Teagle Assessment Project Report for English Literature, 2010-11

Coordinators: Craig Williamson, Peter Schmidt

June, 2011

Student Learning Goals

In Department meetings in 2009-10, we agreed to focus our Teagle assessment project on our first-year introductory seminars in literature. These are not only key courses in generating majors and minors in English Literature but play an important part in the Swarthmore curriculum as a whole. Students with interest in the natural sciences, the social sciences, and the humanities take these courses and we want to make sure that we improve the majority of students’ writing skills long with their ability to think profoundly about literary and cultural history. More than ever before, however, high schools dramatically vary in quality when it comes to giving feedback and basic support to students about their analytical writing skills.

For the purposes of the Teagle assessment plan, we defined the primary goal of introductory first-year seminars (FYSs) in English Literature as follows: to help all students enrolled make a successful transition from high school to College-level writing about literature.

Student Learning Objectives

After extensive discussion in Year 1 of the Teagle project, the English Department came up with the following objectives to be emphasized in all our FYSs – 7 taught in fall 2010 and 5 in spring 2011. (Of the spring seminars, 2 were repeats from the fall; 3 others were “new” FYSs taught by other colleagues.) The writing objectives are included in all first-year syllabi and/or handouts and are emphasized in class assignments and discussions. These objectives are also the focus of our direct assessment strategies. We focused on them because while improving these analytical writing skills are key to crafting strong papers for English literature classes, the skills learned are also transferable to any kind of analytic writing that needs to argue carefully from textual evidence. We would also like to add that the discussion of what our shared writing objectives should be and how we might teach with these objectives in mind has so far been the most valuable aspect of our participating in the Teagle assessment project.
Here is our list of seven primary objectives involving writing for our students in first-year English Literature seminars at Swarthmore:

**Writing Objectives:** Students completing English Literature First-Year Seminars should improve their ability to …
- develop an interesting, specific, supportable thesis
- marshall an argument that is logical, well-developed, and compelling
- support arguments with textual evidence carefully analyzed
- consider alternative readings or counter-arguments
- when appropriate, use criticism, theory, or cultural backgrounds to support the paper’s claims
- craft a conclusion that summarizes and offers new reflections
- use appropriate diction, tone, grammar, spelling, and punctuation

**Strategies to Achieve the Seven Writing Objectives**

The Swarthmore English Department gives individual professors significant freedom in designing their own first-year seminar syllabi, so long as the writing objectives are clearly announced in the course materials and emphasized in the course instruction. Before the Teagle assessment project began, we already had a tradition of sharing successful teaching strategies for improving student thesis sentences, use of quoted evidence, etc., via a shared archive of teaching resources (such as in-class handouts), department meetings specifically devoted to pedagogy and practice, etc. But the Teagle project has definitely heightened our emphasis on these objectives. It has spurred us to organize several meetings per year devoted to sharing strategies for improving student writing. As will be clear in our recommendations for Year 3, in the final year of the Teagle project we will emphasize even more thoroughly the importance of holding regular discussions to share resources and strategies regarding meeting these objectives, for our colleagues find this by far the most useful element in the entire assessment process; it’s a good morale booster and community-building exercise as well.

**How Direct Assessment is Used**

After extensive discussions in Year 1, the English Department decided to have a 3-pronged direct assessment approach that would allow us to compare the professor’s ratings of an early and a later paper in the semester to see what degree of improvement, if any, were found in the students’ writing regarding the 7 identified objectives. We decided to get data on all students in all 12 English first-year seminars from fall 2010 to spring 2011.

First, we decided to have all professors teaching first-year seminars select from each student an early paper and a late paper for assessment. (Students typically write four 3-6pp. analytical papers for these classes, though other forms of writing are involved as well, such as blogs or informal reaction papers, for a
total of at least 20pp. of writing, some of it revised by the student under the
guidance of the professor and, in some cases, a trained student Writing
Associate.) Aside from the normal marking up, commenting upon, and grading
of these 2 sets of papers, the professors would rate the papers using a 1-7
quantitative scale (poor to excellent) for each of our 7 writing objectives.

Second, we set up a blind assessment process online for each course, the
better to create a measure of objectivity, as well as allowing us precisely to
measure the degree to which, in the professors’ assessments, their students’ skills
had improved or not. This system allowed us to collect individual course data
but also to aggregate and analyze data involving all English FYSs from two
semesters, fall 2010 and spring 2011. Professors rated their first batch of papers
online, using a system helpfully set up for us by Robin Shores of Swarthmore
College’s Office of Institutional Research. Once the first ratings were done, they
would not be available to the professor doing the ratings for the later paper in
the semester. After each professor’s grades were turned in, the paper #1 and
paper #2 ratings would be available to her or him and the department chair for
analysis.

A third element is not direct assessment but is an important part of the
process. At the end of each semester we would have all students assess whether
they believed they had improved or not in each of the seven categories via
similar online forms. Then we could cross-tabulate student self-assessments and
professors’ assessments. This approach required the students to log in to the
assessment site after they had completed their semester’s work, a procedure that
was only partly successful. Despite several reminders, approximately 50
students out of 69 in the fall completed the online assessment; 35 of 47 in the
spring.

Analyzing and Sharing Our Findings from Year 2 Direct Assessment
The data from fall 2010 and spring 2011 are presented as Appendices A
and B and the end of this report. We will restrict ourselves here to focusing on
three conclusions from the data.

In aggregate, all papers showed improvement in all 7 objectives,
according to the professors. For instance, the average score on the quality of the
thesis in paper #1 from our 7 FYSs in the fall was 4.09; this improved to 4.71 in
paper #2. The spring scores re a thesis were even more dramatic: 4.89 for #1,
then 6.13 for a paper later in the semester.

Further good news comes when the range of replies from “stayed the
same” to “improved 1-6 points” are studied. In all categories but the thesis over
50% of papers improved, and if “stayed the same” is added to “improved,” the
numbers are quite respectable: in the fall, for instance, according to the
professors, 57.4 percent of papers improved their use of counter-arguments,
while 33.8 percent stayed the same, for a total of over 90%. Obvious room for
improvement remains, particularly if we combine the “stayed the same” and the
regressive (“points dropped”) numbers in all categories. Nonetheless, the overall fall numbers indicate, we believe, that we were doing something right—actually, more than a few things right.

Student self-ratings concur. Regarding the 7 objectives, in both the fall and spring numbers none are lower than a 5 on a 7-point scale where 7 = excellent or “very much improvement,” and many student ratings are closer to 6 than to 5. The lowest rating, a 5.0 in the fall student self-reports for “craft a conclusion,” presents a question for discussion as to why this objective received the (relative) lowest rating. But perhaps it’s an indication that less time was spent instructing and showing strategies for writing conclusions.

Also positive is the fact that the numbers in the spring seminars (Appendix B) are generally compatible with the fall numbers, with no huge negative variants or numerical anomalies cropping up between fall and spring.

Two other details may be noted here. First, the spring data was not ready in time to be discussed at our English Department retreat in mid-May 2011, since the semester had only recently been finished and many professors had not completed grading. (The data has now come in, however, in time to be included in this report. It will also be shared with our colleagues in English in the fall of 2011.)

Second, we are aware that not all our categories are strictly comparable. In particular, the students’ self-ratings presumably cover their work for the entire semester, whereas by design the professors’ ratings focus just on 2 different papers. Still, those 2 papers represent a fair sampling of each student’s work and thus provide a reliable snapshot of student progress or the lack of it.

There are many reasons that we might not see a relationship between student and faculty ratings, including the ones just mentioned plus of course the fact that the students are rating themselves whereas the professors are rating others. Professors’ ratings tend to be lower by about a point on our 7-point scale than the students’ self-ratings. Robin Shores has drawn our attention to one interesting correlation between the student and professor ratings in the spring FYSs, however. For objective #2, “Marshall an argument that is logical, well-developed, and compelling,” Robin finds that there is a statistically significant correlation of 0.48—a “moderate” correlation, she tells us, but a correlation nevertheless. That means the students and faculty members were in some agreement about whether or not the student improved on this dimension, which is perhaps the most important objective of all, since it focuses on a paper’s entire argument.

Re correlations, our own analysis would emphasize the bottom line: both students and professors saw improvement in all the categories, though there was some disagreement about how much and at what level the improvement occurred. We are certainly somewhat heartened by this pattern, though as will become clear from the discussion below we’re not sure that the numbers are giving us important new information. And we worry a bit that an inherent
yearning in both students and their professors to see improvement may inevitably warp somewhat our numerical results.

At the English Literature department retreat in May 2011, the Teagle assessment data from the fall FYSs was gone over in detail among all colleagues. We had a lively discussion of what conclusions were valid to draw from this data, versus what inferences we should definitely not be certain about—conclusions and cautions that I’ve briefly summarized above. We also discussed the issue how to understand that in some cases papers either failed to improve or regressed in some of the targeted categories. For instance, according to professors rating their fall student papers, 50% of second papers did not improve in their use of textual evidence over the first paper in the sample. In the spring, the improvement numbers are lowest in two other categories, “craft a conclusion” and punctuation, grammar, etc. Two explanations seem possible: either our instruction needs improvement in these areas, or students are trying more sophisticated arguments in their later papers and as a consequence may not have the same levels of success. The real answer is probably a combination of both of these explanations. We will certainly focus further next year on improving our instruction related to all 7 objectives for writing improvement. It’s also possible that the second of the student papers in some cases came at the very end of the term close to finals, when typically there is some slippage in student writing because of the press of other work.

When discussing assessment at our May retreat, we spent most of the time assessing the assessment process itself. Unfortunately, the majority of colleagues doubted whether the “payoff” from the data was enough to warrant all the complicated procedures in the online rating system, for these in effect required a second grading or direct assessment process for each paper twice during the semester. In general, we felt the numbers didn’t tell us much that was startlingly new, either in terms of significant successes or failures. Further, we found the focus on defining teaching objectives and strategies much more fruitful than requiring that student papers be judged using a 7-point quantitative scale. We are grateful for a year’s worth of quantitative data regarding our objectives, but no colleague was enthusiastic about gathering more such data for next year, not to mention years and years after that. We definitely think we should stop pressuring our FYS teachers in the name of Teagle to do online quantitative ratings that were both complicated to do and of questionable real value when it comes to assessing ways of learning that are most characteristic of the humanities.

This leads us to the next topic.

Recommendations for English Department Assessment Practices, Year 3

At our retreat we came to these tentative conclusions re assessment in 2011-12.
The online form we developed might be worth trying by other departments to see if they would like to use this process for quantitatively assessing the teaching of writing, but in general our department will stop using quantitative assessment in favor of other strategies—namely, gathering and creatively using a teaching archive of handouts and other in-class exercises focused on our 7 writing objectives. We believe such an approach will be the most help to us as we strive to have our first-year students meet all the 7 writing objectives of our FYs.

More positively, the members of Swarthmore’s English department were unanimous in concluding that our Teagle-generated discussion of shared objectives for our seminars was invaluable and is indeed directly related to our focus for Year 3.

In Teagle Year 3 we will grow our online, sharable archive of teaching resources involving 1) in-class handouts relevant for our 7 objectives, and 2) samples of student papers from our courses that represent examples of good and poor writing related to the 7 categories. These selected early and late student papers will give us examples to study and to mine for future teaching exercises. All student work in our archive will of course also be altered to ensure the students’ anonymity. And the archive will be accessible only to members of our English department.

Among our colleagues, there was definitely interest in growing and using such an archive. It’s valuable as individual professors plan syllabi and class lessons. It can also be used in department meetings when we discuss pedagogic strategies for improving how we impart writing skills. We often use “sample” good and mediocre thesis sentences or quotation-and-analysis examples in our teaching, for it’s important for students to work with examples, not just learn general principles. (For instance, it’s very effective after discussing what a good thesis sentence does to ask students to work in groups to revise and improve poor sample sentences.) The same goes for all the 7 writing objectives listed.

We will keep the list of the 7 objectives in the printed student course evaluation forms that we use for all our FYs, so we won’t entirely lose getting student feedback on whether or not they felt their writing improved in all the targeted areas. Our completion rate for in-class student course evaluation forms is much higher than for the online forms used for the Teagle Year 2 project.

Though we are dropping the online surveys, we are also going to test in a comparative way the use of handouts in FYs focused on some of the 7 writing objectives. This comparative testing will be the key component of our direct assessment strategy for Year 3. Different teachers will choose to use one or another of the handouts (and hopefully discuss them in class), or no handouts. Then from the students’ own evaluation forms at the end of the class (where we ask for their evaluation of our teaching of the 7 writing objectives) we should be able to see what the difference is in the students’ minds between using handout A, handout B, or handout C as opposed to using no handout. We will meet to discuss the
results at the end of the school year, when we will also consider what next steps we’d like to take, such as perhaps applying these writing improvement assessment strategies to other courses and seminars in our curriculum.

**Discussing English Department Teagle Assessment Strategies with Others**

Craig Williamson and Peter Schmidt have discussed our Department’s assessment strategies, experiences, and results with our Teagle colleagues at fall and spring meetings in 2010-11. We have also shared our preliminary results with a wide range of Swarthmore colleagues from all disciplines at a spring 2011 faculty lunch devoted to new “direct” assessment strategies. We will be happy to share our reflections on our assessment procedures, including the Year 3 results, after 2011. And as we said in our previous section, we can guarantee that assessment projects and discussions about how to improve effectiveness will continue among us during future years.