I. Project Summary

My classroom research project was inspired by something Ken Bain writes in *What the Best College Teachers Do*:

"Many outstanding teachers give comprehensive examinations with each test replacing the previous one. The first test covers material from the beginning of the course, but so do all subsequent examinations. . . . In such a system, students can try, come up short, receive feedback on their efforts, and try again on a subsequent examination. What they understand and can do intellectually by the end of the course matters more than anything else."
(Bain, *What the Best College Teachers Do*, p. 161)

The initial goal was to find out whether comprehensive testing stimulated students' self-monitoring practices in Intermediate Spanish (SPAN 201 at Luther College). I developed two research questions and a working hypothesis:

- Do students who are subject to comprehensive testing in Intermediate Spanish employ self-monitoring practices more frequently or to greater effect than students who are not subject to comprehensive testing?
- Which self-monitoring practices are employed by students who perform successfully in Intermediate Spanish when comprehensive testing is implemented?

Hypothesis: Comprehensive testing activates and strengthens self-monitoring practices by recycling course content and goals repeatedly and through frequent feedback on performance.

For the past 2-3 years I have experimented with comprehensive testing in Intermediate Spanish classes, based on the strong hunch that such testing promotes deep learning and leads to better retention. Research on memory recall and testing suggests that frequent testing and repeated retrieval from memory lead to better performance on tests (see Karpicke and Roediger). My goal was to see if this held true in Intermediate Spanish and to see if there might be a relationship between comprehensive testing and metacognitive self-monitoring. Might one of comprehensive testing's benefits for learning and retention be that it stimulates reflective practices, such as self-testing, which in turn lead students to perform better on subsequent tests?

II. Context

This study was carried out during the 2009-10 academic year in four sections of Intermediate Spanish (SPAN 201). Of the two sections I taught each semester, one section served as a control group while the other received the intervention (comprehensive testing). Thus, there were two control groups (38 students) and two intervention groups (40 students). Students in both groups took the same number of tests over the course of the semester; however, the intervention group received tests that were comprehensive: each test measured learning of recent material as well as of all previous material. All students took the same comprehensive final exam. In order to measure performance in the course, I collected data from all tests and the final exam, and students took a standard Spanish placement test at the beginning and at the end of the semester as an indicator of their proficiency level. Self-monitoring practices were measured with the Metacognitive Self-Regulation (MSR) subscale of the Motivated Strategies for Learning Questionnaire (MSLQ). Students completed the questionnaire twice during the semester, once after the third test and again after the final exam. All students received an explanation of the classroom research and signaled their voluntary participation via a consent form.

SPAN 201 at Luther College is a course for students who have completed at least two semesters of college-level Spanish or the equivalent. Students in this course practice vocabulary and structures they have learned previously with the aim of developing an intermediate-low level of language proficiency. Specific course goals include being able to talk about significant events in the past; formulating reactions to and recommendations for environmental problems; giving directions; and describing future plans and career aspirations.

III. Teaching Practice

The metacognitive intervention in this experiment was comprehensive testing. Outside of a brief presentation of the research project when students were given the consent form, I did not explicitly address metacognitive knowledge or self-monitoring practices with students during the semester. My goal was to determine the extent to which comprehensive testing itself activated or strengthened students' self-monitoring practices, so I did not to talk very much with students about employing self-monitoring practices toward better performance. The course syllabus contains a list of strategies for success in the course, some of which are oriented to metacognition, but aside from calling students' attention to these strategies at the beginning of the semester and again at mid term, I did not regularly address self-monitoring skills in class.

Over the course of the academic year and the two iterations of the study, it became apparent that comprehensive testing by itself – unaccompanied by discussions of metacognitive skills or explicit reflection on the relationship between course performance and self-regulation – did not appear to trigger any significant changes in students' self-monitoring practices. A few students in the intervention group commented on the learning benefits of comprehensive testing in their end-of-term course evaluations, but students who were engaged in comprehensive testing during the semester did not appear to think differently about tests or course performance than students in the control group. Although I began to notice the lack of difference between the intervention group and the control group in the first semester, I decided not to change the experimental design in the second semester in order to end up with a larger sample size for the study.

IV. Evidence and Results

Three types of data were collected for all students in this study. First, students completed a standard language placement test once at the beginning of the term and again at the end of the term in order to obtain an indicator of their language proficiency. Second, student test and final exam scores provided a measure of course performance. Finally, students completed a survey of questions from the Metacognitive Self-Regulation subscale (MSR) of the Motivational Strategies for Learning Questionnaire (MSLQ), an assessment tool developed by Bill McKeachie and Paul Pintrich at the University of Michigan.

Below is a summary of the results of comparing course performance data to data from the MSR subscale. Since it appeared that there were no significant performance differences or differences in self-monitoring practices between the two groups, I did not perform a rigorous statistical analysis. Statistical analysis of this kind of data is also beyond the scope of my abilities at present.

a. **Metacognitive Self-Regulation (MSR)**

- MSR increased among all students, regardless of comprehensive testing. The average increase in MSR for in the control group (4.38%) and in the intervention group (4.75%) was nearly the same.

- In the control group MSR increased on all but one question in the subscale. The questions that saw the greatest increases were:
  MSR #2: "If course materials are difficult, I change the way I study." (16.93%)
  MSR #5: "I ask myself questions to make sure I understand material I have been studying." (13.41%)
- In the intervention group MSR increased on seven questions and decreased on three questions. The questions that saw the greatest increases were:
  MSR #4: "Before I study new course material, I often skim it to see how it is organized." (9.15%)
  MSR #8: "I often find that I have been reading but don't know what it was all about." (reverse coded) (10.13%)
- Among first-year students, the intervention group saw a higher average increase in MSR (6.58% vs. 1.01%). Among non-first-year students, however, the control group had a substantially higher average increase in MSR (9.44% vs. 0.54%).
- Supposing that there were no major differences in student make-up or teaching methods between the two groups, and that the only substantive difference was comprehensive vs. partial testing, the data do not show any clear influence of testing method on metacognitive self-monitoring skills.

b. **Student performance (tests, final exam, placement test)**

- The control group outperformed the intervention group on nearly all exams, including the final exam. The intervention group did, however, show improvement in exam performance as measured by the change from the last test to the final exam (2.49% intervention vs. -0.13% control).
- Comprehensive testing may have resulted in more difficult tests for the intervention group and thus lower scores. It is worth noting that the average test scores for the intervention group were not far behind the average scores in the control group.

c. **MSR relative to performance**

- Increased MSR scores did not necessarily correspond to better performance on the final exam or on the placement test in either the control or the intervention group.
- In the control group an increase in placement test performance often corresponded with an increase in MSR. The opposite was true for the intervention group, though, where the largest increases in MSR corresponded with diminished placement test scores.
- Improved MSR did not necessarily correspond with improvement in test scores as measured by the change from the last test to the final exam, nor did improved test scores necessarily correspond to increased MSR. The change in test performance was almost none for the students with highest gains in MSR, both in the control and intervention groups.
- High performance on the final exam did not correspond to a higher average gains in MSR.
- Among all students, the average increase in MSR went up slightly among lowest performers on the final exam (6.82% for those with grades of C or below on the final exam).

V. Conclusions and Implications

Although I was encouraged to see increases in self-monitoring practices and language proficiency (as indicated by placement test scores) among all students on average, there seems to be no correlation between comprehensive testing and increased self-monitoring skills. Comprehensive testing did not appear to do any harm, but its relationship to self-monitoring skills is uncertain. Research on testing frequency and retention suggests that testing can strengthen the ability to store and retrieve information in long-term memory, thus

tests can serve as learning tools as well as assessments of learning. With this in mind, I plan to continue using comprehensive testing in my language classes and look for different means of helping students develop their metacognitive skills.

Working with colleagues in the ACM-Teagle Collegium on Student Learning has been an exciting foray into metacognition, classroom research and the scholarship of teaching and learning. The Collegium group has facilitated much good sharing of teaching practices that help students think about and monitor their own learning, and as I developed by own project on metacognition my colleagues in the Collegium have provided encouragement and valuable feedback at several points during the past year and a half. Their help was especially welcome, since this was my first classroom research project and, as a scholar in the humanities, I am generally unfamiliar with metacognition and research in cognitive science.

VI. Looking ahead

In addition to continuing comprehensive testing in Intermediate Spanish as a means of helping students retain knowledge and learn deeply, the most significant change I plan to implement is explicit attention to self-monitoring practices in classes. For example, this semester (Fall 2010) I have implemented a post-test reflection (test wrapper) that students complete as they receive a graded test. Just prior to the next test in the course, I will return their test wrapper and talk briefly with them about their stated plans to change study strategies or preparation methods. My hope is that through such reflection activities and consideration of self-monitoring in class, students' metacognitive skills will improve.

In Fall 2010 I intend to complete a new round of study in two sections of Intermediate Spanish in hopes of seeing a correlation between reflection exercises (wrappers) and increased self-monitoring practices. I would like to use the results of this study as part of a course portfolio on learning in Intermediate Spanish that I will complete in Spring 2011.

VI. Bibliography

Bain, K. (2004). *What the Best College Teachers Do*. Harvard University Press.

Cross, K. P., & Steadman, Mimi Harris (1996). *Classroom Research: Implementing the Scholarship of Teaching*. Jossey-Bass.

Isaacson, Randy M., & Fujita, F. (2006). Metacognitive Knowlege Monitoring and Self-Regulated Learning: Academic Success and Reflections on Learning. *Journal of the Scholarship of Teaching and Learning*, 6(1), 39-55.

Jacobs, L. C., & Chase, C. I. (1992). *Developing and Using Tests Effectively: A Guide for Faculty*. Jossey-Bass.

Karpicke, J. D., & Roediger III, H. L. (2007). Repeated Retrieval during Learning is the Key to Long-term Retention. *Journal of Memory and Language*, 57(2), 151-162.

Ross, M. E., Green, S. B., Salisbury-Glennon, J. D., & Tollefson, N. (2006). College Students' Study Strategies as a Function of Testing: An Investigation into Metacognitive Self-Regulation. *Innovative Higher Education*, 30(5), 361-375.

Weimer, M. (2002). *Learner-Centered Teaching: Five Key Changes to Practice*. Jossey-Bass.

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into Practice*, 41(2), 64-72.