during delivery of the child.
 - The use of anesthetics to remove all pain felt by the mother during delivery of the child.
 - Birth is treated as a celebration of life and not a medical procedure.
9. A pregnancy is most sensitive to teratogens during:
- the germinal period.
 - the fetal period.
 - the embryonic period.
 - conception.
10. Melissa, who has active genital herpes, had her baby delivered by Caesarean section. This form of delivery was used because:
- genital herpes deteriorates the birth canal.
 - it ensures fast treatment for the infant who is also infected.
 - it prevents transmission of the disease to the infant during birth.
 - factors unrelated to the herpes were in place.

Please answer the following short-answer essay questions.

- Describe a physical characteristic of a mother during pregnancy which can lead to increased chances of spontaneous abortion, or miscarriage.
- Describe the effects of the Rubella virus if it invades during the first two months of pregnancy.
- Maria is pregnant and has experienced severe and prolonged anxiety during her pregnancy. How will this anxiety affect Maria during her pregnancy?

Joan is pregnant and has been smoking throughout her pregnancy. What effects might Joan’s smoking have on the fetus?

On-line Quizzing and its Effect on Student Engagement and Academic Performance

**Mark G. Urtel¹, Rafael E. Bahamonde², Alan E. Mikesky³,
and Jeff S. Vessely⁴**

Abstract: The goal of this study was to determine if on-line out-of-class quizzing would lead to increases in (a) classroom engagement (b) academic performance and (c) preparation perception of college students. Twenty-four sophomore level students enrolled in a required functional anatomy course participated in this study. Results from this study indicate that on-line quizzing had a statistically significant impact on some indicators of student classroom engagement. However, on-line quizzing had no statistically significant impact on academic performance as evidenced by test scores. Finally, student perceptions, as a result of on-line quizzing significantly improved.

I. Introduction.

While research on in-class quizzing and student academic indicators is moderately represented in the literature, research regarding out-of-class on-line quizzing and student academic indicators is under represented. The authors in this study have utilized, at some point, traditional in-class quizzing in their courses, yet, none have trialed out-of-class on-line quizzing. Anecdotally, the measured outcomes from using in-class quizzes have varied among the authors, but, it is agreed that in-class quizzes can be somewhat cumbersome.

For example, in-class quizzes take away from instructional time (anywhere from 10-20% of the allotted class), in-class quizzing adds a layer of class management by taking up instructional time so one can distribute and collect the quizzes. Furthermore, this time needed to take and manage the quiz also decreases the opportunities for students to directly engage the instructor on class topics and conversation and, visa versa, for the instructor to engage the students directly.

When combining our collective experiences regarding in-class quizzing with the noticeable lack of research focusing on out-of-class on-line quizzing, we felt a study should be completed. This research need was further justified based on how uncertain the current literature is with regard to the effect quizzing has on various student academic indicators. We hypothesized that the impact of out-of-class on-line quizzing on student academic indicators would yield more unequivocal findings than what is currently documented on in-class quizzing. Thus, the present study was initiated. The hypotheses we tested were:

1. Out-of-class on-line quizzes would lead to an increase in student classroom engagement.
2. Out-of-class on-line quizzes would lead to an increase in academic performance.
3. Out-of-class on-line quizzes would lead to an increase in student perception.

¹ Assistant Professor of Physical Education, murtell@iupui.edu.

² Associate Professor.

³ Professor.

⁴ Professor, the above at Indiana University Purdue University Indianapolis, 901 West New York Street, Indianapolis, IN 46202.

II. Review of Related Literature.

The concept of quizzing in higher education is not novel. In fact, quizzing has been used by the faculty of college classrooms as an assessment tool, a teaching technique, or some combination for over thirty years (Mawhinney, Bostow, Laws, Blumenfeld, and Hopkins, 1971; Olsen, Weber and Dorner, 1968; Standlee and Pashan, 1960). While there is not an abundant amount of research on quizzing in higher education, it has not gone completely unnoticed either. Studies completed by Anderson (1984), Burns and Vinchur (1992), Connor-Greene (2000), Hagen (2000), and Peckham and Roe (1977) demonstrate the consistent and equivocal findings, over time, of the research on quizzing in higher education. These authors, as a whole or in separate, suggest that quizzing has demonstrable impact on student engagement and student perception toward learning, yet, is not independently linked to student learning / performance.

Studies on traditional in-class quizzing have served to inform faculty of the effectiveness of various teaching and assessment practices on student engagement, performance, and perception. For example, some studies suggest that in-class quizzing leads to an increase in the amount of time students spend reading course material out-of class (Connor-Greene, 2000; Mawhinney et al. 1971). This finding is especially encouraging as out-of-class reading by a student tends to promote academically engaged time (paying attention, contributing to discussions, and answering questions) by the student during the class (Ehrlich, 1995). Moreover, this academically engaged time is second only to general ability regarding having documented positive effects on classroom achievement or academic performance (Berliner, 1979; Denham and Lieberman, 1980).

In addition, according to Barbarick (1998), Connor-Greene (2000), Crooks (1988), Ehrlich (1995), and Hagen (2000), when quizzing is used in a course, students perceive themselves as being more prepared for class meetings and more prepared for class assessments. This increase in self-perception may motivate the learners to participate in class more by adding to discussions and asking questions (Connor-Greene, 2000).

However, with regard to academic performance, the findings have been more balanced. Anderson (1984), Connor-Greene (2000), and Olsen et al. (1968) offer views that, while admittedly not highly generalized, indicate that quizzing does not automatically lead to higher test scores or increases in student performance as indicated by final grades. This is in contrast to Barbarick (1998), Beaulieu and Zar (1986), Crooks (1988), Hagen (2000), and Martin and Srikameswaran (1974) who suggests that, with appropriate content overlap between the quizzes and performance assessments (tests), quizzes may lead to higher student performance on tests or in the form of final grades.

What is novel, however, is utilizing the current technology found in higher education today to deliver the quizzes. More specifically, college faculty has at their disposal today technologies and expertise that was simply not available forty years ago. As a result, when considering quizzing as an assessment tool, a teaching technique, or some combination, faculty currently can construct and deliver quizzing which looks substantially different to that found in college classrooms thirty, twenty, or even ten years ago. As a result, the use of out-of-class and on-line quizzes by instructors in higher education and their effect on student engagement, learning, and perception warrants further investigation.

Therefore, the theoretical framework adopted for this study views student academic indicators, first and foremost, as being able to be influenced by the classroom instructor.

Secondly, that the student academic indicators are engagement, performance, and perception. This view is consistent with the theory as presented by Berliner (1979) and Denham and Lieberman (1980).

III. Method.

A. Subjects.

Subjects (N=24) were enrolled in a required sophomore-level functional anatomy course. Students were informed on the course syllabus and during the first meeting day about their participation in the study.

It should be noted that this course is required for the enrolled students. As a result, there was a concern that manipulating the graded assignments and offering the course in a different way than how it is normally offered may unnecessarily add stress to the enrolled students. This course is normally delivered with daily in-class quizzes over the entire semester. For this study we gave quizzes for only one-half of a semester and made them out-of-class and on-line. The absolute impact of the quizzes as a course grade assignment was similar in both instances however. In the traditional course offering the value of the quizzes is 33% of the student's total grade and for this study the quizzes equaled 24% of the students' total grade.

B. Materials.

The on-line quizzes were generated by the course instructor and made available to the students using the university portal known as ONCOURSE™. All registered university students have access to this medium and when a student accesses the site their log-on and log-off times are recorded. Moreover, to decrease the chance of cheating on the quizzes there is a time limit to finish the quiz.

In addition, the students needed to complete the quiz prior to the corresponding lecture and could do so no earlier than 72 hours before class time and no later than 30 minutes prior to class. Moreover, the content of the quizzes directly reflected the assigned reading.

C. Definitions.

Student Classroom Engagement: Engagement was primarily defined by the student interactions that occurred during class (questions asked and questions answered). The first way we recorded student interactions was by a simple percentage of the class that asked a question to the instructor during the lecture sessions. Second, we assessed the total number of questions generated by the students to the instructor during the lecture sessions. Finally, we compressed Bloom's (1956) taxonomy into the three categories of Knowledge, Comprehension, and Application and evaluated the level of each question that students asked to the instructor. This compression is similar to the use of Bloom's (1956) simplified taxonomy by Cox and Clark (1998) and Crooks (1988). We trained a recorder to code every student-generated question directed to the instructor. If the question had a "recall" aspect to it was coded as a "Knowledge" question; if it had an "understanding or translation" aspect then it was coded as a "Comprehension" question; and, if it contained a "use or application" aspect then it was coded as an "Application" question. Herein, these are known as K, C, and A, respectively.

Each course meeting the trained recorder was present and tallied the student interactions as they occurred. However, to ensure accuracy we videotaped each course meeting and the trained recorder would validate their recordings by reviewing the videotape. Again, this was completed for both the quizzed and non-quizzed parts of the semester.

Regarding questions answered by the student, the instructor posed three questions to the entire class at the conclusion of the class session. There was one question at the K level, one at the A level, and one at the C level. Students were made to respond via paper/pencil and turn in their paper at the end of class. The responses were evaluated and recorded on a scale of 0-3. If a student was present yet did not / could not answer the questions it was recorded as a zero, if they answered 1 correct, then that equaled a 1, etc. If a student was absent it was recorded as "missing". An individual running average was recorded as was a daily class average. However, it should be noted that the students were fully aware these scores had NO impact on their course grade.

Finally, attendance was used as an indicator of student engagement. Attendance was tracked throughout the entire semester and stratified into "no-quiz" attendance and "quiz" attendance.

Academic Performance: A students' academic performance was evaluated by multiple choice tests completed by paper and pencil. Tests were administered at three points in the semester (pre, mid-term, and final). The pre-test contained 20 questions and was given the very first class meeting. The mid-term test was administered at the halfway point of the semester and the final test was administered during the last class meeting.

Ten of the pretest questions were reflective of material to be covered during the first half of the semester and were then embedded in the mid-term test. Similarly, the second ten questions of the pre-test were reflective of material to be covered during the second half of the semester and were then embedded on the final exam.

Again, each respective bank of ten questions came directly from the related reading and also varied with respect to the taxonomies used: K, C, and A. Furthermore, these questions were selected from questions generated from the previous semesters' test question analysis. Only questions associated with an index of discrimination of 40-60% were used for the pre-test, and subsequent mid- and final tests.

Student Perceptions: Student perceptions of their own level of engagement in the course were assessed following the first half of the course and, again, following the second half of the course. The instrument used was adapted from a university-specific student engagement project. Differences in perception were compared and reported as both individual and class average changes in class perception from the quiz and no-quiz portions of the course.

D. Procedures.

The first half of the semester (7.5 weeks) was dedicated as the no-quiz portion of the course and the second half of the semester (7.5 weeks) employed on-line out-of-class quizzing. To ensure consistency throughout the semester the reading schedule and class structure were controlled.

The reading schedule was formatted the same on the syllabus for the entire semester. Students clearly knew what reading assignment was to be completed prior to attending class throughout the entire semester.

The instructor structured the class sessions the same throughout the semester. The instructor started each class with a review of the reading and a chance to ask questions, then proceeded with the lecture (fielding questions throughout), and ending with the “informal quiz”.

The pre-test was administered the first day of class. At the mid-point of the semester the students were given the mid-term test and the self-perception inventory.

Starting at the eighth week of the semester, the students were reminded that graded quizzes were being implemented and that they needed to complete the quiz prior to coming to class. At the end of the semester, the students took the final test and the self-perception inventory was repeated.

E. Statistical Analysis.

A paired samples t-test and Pearson Product Moment Correlation were performed on the data collected for the quiz and no-quiz models. The level of significance was set at $p < 0.05$ for all analyses.

IV. Results and Discussion.

A. Engagement.

Analysis revealed no statistically significant difference between the percentage of students asking questions during the no-quiz and the quiz portions of the course. However, the total number of questions asked between the no-quiz and the quiz portion of the course was significantly different. During the no-quiz period there were 34 total questions asked in class during the first-half of the course. This is opposed to 74 total questions asked in class during the second half of the course where quizzing was utilized.

As such, the same “pool” of students tended to ask questions whether there was a quiz or not, yet the number of questions generated by this “pool” of students increased after the quizzes started. This increase could be explained away by arguing that students got more comfortable with the course or instructor over time, as such, the students asked more questions during class time. However, the results do not suggest this.

If the students got used to the course and instructor over time (7.5 weeks) one would predict that the distribution of questions would gradually increase over this time. Where the fewest questions were asked during week one, when the students would have been the least comfortable with the course or instructor. And, the most questions would have been asked during week seven, when the students should have been more comfortable with the course and instructor. Yet, the results do not reflect this trend. Weeks two and three yielded the most K, C, and A questions from the students, then there was a slight decrease and stabilization of student generated questions from weeks four through seven. One could speculate that this question distribution indicates that the students were comfortable with the course and instructor by at least week three.

Moreover, when factoring in the nominal increase with the percentage of students asking questions over the first-half of the semester, one might conclude the students were comfortable with the course and instructor early on. As a result, it may very well be that on-line out-of-class quizzing, and the necessary pre-class reading associated to that, may lead to more engaged students during class time.

Regarding the level of student questions asked, it was found that no statistically significant difference existed regarding the number of K level questions asked by the students during the no-quiz portion of the course and the quizzed portion of the course ($p = 0.824$). However, there was a statistically significant difference in the number of C and A level questions asked between the no-quiz and quiz portions of the course ($p = 0.026$ and $p = 0.002$, respectively). While this absolute and significant increase in level of question asked cannot be definitively linked to the introduction of the quizzes, the results do illustrate another positive trend toward more in-class student engagement when on-line out-of-class quizzes are introduced.

There was also a statistically significant difference in regard to student answers of the instructor questions at the end of each class session. It was found that the number of “zero” or “present but could not answer” responses significantly decreased from the no-quiz to the quizzed portion of the course ($p = 0.000$).

This finding supports the idea that by assigning graded quizzes there will be an increase in student preparation and a subsequent increase in a student’s ability to answer questions about the course material. More important though, this finding suggests that on-line out-of-class quizzes yield the same benefit as in-class quizzes but without the “expense” of class time and management.

It was found that on-line out-of class quizzing had no impact on class attendance ($p = 0.68$). The same students missed class whether it was the no quiz or quizzed portion of the course.

B. Academic Performance.

There was no correlation between the pre-test score (average of 2.4/10) and the mid-term score (average of 5.4/10) in regard to the ten-question component ($r = 0.267$). Moreover, there was no correlation between the corresponding ten-component pre-test score (average of 2.2/10) and the corresponding ten-component final test score (average of 4.9/10), the $r = 0.231$.

Both of these results were expected and favorable as there should be a significant difference, and subsequent low correlation, between a pre-test score and, in essence, a post-test score. This implies that the “treatment” (quiz) had a positive impact.

For the purposes of this study, however, the fact that the rates of improvement were so similar for both the no-quiz and quizzed portions of the course indicates that on-line quizzing may not automatically lead to better test scores. There are many factors that influence test-scores. These may include but not be limited to (a) content overlap between the reading and the test questions, (b) format of test questions, (c) the wording of the questions, and (d) the taxonomy of the test questions. Without careful control, test questions themselves could influence test scores more than anything, interventions included (in our case, quizzes).

Admittedly, our criterion for test question selection may have been too restrictive and aggressive, thus, placing a low ceiling on how much improvement one may expect from pre-test to post-test. As a result, any conclusions regarding on-line quizzing and academic performance, via test score, are preliminary as stated here.

C. Student Perceptions.

Some student perceptions changed significantly from the no-quiz portion of the course to the quizzed portion of the course; however, not all of them were favorable. Regarding the prompt “I plan specific study times for this class”, during the no-quiz portion slightly over 50% of the students either agreed or strongly agreed, however, during the quizzed portion slightly over 80% of the students either agreed or strongly agreed to that statement. This indicates a formal acknowledgement by the student to dedicate time to reading and studying.

However, responses to “Attending class is critical to achieving the grade I expect in this class” went from 95% of the students either agreeing or strongly agreeing to that statement to about 80% during the no-quiz portion and quiz portion of the course, respectively. It is difficult to ascertain the cause for this shift, however, one could speculate that by doing the reading in advance for the quiz, the actual class session and related material would be more familiar and less novel to the student.

Finally, 83% of the twenty-four students responding to the student perception survey felt that the on-line out-of-class quizzes made them “learn the material better”. Moreover, 80% of the respondents felt that graded on-line out-of-class quizzes motivated them to “read prior to coming to class”, and, finally, 71% of the students felt the on-line out-of-class quizzes “got them ready for class discussions”. It is clear that student perception of on-line out-of-class quizzing was favorable.

V. Conclusions.

The results of this initial investigation are encouraging. We have concluded the following:

1. The hypothesis statement: “On-line quizzes would lead to an increase in student classroom engagement” is partially supported. Specifically:
 - a. The total number of questions asked during class by students increased significantly as a result of on-line quizzing.
 - b. The C and A level questions asked by the students increased significantly as a result of the on-line quizzing.
 - c. Students answered the instructor questions at a significantly greater rate during the quizzed portion of the course than the no-quiz portion.
 - d. The percentage of students who asked questions during class did not increase during the quizzed portion of the course.
 - e. The K level questions asked by the students did not increase as a result of on-line quizzing.
 - f. On-line quizzing did not influence class attendance.
2. The hypothesis statement: “On-line quizzes would lead to an increase in academic performance” was not supported.
3. The hypothesis statement: “On-line quizzes would lead to an increase in student perception” was partially supported.
 - a. There were four prompts that students positively attributed to on-line quizzes and one prompt which students negatively attributed to on-line quizzes.

VI. Reflective Critique.

As a result of completing this preliminary study, the authors have continued their professional dialogue on the merits of quizzing, generally, and out-of-class on-line quizzing, specifically. While each of our areas of expertise and individual pedagogical beliefs drive our instructional and assessment methods, this research project has undoubtedly influenced all of us. The findings from this study have confirmed some of our intuitions while at the same time tempered other intuitions. Moving forward, we (collectively and individually) have confidence of what quizzing may or may not accomplish as part of a college classroom. This perspective will help guide the development of instructional strategies to help improve the teaching and learning process.

In addition, we generally agree that on-line out-of-class quizzing is an attractive alternative to traditional in-class quizzing that, at best, contributes to the improvement of certain student academic indicators. As important, we generally agree that, at worst, on-line out-of-class quizzing does no harm to the student academic indicators. Given the unintended benefits of out-of-class on-line quizzing, such as (a) timely feedback to the student (b) automatic grading and corresponding ease of entry into an electronic grade book and (c) saving in-class instructional and discussion time, we feel that out-of-class on-line quizzing can be a significant tool for educators to utilize.

References

- Anderson, J.A. (1984). "Frequency of quizzes in a behavioral science course: An attempt to increase medical student study behavior." **Teaching of Psychology**, Vol. 11, No. 1, p. 34.
- Barbarick, K.A. (1998). "Exam frequency comparison in introductory soil science." **Journal of Natural Resources and Life Science Education**, Vol. 27, pp. 55-58.
- Beaulieu, R.P. and Zar, M.C. (1986). "The effects of examination frequency on student performance." **Journal of Instructional Psychology**, Vol. 13, No. 2, pp. 81-89.
- Berliner, D.C. (1979). *Tempus Educate* in P.L. Peterson and H.L. Walberg (Eds.) *Conceptions of Teaching*. Berkeley, CA: McCutchan.
- Bloom, B.S. (1956). *Taxonomy of Educational Objectives: Handbook 1. Cognitive Domain*. New York: McKay.
- Burns, D.J. and Vinchur, A.J. (1992). "Effects of evaluative quizzes on test performance." **Journal of Instructional Psychology**, Vol. 19, No. 3, pp. 148-154.
- Connor-Greene, P.A. (2000). "Assessing and promoting student learning: Blurring the line between teaching and testing." **Teaching of Psychology**, Vol. 27, No. 2, pp. 84-88.
- Cox, K. and Clark, D. (1998). "The use of formative quizzes for deep learning." **Computers in Education**, Vol. 30, No. 3/4, pp. 157-167.

Crooks, T. (1988). "The impact of classroom evaluation practices on students." **Review of Educational Research**, Vol. 58, No. 4, pp. 438-481.

Denham, C. and Lieberman, A. (1980). Time to learn. Washington, D.C.: National Institute of Education.

Ehrlich, R. (1995). "Giving a quiz every lecture." **The Physics Teacher**, Vol. 33, pp. 378-379.

Hagen, J.P. (2000). "Cooperative learning in organic II: Increase in retention on a commuter campus." **Journal of Chemical Education**, Vol. 77, No. 11, pp. 1441-1444.

Mawhinney, V.T., Bostow, D.E., Laws, D.R., Blumenfeld, G.J., and Hopkins, B.J. (1971). "A comparison of students studying behavior produced by daily, weekly, and three week testing schedules." **Journal of Applied Behavior Analysis**, Vol. 4, No. 4, pp. 257-264.

Martin, R.R. and Srikameswaran, K. (1974). "Correlation between frequent testing and student performance." **Journal of Chemical Education**, Vol. 51, No. 7, pp. 485-486.

Olsen, R.E., Weber, L.J., and Dorner, J.L. (1968). "Quizzes as teaching aids." **Journal of Medical Education**, Vol. 43, pp. 941-942.

Peckham, P.D. and Roe, M.D. (1977). "The effects of frequent testing." **The Journal of Research and Development in Education**, Vol. 10, pp. 40-50.

Standlee, L.S. and Pashan, W.J. (1960). "Quizzes' contribution to learning." **Journal of Educational Psychology**, Vol. 51, pp. 322-325.

Service Learning in English Composition: A Case Study

Maria Mikolchak¹

Abstract: Although service learning has gone a long way since the time when, until the mid-1980, it was practically unknown as a pedagogical practice beyond a closed circle of practitioners, in many higher education institutions it still remains a domain of few faculty interested in integrating service learning in their disciplines. While experimenting with teaching English 191 Composition course at St. Cloud State University in Minnesota and searching for a way to relate the abstractions of my discipline to the realities of the world, I stumbled into service learning and created a course that truly convinced me of the revolutionary potential of service learning in transforming the stale educational practice. This article will record my experience with the course and the lessons I learned from it.

I. Course Objectives.

“The purpose of English 191 is to introduce you to college-level writing, reading, and thinking. By the time you finish this course you should be able to read critically various texts that make up American culture and literature and write about them. Also, the course will help you to make informed judgments based on research as well as interpretation of your personal experience,” explains the standard course description for my freshmen composition course. It is this last part, “your personal experience,” that gave me most trouble. No matter what topics we covered in the course, not all of the students could relate to them, which made interpretation of the personal experience an empty promise.

In studying a particular topic of violence the semester before I first thought about service learning, I heard from my students that most of the things we read about or watched in the videos do not happen where they live (mostly Minnesota), which for many students means that they do not happen at all.

To help students relate to the issues discussed in the course – and violence was one of them – service learning seemed to be the perfect pedagogy. Trying to define service learning for myself and going through literature, I realized that much of literature on service learning records a prolonged debate on what service learning is. In 1990 Jane Kendall wrote that there were 147 definitions of service learning in literature [Eyler (1999)], which for me was a clear sign that if I do want to teach English 191 as a service learning course in the summer semester, I cannot get into the definition polemics but should accept something that makes sense to me as a working definition. I ended up with the definition offered by Robert Bringle and Julie Hatcher of the Office of Service Learning at Indiana University-Purdue University Indianapolis: “Service Learning is a course-based, credit bearing educational experience in which students (a) participate in an organized service activity that meets identified community needs and (b) reflects of the service activity in such a way as to gain further understanding of the discipline, and an enhanced sense of civic responsibility”[Bringle (1995)].

¹ Associate Professor in the Department of Foreign Languages and the English Department, St. Cloud State University, 720 4th Ave S, St. Cloud, MN 56301. MMikolchak@stcloudstate.edu

II. Course Structure.

The course consists of three interconnected major parts: teaching an argument, interpreting images, and analyzing violence in society. The first part of the course introduces students to the Toulmin model of argument. My purpose is to show the students that all language, including the language of visual images, can be seen as an argument. We spend several class periods on practicing making arguments, critiquing arguments, recognizing fallacies, and, more generally, figuring out how to write a persuasive essay.

The second part of the course builds on the theory of the argument and asks students to think more specifically about images in media and how these images can be seen as arguments. We watch Judy Kilbourne's "Killing Us Softly" and Jackson Katz's "Tough Guys" through the lens of the argument. The topic of violence is introduced through those videos as well as additional readings, so discussions flow naturally from the material students are exposed to.

By the time we cover the first two parts of the course, I aim at reaching two major goals: teaching the students to critically read arguments (analysis) and to write about them (production).

The third part of the course is devoted to a major research project with the general topic of Violence Against Women. To teach the students research strategies, I arrange for special library research sessions (a total of four hours) where a librarian introduces the class to the library, resources, and research process. The lecture is accompanied by a hands-on experience when students practice conducting research on sample topics. This semester I was fortunate to have Pamela Salela, who is not only an experienced librarian but also is a specialist in women's studies, conduct the library research sessions for my class, and the sessions were extremely productive.

III. Service Learning Project.

Having initially outlined the course for myself, I started considering how to integrate service into it. I found it useful to think of service, as Keith Morton suggests, in the sense of a "text" [Morton (1996)]. As Morton further points out, service is not a traditional text and, most importantly, it is written concurrently with the course, but thinking of it as a text has a number of benefits. "First, it suggests that service is equal to written work in its learning potential," Morton points out. Second, "the analogy of texts implies that faculty must decide what texts are appropriate for the course and whether they are required or optional" [Morton (1996)]. The analogy of the text fits especially well a composition course where students read and critically analyze texts to be able to produce texts of their own.

Obviously, there are certain limitations in what instructors can choose as a service text. There are a number of practical considerations involved. What organizations need this service? How many hours are needed? Will students be able to make those hours? – to name just a few.

Working on the preparation for the course, I found the following principles suggested by Campus Outreach Opportunity League (COOL) and reproduced in the book *Service-Learning in Higher Education*, edited by Barbara Jacoby, especially useful: Community voice (service learning should aim at meeting the needs of the community); Orientation and training (students should be provided with information on their service, organization for which they do the service, and the issue); Meaningful Action (the service should be necessary and valuable to the community itself); Reflection (this crucial component of the service learning experience should happen immediately after the experience to discuss it in order to place the experience into a

broader context); Evaluation (students should evaluate their learning experience and agencies should evaluate the effectiveness of the students' service) [Mintz (1996)].

In coordination with the service-learning center at SCSU, I have found a place that needed service and that would be directly related to the text of violence studied in the course – Annemarie's Shelter for battered women.

For any service learning project to make sense it has to be oriented not only towards the students and their learning goals, but first and foremost, towards the needs of the community. What Annemarie's needed was help with cleaning and painting the transitional house belonging to the shelter (transitional house is a place where women can live for a while for a nominal fee after they leave the shelter as a transition to a regular housing situation).

As much literature on service learning points out [Mintz (1996)], good organization is time consuming and requires a lot of energy on the part of the instructor. Without effective administration, integrating service learning into a course might become just another failed revolution in pedagogy. I am particularly grateful to Judy Gay from Annemarie's shelter and to Eveily Freeman, Service Learning Coordinator at SCSU, who assisted me with service arrangements and with adapting our service to the learning needs of the students. We had several meetings long before I even started working on the course syllabus in which we discussed the needs of the shelter, the needs of the course, and the best ways to integrate service into the discipline. We planned several presentations for the students before actually exposing them to the service. They included Eveily Freeman's presentation on service learning early in the semester and two presentations by Annemarie's social workers as we went along. This allowed the students ample time to think about our expectations from them and to address possible problems with service learning.

Minor and major organizational problems included time management, transportation, work supervision, and providing a lunch, to mention just a few. Several students had class conflict since service was arranged for two specific days at a set time. As an instructor, I helped solving the problem with other instructors, in most cases arranging for individual assignments to cover for the missed class. I am happy to say that I met with full understanding and cooperation on the part of my colleagues at SCSU. Although service learning is not institutionalized as a program at our university, individual endeavors are certainly supported by other faculty and administration.

The students were to spend ten hours total (two five-hour days, plus one hour for lunch break) working on the transitional house. Annemarie's provided tools and lunch. Also, on the second day of work, students were invited on a tour of the shelter where Judy Gay talked to them about the history the shelter and showed the facility. In addition, before the actual service days, we had two presenters from Annemarie's who talked about the shelter and the problem of domestic violence and, in particular, violence against children. Part of the presentation was a video – another text the class had to analyze.

IV. Course Outcome.

Participation in the service learning project remarkably increased students' interest in the topic discussed in class and their understanding of the issue. It made the whole learning process more meaningful. While choosing the concrete topic for research papers, students had a very good idea what they wanted to research – unlike a typical class where a lot of students totally depend on the instructor for the topic choice. Research papers also demonstrated personal

involvement and allowed students to use their service learning experience to relate to the data they researched.

The goal of the final presentation in class on research project was to allow the students to share with the class how they chose their topic and conducted research. This allowed for more reflection. Thus, for example, one student wanted to research a history of the shelters in Minnesota, but was surprised to find out that the information was almost non-existent. His conclusion was that although domestic violence is one of the major problems in the country, there is an amazing lack of interest to it on the part of the government and a very poor funding.

Service learning allowed the students to connect all parts of the course into a meaningful whole: analyzing texts, viewing all texts as arguments, producing a critique thereof, and conducting research on the topic to which service learning allowed them to have more sensitivity and understanding.

V. Reflection.

Importance of reflection in learning and in service learning in particular has been acknowledged broadly in the literature on service learning. Hutchings and Wutzdorff, for example, write that “[t]he capacity for reflection is what transforms experience into learning” [Hutchings (1988)]. Through class discussion I tried to encourage the students to think critically about their experience and to generate ideas as a community of readers and writers. Class discussions also offer opportunity for instructor to challenge certain ideas and offer others without forcing any ready-made answers on the students. Probably, the most frequent and important comment they heard from me was that there is no one correct answer to the question asked. Generally, a persistent problem with class discussions in English 191 is that since it is a freshmen class, most of the students are still shy to talk in public or, to be more specific, in a college environment to which they are new. That is why students oftentimes prefer journals as a form of reflection. They consider journals a safe place for honest reactions. The fact that this reflection is put in writing is also highly beneficial for the class that essentially is a composition class. One of the students wrote in her journals that, in fact, journal writing was the only form of writing that she liked, and not surprisingly her journals were much longer than the suggested one-page length.

Likewise, although the syllabus required students to have a certain amount of journal entries, with most of the students the number of journals in their portfolio exceeded the requirement. In fact, one student even made fun of herself and her developing addiction to journal writing.

All students expressed a positive attitude to the service learning project, although for different reasons and in very different ways. For one of the students the first service day was “one of the funniest days of class ever.” The student explains that it is during the service work that people in class really got to know each other and to interact: “Everyone was having an awesome time and the work was easy even though I don’t even mind work like that. It didn’t even seem like work because we did not have someone over our heads watching us and criticizing us. It was volunteer work and I really enjoyed the feeling I got when I got done. You cannot buy that type of feeling anywhere. There is no price for it. It makes me really want to start volunteering so I can help make a little different in someone’s life. I might apply at Annemarie’s.” This student also mentions how happy he was that people at the shelter appreciated this work.

Although this is a positive evaluation of the experience, it is easy to see that it mostly revolves around the student himself and the way he and other people see him. While this may seem discouraging, most researchers notice that egoism as acting toward the ultimate goal of increasing one's own welfare with rewards such as feeling a sense of accomplishment and satisfaction, gaining skills, and affiliating with others is critical to initial and continued involvement in service [Winniford (1995)]. The following journals of the same student, however, show the progress from what service learning does to the student to how it helps other people: "I and my friend both wanted to help out and make a difference in St. Cloud where we will be living for the next three years. We don't feel that violence in the answer in families and we should try and change that for the better." Thus, we can see a movement from egoism to altruism (acting with the ultimate goal of helping others) as a result of a continued involvement with the project.

While most students find the service learning experience helpful in achieving class goals, one student writes that he does not know what this project has to do with the class. Nevertheless he says he is happy to have worked at the shelter, since on his own he would have never done any volunteer work. For this student service learning did not work out. Enjoying volunteer service has its merits, but within the academic context the purpose of service learning is not to push students into volunteerism but to help them acquire academic knowledge in the discipline through service. However, while for this particular student the class academically the class did not bring the expected results (largely owing to absences from class and insufficient time investment in studying), his positive experience at Annemarie's and his desire to do more work for the shelter was at least one positive outcome of his total class experience.

As an instructor, I was fascinated to see how students reflections developed from "the fun class" and "making friends" through thoughts on the poor living conditions of women and children in the shelter to trying to deal with the problem itself. Students start talking about domestic violence and abuse as a social evil, about budget cuts that affect the shelter, about responsibilities of politicians, and about the upcoming elections. As Janet Eyler and Dwight Giles summarize it, "[s]ervice-learning aims to connect the personal and intellectual, to help students acquire knowledge that is useful in understanding the world, build critical thinking capacities, and perhaps lead to fundamental questions about learning and about society and to a commitment to improve both" [Eyler (1999)]. Marylu McEwen names as anticipated learning and development outcomes for the students the following: greater complexity in thinking, ethical commitments regarding themselves, and what they know and believe; greater awareness of themselves as and of their own racial, ethnic, and cultural heritage; greater sense of their place in the United States; increased tolerance and empathy; greater clarity about themselves and their life purposes; and development and maturity of their values [McEwen (1996)]. I am happy to say that the design of the course I taught led to most of the above-mentioned outcomes.

There is evidence that students can better analyze a social problem when they combine academic knowledge gained in class with personal experience through service [Eyler (1999)]. That is exactly what happened in the class I taught. One of the most positive outcomes, noticeable to me as instructor, was students ability to cope with problems for which there is no obvious solution. Thus, from easy suggestions at the beginning of class that abused women should "simply leave the abuser" students developed a much more mature and informed understanding of the complexity of the issue and of the lack of a radical solution to the problem.

VI. Final Research Project.

Part of composition class is teaching students to do a major research paper. The problem I usually run into as instructor is the lack of interest on the part of the students. I can teach them how to do research and I can offer topics that, I assume, should both reflect the content of the course and be of interest to the students. On many occasions I truly felt I can lead the horse to the water but I cannot make him drink. With the service learning experience involved, the students were clearly interested in researching topics discussed in class during the reflection class time. They were given much freedom in choosing the topic of their liking while connecting it to the service learning experience was not a requirement. Interestingly, all the students in class ended up researching a topic that was connected to their service. Final projects were, generally, of a better than average quality, which I think to a great degree reflects the genuine interest the students had in their research. Another quality common to all of the projects was their more personal character not only in the sense of referring to the service experience and reflection on it, but also in the sense of talking about their own families and communities in which they have grown up. Also the papers contained a richer mix of sources, including interviews, recalled experiences of service, lectures, and videos. Thus, service learning helped eradicate the common scourge of English composition classes -- the lack of engagement with the material.

VII. Evaluation and Grading.

While oftentimes faculty is concerned about not being able to observe the service directly, I was lucky to participate in the service project together with my students on both days. Nevertheless, as I mentioned before, students commented on the lack of pressure and explained that nobody was watching them or criticizing. I am happy that my students accepted me in their group as an equal, not as a supervisor, which in fact I was not. Our work was supervised by a person from the shelter.

Another factor that I think contributed to the success of the experience was that I did not grade the service per se. The syllabus explained that since it was a service learning class, participating in the service was a necessary prerequisite for passing the class. However, no particular percentage of the grade was assigned to the work at the shelter. While students were not graded for how well they worked, they were graded for demonstrating what they have learned both from service and from other class assignments, that is they were graded for the work specific to the discipline.

VIII. Final Evaluations.

Final evaluations for the class demonstrate the overall success of the service learning experience. There is a lot of evidence of the growing social consciousness on the part of the students and willingness to actively offer help to the community they live in. While I do not want to overestimate students' willingness to do more for the shelter and to volunteer on a regular basis (not everybody will eventually do it), evaluations convey a feeling of awakening sensitivity to the societal needs.

From the thirteen students registered for the class, all but one expressed desire to do volunteer work for the community in which they live. Seven people said they want to contact Annemarie's shelter and offer their services in the coming school year. Two male students who

asked most questions about people who work at Annemarie's during our tour and learned that Annemarie's were looking for male volunteers (to give children a positive example of male involvement), expressed desire to work with the children.

In general, students, for most of whom this course was the first college experience, gave a very high evaluation (5 on a scale from 1 to 5) of learning through service and said they would be willing to take more classes with service involved.

IX. Conclusion.

While English composition has often been criticized for working in unreal rhetorical situations [Heilker (1997)], service learning creates a very real situation with a very real audience and very real needs. It also gives students ideas they want to research and write about, an asset in a class where the majority declares from the very beginning that they hate to read and to write and are taking this class because it is a requirement.

While I have no doubt that service learning is a beneficial pedagogy, I also understand that it is not successful automatically. It requires a lot of time and effort to make it work. But my experience with a service learning composition class convinced me that the outcome is worth the effort.

References

- Bringle, Robert and J. A. Hatcher. (1995). "A Service Learning Curriculum for Faculty," *Michigan Journal of Community Service Learning*, Vol. 2, pp. 112-122.
- Eyler, Janet and Ciles E. Dwight E. Ciles, Jr. (1999). *Where's the Learning in Service-Learning?* San Francisco: Jossey-Bass Publishers, p. 3.
- Heilker, Paul. (1997). "Rhetoric Made Real: Civic Discourse and Writing Beyond the Curriculum." In Linda Adler-Kassner, Robert Crooks, and Ann Watters, eds. *Writing the Community: Concepts and Models for Service-Learning in Composition*. Urbana: AAHE, p.71.
- Hutchings P. and A. Wutzdorff. (1988). "Experiential Learning Across the Curriculum: Assumptions and Principles. In P. Hutchings and A. Wutzdorff, eds. *Knowing and Doing: Learning Through Experience: New Directions for Teaching and Learning*. San Francisco: Jossey-Bass, p. 15.
- McEwen, Marylu K. (1996). "Enhancing Student Learning and Development Through Service-Learning." In Barbara Jacoby and Associates. *Service-Learning in Higher Education*. San Francisco: Jossey-Bass Publishers, p. 87.
- Mintz, Suzanne and Garry Hesser. (1996). "Principles of Good Practice in Service-Learning." In Barbara Jacoby and Associates. *Service-Learning in Higher Education*. San Francisco: Jossey-Bass Publishers, pp. 30-31.
- Morton, Keith. (1996). "Issues Related to Integrating Service-Learning into the Curriculum." In Barbara Jacoby and Associates. *Service-Learning in Higher Education*. San Francisco: Jossey-Bass Publishers, p. 282.

Mikolchak, M.

Winniford, J. C., D.S. Carpenter, and C. Grider. (1995) "An Analysis of the Traits and Motivations of College Students Involved in Service." ***Journal of College Student Development***, Vol. 36, pp. 27-38.

Student Compliance with Assigned Reading: A Case Study

Brian D. Brost¹ and Karen A. Bradley²

Abstract: Educators value reading as an academic practice and express frustration with student noncompliance with assigned reading. Some research has addressed this issue and recommends multiple strategies for overcoming the problem of student noncompliance. Much of this research, however, treats the problem as exclusively student-centered. This paper presents a small case study designed to engage explanations for student noncompliance and to reflect on strategies used by faculty to incorporate reading assignments into their course design. This study suggests that more attention be placed on the reciprocal nature of the problem and how faculty behavior can contribute to reading apathy. How faculty members conceive, integrate and utilize assigned reading in the classroom does affect how students respond and take responsibility for the practice.

I. Introduction.

Though the classroom is the focal point of the educational experience, learning and teaching does not end there. When students leave the classroom, they continue the learning process by studying their lecture notes, conducting research, writing papers, solving problems and, of course, reading assigned texts. For many classes, the learning process is thought to hinge on students reading assigned material in preparation for and in tangent with the learning activities in the course (Altman and Cashin 1992, Nilson 1998, Grunert 1977). The time required to read course materials might easily exceed the amount of time that an instructor meets with students in class (Lang and Gore 1988). Even when teachers do not use readings as an integral part of their classroom pedagogy, they often recognize that they have a responsibility to produce active, critical readers as lifetime learners, and they may also expose students to important, challenging texts in order to promote cultural fluency.

Given the importance that teachers attach to assigned reading, it is not surprising when educators report frustration over their perception that students are not preparing assigned reading with the care they would like. Nor does it appear that teachers' perception of non-compliance is mistaken. Burchfield and Sappington (2000) report that, on average, only about a third of all students complete their text reading assignment on any given day. This recent finding confirms a body of research conducted over the last 30 years (McDougall and Cordiero 1993, Self 1987, Marshall 1974). Hobson (2004) clearly captures the dilemma stating, "faculty face the stark and depressing challenge of facilitating learning when over 70% of the students will not have read assigned course reading" (p.1).

While many reasons for noncompliance are offered, researchers repeatedly cite the *unpreparedness* of students to read the kinds of materials typically assigned in college coursework. The idea is that students don't read because they can't read; at least, they do not read well enough for the kind of texts most faculty assign. Leamson (1999), for example, argues that today's students are entering school with major deficiencies, and he believes that "the major

¹ Associate Professor of Philosophy, brost@cmsu.edu

² Associate Professor of Sociology, both at the University of Central Missouri, Warrensburg, MO 64093.

deficit is in language use and language understanding” (p. 43). Bean (1996) makes a similar point with the following analogy: “Armed with a yellow highlighter but with no apparent strategy for using it and hampered by lack of knowledge of how skilled readers actually go about reading, our students are trying to catch marlin with the tools of a worm fisherman” (p. 133). And this problem has not gone unnoticed by others (Lowman 1995, Healy 1990, Nist and Kirby 1989).

Since many believe that the problem of noncompliance stems from the problem of unpreparedness, remedies for noncompliance tend to be student-focused approaches. For example, Hobson (2004) asserted that teachers can get students to read by taking a “less is more” approach to course reading and by aiming material at “marginally-skilled” students. Other researchers echo these sentiments (Leamson 1999, Bean 1996).

Much of the research done on the problem of noncompliance has focused on students relatively new to the university—first-year students and students in general education classes (Burchfield and Sappington 2000, McDougall and Cordiero 1993, Erickson and Strommer 1991). Less has been done to understand the problem in advanced courses. This paper, however, examines student reading compliance in advanced level, elective coursework. Looking at rates of compliance in this context is important for several reasons. First, since advanced students are presumably skilled readers, unpreparedness (what many see as the chief problem to compliance) should not form a substantial barrier to the completion of assigned reading. Additionally, when students get to *elect* their course (presumably choosing a course that interests them), *prima facie* we should expect the students to be more motivated to read the assigned materials. Thus we can move our focus away from the problem of unpreparedness and toward the problem of noncompliance.

The objective of this research is to examine rates of compliance with assigned reading in advanced students and, more importantly, we seek to understand students’ attitudes toward the use of assigned reading. In addition, we explore the way in which faculty utilize reading assignments in the learning process. Our overall goal is to become more aware of the role instructors play in reading noncompliance and to consider ways of improving student receptiveness to the practice of reading. This is one small step towards understanding how we help students reach the university goal of becoming critically engaged readers both in the classroom and beyond.

II. Methodology.

This research is a case study of a course designed as a lecture series; that is, weekly classes in which professors from different disciplines were asked to participate as a one-time guest lecturer. The course was conducted at a comprehensive regional university as part of the philosophy program. It was organized and implemented by the first author and was attended by both. The course was not conceived with any intent to do research. Rather, our interest in this research was sparked both by observing the wide variety of ways in which the faculty members conceived of reading and its relation to their classroom activities and by observing how students responded to these various methods. We recognize the limitations of its conclusions but still feel the study is useful in terms of highlighting some of the dilemmas faced by faculty and students.

As part of the course, we observed the variety of sources, length, and difficulty of the reading material and assessed the way faculty engaged the material in their lecture. Thirteen faculty members participated in the lecture series. Each was selected based on their expertise.

Before the term began, lecturers were to submit readings to be assembled into a course packet. Guest instructors were directed that the readings should reflect a typical reading load for a week's worth of class time, and that the course was listed as a junior-level course. It was assumed that students would complete the reading assigned each week *before* the class met so as to make the in-class meeting more productive. After each class, students were asked to write a critique of the lecture and accompanying reading. Together, these critiques determined half of a student's final grade. Two midterm essay exams accounted for the remainder of the final grade. Students were given the option of substituting one research paper for one of the exams (only one student opted for this alternative). Students were not graded on in-class participation.

Twelve students participated in the course, although the lectures were open to other observers. We assessed levels of compliance with and attitudes toward the reading assignments by administering an end of the semester survey. The second author who had no grading responsibility over the students administered the survey. Even so, the results of the survey were not examined until after course grades were assigned. In addition, we observed the behavior of the students at the lectures over the course of the semester, the quality of the work they produced, and we reflected upon our own receptivity to the assigned readings.

We did not collect any demographic information from the students in order to protect the anonymity of the survey process as much as possible. However, we can make a few generalizations about the students in the course. All but two of the students were white. All were American, and most were from the Midwest. Most students were philosophy majors or minors, and most were traditional students in terms of age. Only one had dependents and only two worked full time. Generally, we would gauge the students to be average or above average students in terms of academic ability for this institution.

Doing the research in this context was useful in that it allowed us to look at both the students and the professors in the context of a course. In addition, the course as a lecture series allowed us to look at a cross-section of faculty from various disciplines across campus. The limitations of a case study this small, of course, is that the results are specific to the case studied. Even so, we feel the case provides important clues as to how to proceed in more generalizable studies in the future.

III. Findings.

A. Faculty.

Our first observation was simply that the type and amount of reading varied dramatically. For example, in terms of length, one speaker asked students to prepare an assignment of only about 5,000 words, while another asked students to read closer to 77,000 words—a significant difference in terms of expectations. Additionally, the *types* of reading assigned widely varied. We noted five different types of reading that included: 1) simple handouts produced by the lecturer; 2) non-scholarly articles (a book review in the magazine, *Nature*, for instance); 3) textbooks chapters and secondary sources; 4) scholarly articles (for example, one speaker assigned three articles of her own published research); and 5) primary and/or classic sources (“Self Reliance” by Emerson and part of Freire’s *Pedagogy of the Oppressed*, for two examples). Some of these readings were quite elementary (textbooks and non-scholarly articles) while others we would deem as very challenging (primary and scholarly articles).

Lecturers had diverse strategies for utilizing these readings in class. On the one hand,

several speakers geared their lecture to summarizing, reviewing and highlighting material covered in the assigned text. On the other hand, one lecturer read a scholarly paper in class, a paper that had little overt connection to the assigned reading. A majority of the professors presented lectures where elements in their presentation were drawn from topics covered in the reading they had assigned. Nevertheless, of the 13 lecturers, less than half directly used the texts in their lecture; that is, the faculty did not utilize specific passages during class nor ask students questions specifically about the reading. Generally, we judged the lectures to be accessible to students whether or not the reading had been done.

B. Students.

As a group, the students were quite active in discussion, asking for clarifications and raising objections. However, it seemed to us that student questions were rarely drawn from the reading, and at times, students asked questions that made it quite evident that they had not read the material before the lecture (at least, not carefully). Student in-class responses were most often stimulated by the oral presentation of the faculty member or drawn from the students' own experiences with the subject matter.

The survey given at the end of the class yielded interesting results. We began by simply listing all the reading assignments and asking students how much they read of each piece. Many of the lecturers assigned several distinct pieces and these were listed separately. Their response options included: none (0), some (1), most (2), and all (3). The average responses across the reading assignments ranged from .5 to 1.9 meaning that the students tended to read "some" or "most" of the assigned readings but rarely "all." The assignments with the most reported readers were those with one or two pieces, meaning that most students read at least some of each article. However, the lowest mean scores were given to the reading assignments with more numerous articles, indicating that students didn't begin many of the articles. One professor had six different readings assigned, and three of the article's means were rated below 1, suggesting that hardly anyone in the class even looked at the articles. When professors assigned only one or two pieces, the overall reported reading level ranged between 1.2 and 1.8. It appears students are more likely to read some of each piece when the overall number is lower. Of course, the number of pages actually read may be the same in both cases. In fact, our data is consistent with students generally starting the reading but losing momentum somewhere in the process. At the least, we can conclude students rarely completed all the reading assigned.

When we asked the students what the important factors were when deciding how much of the assigned material they read, they reported that of the listed options, "personal desire to learn" was the most important factor (average ranking was 7.3 on a scale of 1-10, 10 being most important). In addition, they noted that "time" (6.9), "the actual subject matter" (6.4), "the difficulty of the reading material" (5.6), and the "desire to participate in class discussion" (5.6) were variously important to their decision. A "sense of obligation" was only ranked 4.5.

We asked students to assess how important actually reading the assigned material was in terms of their overall learning in the course. Two students reported it was "very important." However, most (6) replied it was only "somewhat important" while two indicated it was "not important at all." While we can't rule it out, there is no particular reason to suspect that this is a peculiar finding that applies only to this course. At the least, it raises the question, at what point students develop the idea that reading is not essential to learning.

With time being listed as an important factor in students not reading, we checked to see if

these students had substantial time-consuming obligations besides school. Only a few reported that they worked full (2) or part (4) time, only one cared for dependents. In addition, we asked the students about their reading practices and were comforted to discover that almost all the students considered themselves to be “readers.” If they were unable to read easily or didn’t like to read, we might understand why, in practice, they didn’t read. However, five reported that they read “all the time—for enjoyment and to be well-read while the other half (5) reported that they read “occasionally” or “a lot.” While this data on their reading habits may be skewed (self-diagnosed), we merely interpret the results to show that, at least in principle, they have a positive attitude toward reading.

We asked students, what faculty should do to get good readers to actually read in the classroom. The students responded to a list of options, selecting ones they felt would make it more likely that they would read assignments. Those options that ranked highest were “give less reading (7),” and “use the reading better in the lecture/discussion (7).” Six indicated “providing a reading guide” would help, and five reported faculty should “pick better things to read.” Reading quizzes were not very popular (3).

We ended the survey with an open-ended question asking students to reflect on their reading behavior and why faculty may have difficulty getting students to read material carefully before class. The responses suggested a few things. While some students attributed the problem to “laziness on part of the student” or to the fact that the reading is “too boring,” others described the problem as more dynamic. One student reported, “It’s hard to get students to read, because we know that the material will be summed up in class anyway.” Another student echoed this sentiment saying, “I think some students think, since it will all be gone over in class that they don’t have to read it.” This suggests that students may very well strategize their reading decisions. As one student concluded, “students are smart and know what they are doing by reading or not reading.”

In summary, the faculty in our study had very different perceptions of what constitutes appropriate required reading for college students in an advanced course. While there is much to be said for a diversity of teaching styles, inconsistent load expectations may pose difficulties for both faculty and students. Students do form expectations about what “reasonable” reading loads may be (accurate or not), and in turn, they impose these expectations on faculty via their willingness to cooperate. While the assigned reading usually supplemented lectures, many instructors did not directly utilize the assigned reading in their classroom meeting in ways that we think would have enhanced support for student reading. In fact, some faculty strategies reinforced the benefit of not reading by repeating the material for them or by simply not actively using or drawing on the information in class.

The grading criteria for this class also appear to have failed to motivate students to adequately complete assigned reading. Since response papers were due after each weekly presentation, students could avoid completing the assigned reading by drawing mainly from the lecture and/or in-class discussion. When confronted with exams, students could then go back and revisit the texts (though it appears from our survey that many did not). This instrumental approach to reading (reading the bare minimum at the last possible moment) missed a primary point of the assignments—that point being to make class meetings (whether it be lecture, discussion, etc.) more productive.

For their part, the students in this course were faced with the challenge of adjusting reading strategies to a variety of academic subjects and kinds of texts. Even so, these were advanced students in an elective course. The end result, however, were compliance rates that

roughly approximated the rates for less skilled readers in general education courses (Burchfield and Sappington 2000, McDougall and Cordiero 1993, Self 1987, Marshall 1974). We found little evidence that this group had significant outside commitments that would restrict their time available to read. As a group, they did not complete the class reading even as they expressed very positive attitudes toward reading. Open-ended student responses suggested that many made a conscious decision *not to read* the assigned texts and there was some evidence that students did not view the reading as important to the pedagogical process.

IV. Discussion.

Admittedly, the course studied, comprised as it was of distinct lectures by different instructors, is not typical. We are not suggesting that this case study is generalizable. Nevertheless, we suspect that the peculiarities of this course allowed a unique glimpse into a pervasive problem with student reading, and it was helpful to see how instructors in a wide range of disciplines used assigned reading in their instruction. We found the lectures informative, but we were not always sure how the assigned reading enhanced them. We even felt frustrated when instructors spent a significant amount of time merely reviewing and summarizing material we, like the students, had been asked to prepare for the class. We sympathized with students when the lecture had no obvious connection to the reading, and understood how they might question why they spent the time preparing irrelevant material. However, when the reading material *was* drawn on in a constructive way, and students who appeared not to have read the material appropriated discussions in counterproductive ways, we also felt disconcerted. Consequently, we were left with the understanding that the role of reading assignments in learning constitutes a problem that has multiple sources. Our impressions were supported by student responses on the survey.

As we mentioned earlier, many writing on the issue suggest that the barrier to student compliance with reading assignments lies in the area of student preparedness. While they suggest there are things that educators can do address this barrier, they place the *source* of the problem squarely with the student. However, if their hypothesis were true, then we should expect skilled readers, like the students in our study, to comply in higher proportions than less skilled readers. However, we found our subjects to comply at roughly the same rate as less skilled readers. It is possible that the students in our study were not actually skilled readers and, more generally, advanced students are not acquiring reading skills as they advance in their studies. Perhaps this would not be too surprising: if students are not reading in their lower-division courses, then where are they learning to be skilled readers? Alternatively, it is possible that universities select for students who can succeed without doing assigned reading. In other words, the students who advance in college are just those students who were able to succeed in lower division classes without having read the assigned reading.

We are willing to grant that unpreparedness may have something to do with the problem of noncompliance (note the diversity of types of reading they are asked to do), but we gather from our case study that noncompliance is not simply a *student-centered* problem. Rather, we suspect that *faculty members* deserve our share of the responsibility as well. Our survey confirmed the point that many students do not understand the pedagogical role of assigned reading (Lowman 1995, Cannon and Newble 2000). In fact, our experience in this course was that much of the assigned reading did *not* have an overt pedagogical role; over half the faculty didn't even use the assigned reading in apparent way within their class time. Moreover, the way

the class was graded also failed to emphasize the importance of reading assigned material since students could still write successful essay responses to the lectures with substantial noncompliance. Also, many students in our survey choose not to read even when preparing for exams, choosing to base their answers on lecture material only. We might, then, forgive students for their perception that assigned reading is not essential to academic success.

In the course, we witnessed (among other things) what Bean (1996) has referred to as the “vicious circle” of the reading process (p. 134). We offer an example to illustrate this phenomenon. Suppose in order to teach Aquinas’ proofs for the existence of God, we assign the corresponding passages from *Summa Theologica*. How are we to use this reading? Do we expect the students to understand the arguments without further explanation? We recognize that this is probably too much to expect from the students, or worse, we suspect that too many students failed to read the assigned passages. Instead, we are likely to explicate the arguments in class and directly walk them through the text. Students, in turn, may simply not read, waiting for the instructor to cover the reading for them in class. We feel compelled to cover assigned reading material because we cannot assume that a majority of the students have read and understood the material. They see no reason to read if instructors will, as students sometimes put it, “tell them what they need to know” in class. Of course, there should be, and often is, direct discussion of the reading in class; the question is how to do it in such a way that we do not undermine students’ need to critically read on their own.

What is to be done about the problem of non-compliance once we understand that the problem is perhaps more dynamic than is often recognized? This study does not pretend to answer this difficult question. However, we do make the following observations.

(i) The relationship between non-compliance and preparedness needs to be better understood. Many assume that preparedness is the chief barrier to compliance. We feel that this may get the cart before the horse: the acquisition of reading skills depends, after all, upon compliance with assigned reading. Moreover, we feel that some of the suggestions for improving compliance--suggestions that view unpreparedness as the chief obstacle to compliance--may worsen the problem. For example, using class time for reading material with a high priority (Lowman 1995, Hobson 2004) simply exacerbates the vicious circle of the reading process. Students will surely be less likely to read outside of class if important materials will always be read in class. Solutions like aiming readings at remedial readers (Leamson 1999) or advocating that teachers assign less reading, perhaps skipping a textbook at all (Grunert 1997, Maleki and Heerman 1992) seems to us to devalue the role of reading in learning, reducing students expectations to read independently. A better understanding of the relationship between preparedness and compliance is needed before we can hope to discover real solutions to these issues.

(ii) We do not dispute that preparedness is at least a partial explanation for non-compliance. However, while there has been much research for improving student reading skills, dissemination of these innovative theories as well as *practical* teaching ideas are lagging. There has been much research on teaching reading skills within a classroom focused on reading acquisition (Dillard 2003, Nist and Simpson 1996, Nist and Diehl 1985). Still, researchers generally fail to explain how to integrate the strategies they recommend in courses where time allotted to reading skill development is very limited (Stahl, et. al. 1992). Improving reading skills is a tremendously resourceful field, but the format of the help (geared towards whole classes devoted to reading training) makes it much more difficult and less likely that a typical college professor will be able to draw easily from the material.

(iii) Our brief excursion into faculty teaching habits demonstrated in this lecture series was enough to convince us that faculty are clearly a piece of the compliance problem. It has been argued that many of the common ways instructors deal with the problem often creates more problems. For example, Bain (2004) argues that focusing on point accumulation fosters strategic learning rather than “deep learning” (p. 151). Assigning quizzes as a way to force students to read sets the wrong precedent for classroom learning. What faculty can do to resolve this issue is complex and full of wrong turns. As Bain notes, the best college teaching does not come from templates or ‘simple list of do’s and don’ts” (p.15). Unfortunately, college professors are simply not taught enough about pedagogy and teaching as a theory driven enterprise. More training in this area for all college instructors is probably needed.

(iv) Student perception of the reading-learning process is important. The second author of this paper has experimented with letting introductory sociology students engage the course’s textbook outside of class while reserving class time for exercises addressing more conceptual issues. What would happen if we held students responsible for material in a very accessible textbook without spending class time lecturing over it (she did reserve class time for questions regarding the text)? While students’ performance on this material was judged to be similar to the performance of students in classes with explicit lecturing on the material, students resisted the change. One student evaluation stated, “You turned this course into an independent learning course and didn’t do your job of teaching.” Some students were able to accept the new expectations but many were hostile toward them. So while the students generally appeared capable of learning from the textbook on their own, many did not feel that this expectation was reasonable. Ways of dealing the compliance problem will need to be sensitive to students’ attitudes toward reading.

(v) We end on a positive note. It should not be overlooked that the students in the course studied have a positive attitude toward reading; they did not hesitate to identify themselves as “readers,” whether or not this actually reflects their actual reading habits. Moreover, “personal desire to learn” was the most common reason cited for reading assigned material. In these respects, we hope that our students are representative. If so, there certainly is hope that given appropriate learning conditions—including maybe giving students a sense of “ownership” of the reading process-- students can be motivated to prepare the reading we assign them.

Our ultimate purpose in writing this paper is to call attention to the direct role faculty may have in perpetuating reading noncompliance in our students. It is easy to finger students as the source of the problem, ignore the role instructors play. Faculty need to ensure that we create the appropriate learning conditions that actually foster and reward the behavior we seek, and we feel the dynamic nature of the compliance problem is something that merits continued attention.

References

- Altman, H. B. and Cashin, W. E. (1992). “Writing a syllabus.” IDEA Paper No. 27, Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Bain, Ken (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard University Press.
- Bean, J. C. (1996). *Engaging Ideas: The Professor’s Guide to Integrating Writing, Critical Thinking and Active Learning in the Classroom*. San Francisco: Jossey-Bass.

Burchfield, C. M. and Sappington, J. (2000). "Compliance with Required Reading Assignments." *Teaching of Psychology*, 27(1), 58-60.

Cannon, R. and Newble, D. (2000). *A Handbook for Teachers in Universities and Colleges: A Guide to Improving Teaching Methods* (4th ed.). London: Kogan Page.

Dillard, M. (2003) *Reading in the Academic Environment*. New Jersey: Prentice Hall.

Erickson, B. L. and Strommer, D. W. (1991). *Teaching College Freshman*. San Francisco: Jossey-Bass.

Grunert, J. (1997). *The Course Syllabus: A Learning-centered Approach*. Bolton, MA: Anker.

Healy, J. M. (1990). *Endangered Minds: Why our Children Don't Think*. New York: Simon and Schuster.

Hobson, E. H. (2004). *Getting Students to Read: Fourteen Tips*. IDEA Paper No. 44, Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

Lang, F. and Gore, T. A. (1988). "A Study of the Reading Habits of Fourth-year Medical Students." *Journal of Medical Education*, 63, 567-568.

Leamson, R. N. (1999). *Thinking about Teaching: Developing Habits of Learning with First Year College and University Students*. Sterling, VA: Stylus.

Lowman, J. (1995). *Mastering the Techniques of Teaching* (2nd ed.). San Francisco: Jossey-Boss.

Maleki, R. B. and Heerman, C. E. (1996). *Improving Student Reading*. IDEA Paper No. 26, Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

Marshall, P. (1974). "How much, how often?" *College and Research Libraries*, 35(6), 453-456.

McDougall, D. and Cordiero, P. (1993). "Effects of Random-Questioning Expectations on Community College Students' Preparedness for Lecture and Discussion." *Community College Journal of Research and Practice*, 17, 39-49.

Nilson, L. B. (1998). *Teaching at its Best: A Research-based Resource for College Instructors*. Bolton, MA: Anker.

Nist, S. and Diehl, W. (1985) *Developing Textbook Thinking*. Lexington, MA: D.C. Heath and Company.

Nist, S. L. and Kirby, K. (1989). "The Text Marking Patterns of College Students." *Reading Psychology: An International Quarterly*, 10, 321-338.

Brost, B. D. and Bradley, K. A.

Nist, S. and Simpson, M. (1996) *Developing Textbook Fluency: Active Reading and Studying in the Content Areas*. Lexington, MA: D.C. Heath and Company.

Self, J. (1987). "Reserve Reading and Student Grades: Analysis of a Case Study." ***Library and Information Science Research***, 9(1), 29-40.

Stahl, N., Simpson, M. and Hayes, C. (1992) "Ten Recommendations from Research for Teaching High-Risk College Students." Retrieved July 12, 2005.
http://www.ncde.appstate.edu/reserve_reading/10_Recommendations_Article.htm.

Appendix A: Questionnaire

This survey is designed to assess your reading behavior in this lecture series and your perceptions of the reading assignments. Completing the survey is voluntary, anonymous, and will have no effect on your course grade. If you are willing to participate, please write the answer choice in the blank beside the question. We do appreciate your willingness to respond.

Below is a complete list of all the reading assignments given in this course. Please indicate, as best you can remember, how much of each assignment you read and how adequately you felt the material was utilized by the professor in the actual lecture.

0=none, 1=some, 2=most, 3=all

(NOTE: We are omitting reporting this section (questions 1-11). It simply lists particular faculty and the reading assignments they gave.)

12. In your own opinion, which of the faculty members who lectured in this course, if any, used the assigned readings most effectively?
13. In deciding how much of the assigned reading you were going to read each time, how important were each of the following factors in your decision? Rank each item from 1-10, with 1 meaning "not important" and 10 being "very important."
____ time, ____ difficulty of the material, ____ actual subject matter, ____ desire to participate in class discussion, ____ personal desire to learn, ____ sense of obligation, ____ wanted to please professor
14. In your own experience, how important is actually reading assigned material in terms of your overall learning in this class? Not important, somewhat important, very important, it was essential.
15. What could faculty do to make it more likely that you would read the assigned reading (put an X by all that apply)? ____ give less reading, ____ pick better things to read, ____ use the reading material better in the lecture/discussion, ____ provide a reading guide or questions to use as we read, ____ give points for reading or quizzes, ____ other, please specify: _____.
16. Outside of assigned course reading, which of the following best describes your reading lifestyle (put an X beside the comment that best fits you). ____ Never read unless I have to, ____ read some here and there, mostly comics, ____ read occasionally, things I want to read, ____ Read a lot, only for my own enjoyment, ____ Read all the time, to enjoy and to be "well read."
17. Check the statements that apply to you. ____ I work full time, ____ I work part time, ____ I care for any dependents (children, aged parents, etc.).
18. Finally, in your own words, explain why you think getting students to read material before class may or may not be a problem for faculty.