

Annual Assessment Report Summer 2006

Introduction

Sterling College's current assessment program began 6 years ago. We designed the program to be clear and simple, and to provide faculty with data that permitted them to make instructional improvements. Students in most academic programs have clearly benefited from such improvements. Students in a few programs have not yet benefited, as some faculty have resisted the reflective and adjusting exercises the data suggest.

Assessment data continue to reflect the high quality of Sterling's general education and academic major programs. Yet, in every program there is room for improvement. Assessment is a continuous task. After changing a program, faculty again need to evaluate whether the change generated improved student learning outcomes. This ongoing modification, improvement, evaluation cycle is the hard work and the beauty of assessment. As faculty continue to seek meaning in the data and commit themselves to improving student learning through careful data evaluation and instructional improvements, the value of assessment activities will increase.

This assessment report documents data collected from students and faculty during the 2005-2006 academic year. The report presents the following information:

- ◆ Compliance rates
- ◆ Institutionally aggregated data on knowledge, skills, and values for graduates
- ◆ A description of "Success Stories" in assessment, highlighting departmental achievements during 2005-2006 produced by assessment processes.
- ◆ Departmentally aggregated data on knowledge, skills, and values for graduates

The Assessment Committee has examined and approved the full report. The Committee delivers that full report to the President and the Vice President for Academic Affairs. A shorter report, excluding the departmentally aggregated data, is public information and is distributed to all departments through their department chairs and to students through Student Government Association. The public report is also available on the Sterling College web page. Each department chair or assessment director for a major also receives a copy of his/her departmental report.

Compliance Rates

During the 2005-2006 academic year, 75 students graduated with 19 majors from 14 departments. All 14 departments submitted assessment reports. Some assessment data were missing for the one Independent Interdisciplinary major which has no single assigned department. Almost all academic units are responding to assessment data collection assignments.

Compliance for students is different than it has been in past years. All students entering the college in 2002 or later were required to complete assessment assignments

for graduation. Almost all of the 2005-2006 graduates fit this category. In past years, even without required participation, students responded cooperatively to give us compliance rates at or above 95%. This year, of 405 data items we should have collected, only 23 were missing. Reasons for missing items fell into a few categories as indicated in Table 1.

Table 1. Reasons for missing data

Reason	Number of cases
Entered before 2002. Chose not to comply	4
Earned double major; produced single product	3
Department failed to collect data	13
Product pending completion of experience	2
Excused for disability	1

Only 2 of the reasons for missing data are under institutional control—students completing double majors but providing a single product and the departments failing to collect the data. The Assessment Committee has decided that students completing double majors should demonstrate their skills for each major. This expectation should be clarified. Most of the double majors are in closely related departments. Some students use the same seminar experience for creative products for both majors. The college may want to review whether double majors are justified within a single department, or within content areas that overlap greatly. This issue will emerge more frequently in the future as the college adds majors within departments, as it has done for Exercise Science. The second controllable reason for missing data is departmental failure to collect required data. This issue needs to be resolved. Some departments have demonstrably inadequate systems for collecting their data from students. Prior assessment reports have noted these problems, and departments must adapt their procedures to effectively obtain necessary data.

The missing data fell into 4 categories—Departmental creative product scores, creative product self-assessment scores, student ethics instruction reports, and general education tests. Most of the missing data involved student self-reports (items the departments failed to collect above) of creative products (7 items) and service projects (7 items). Three students did not complete their general education test (CBASE). One of these was excused due to disability. The other 2 chose not to comply and had entered college prior to 2002. Two students did not complete the ethics instruction form. These 2 were the same students who missed the CBASE test. We collect ethics forms at the same time we administer the CBASE. Four students did not have faculty evaluation scores for their creative products. Two of these were double majors who completed a product in only 1 department. Another was the Independent Interdisciplinary major, for whom the creative product assignment was mishandled. The fourth person graduated, but will complete her creative product in the context of her student teaching assignment in fall.

Compliance problems are minimal. There are several correctives that will improve compliance, and the departments which do not effectively collect student data need to correct these problems. Furthermore, faculty should consider the larger issues related to double majors. If the problem is only in refining data collection procedures, we

can easily correct it. If there are larger issues about what work qualifies for a double major, those will take greater deliberation to correct.

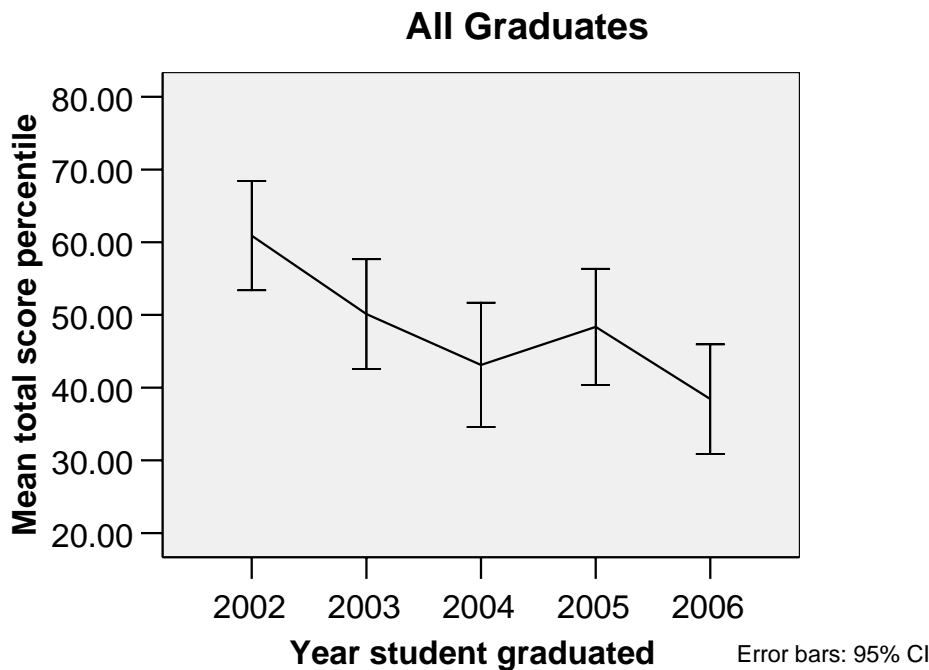
Institutional Data

Majors

Knowledge

Nationally standardized major field test scores for 2006 are available for 11 departments. Figure 1 shows mean percentiles and 95% confidence intervals for major field test scores earned by Sterling College students for the last 5 years.

Figure 1. Means and confidence intervals for major field test scores



Last year's report suggested that a downward trend for the prior 3 years did not look like an item of concern. However, the 2006 scores are considerably lower than the 2002 scores. The confidence interval lines for 2006 do not overlap those for 2002, suggesting that these are truly different scores. Analysis of variance confirms that 2006 scores are significantly lower than 2002 scores, $F(4, 284) = 4.66, p < .01$. We have reasons for concern about instruction in the majors for our students. Individual departments should examine their cumulative scores across the 5 years of testing and if scores or sub scores are consistently low, should plan corrective curricular modifications.

For scores from all years to date, averages for departments range from the 24th to the 75th percentile. Aggregating scores across years eliminates extreme scores that arise from very low sample sizes for a single year. One department—the one at the 75th percentile—had only 1 student, making that high percentile non-representative. The next highest average percentile for a department was 68th.

Sub scores are available for 7 departments for 2006 graduates. Sub score variation can help a department identify weak and strong areas of their curriculum. If the range of sub scores is high and persists across several years, the department can assume that its low scores are targets for curriculum improvements. For 2006 data, sub score variation ranged from a low (indicating balanced instructional emphasis across areas of a discipline) of 5 percentile points to a high (indicating less balance) of 30 percentile points. This range of variation is reduced when data from all years of testing to date are included. With data aggregated across years, sub scores varied from a low of 2 to a high of 11 percentile points if the department with only a single student is omitted. These data suggest balanced instruction for these departments, though the quality of the instruction may still vary as some of these departments have low and some have higher average percentiles. Each department that has sub score reports should examine its range of sub scores carefully, considering the following:

- ◆ Large sub-score variations indicate that students are not learning some of the discipline's content knowledge. Intentionally or inadvertently, the department may have left this content weak in its curriculum.
- ◆ Small sub score variation, if all the sub scores are high, indicates balanced curricular offerings encompassing the range of content in the discipline.
- ◆ Small sub score variation, if all the sub scores are low, indicates general weakness either in the students or in the curriculum design.
- ◆ Scores aggregated across all years of testing provide a more reliable indicator of program balance.

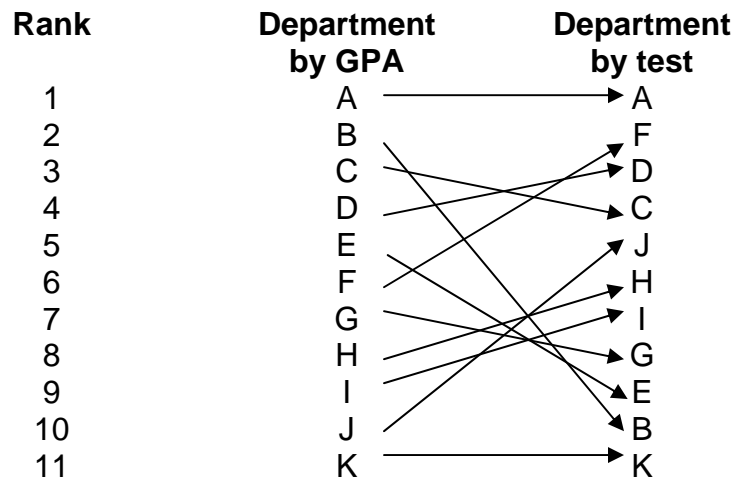
Major field tests and grade point averages independently assess content knowledge in a discipline. Faculty should expect the measures to be correlated, especially if the major field test content coverage corresponds with the department's curriculum design. For 2006 data, that correlation was significant, $r(68) = +.31, p < .05$.

Aggregating data across all years of testing provides more reliable estimates of the real relationships between GPA and test scores. For all graduates ($N = 298$), the relationship between GPA and knowledge test scores was significant and positive, $r(296) = +.47, p < .01$). This correlation accounts for 22% of the variation in GPAs. Seven of 11 departments showed significant positive correlations. A positive correlation means that students who scored high on the major field test more likely had higher GPAs. Cumulative departmental correlations ranged from $r = +.1$ (low and not significant negative relationship) to $r = +.71$ (strong positive relationship). Five departments had significant correlations ($p < .05$) that were higher than $r = +.4$.

We can examine inter-institutional consistency of grading practices by comparing average departmental GPAs and average departmental major field test percentiles. If all departments have similar grading standards, and grading depends on knowledge of discipline content, we should see that the rank order of grades corresponds with the rank

order of major field test score percentiles. We present these ranks for 11 departments for which these data are available in Figure 2. Non-horizontal arrows indicate changes in rank from one measure to the next. Rising arrows indicate departments in which students earn lower grades but score higher on the major field tests. Descending arrows indicate departments in which students earn higher grades, but score lower on the major field tests. Readers should remember that average GPA will reflect more than just content knowledge.

Figure 2. Comparing ranks for average departmental GPAs and standardized knowledge test percentiles.*



*Each letter represents a given department. Data represent departments for which major field test scores are available through 2006.

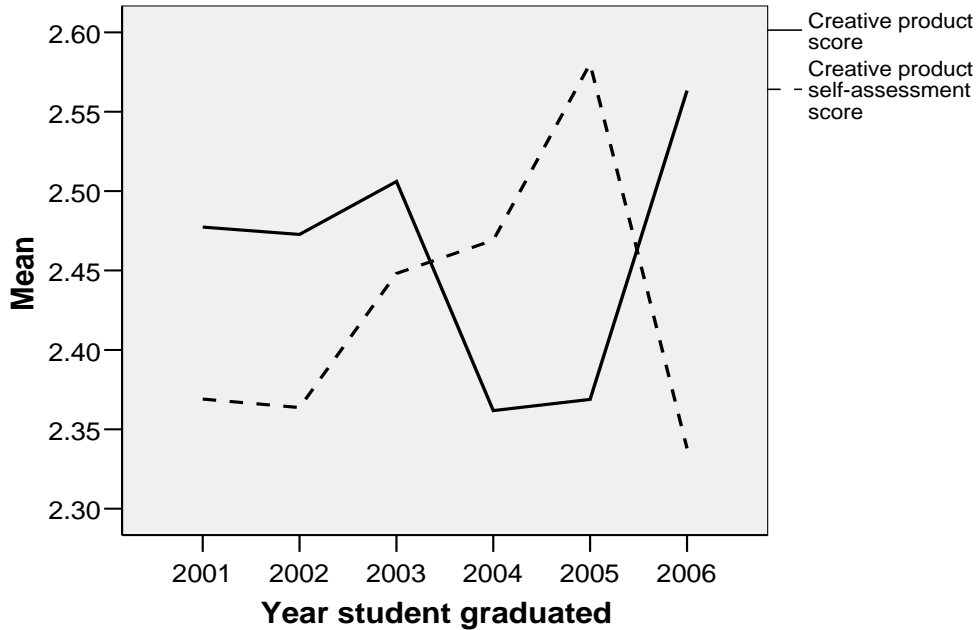
Skills

Sterling College assesses skills in the major by asking faculty and students to rate creative products that students complete late in their academic programs. All raters assign numerical scores to the products on the following 4-point scale:

- 0 = product did not reflect essential skills for a professional
- 1 = product minimally reflected essential skills for a professional
- 2 = product moderately reflected essential skills for a professional
- 3 = product clearly reflected essential skills for a professional

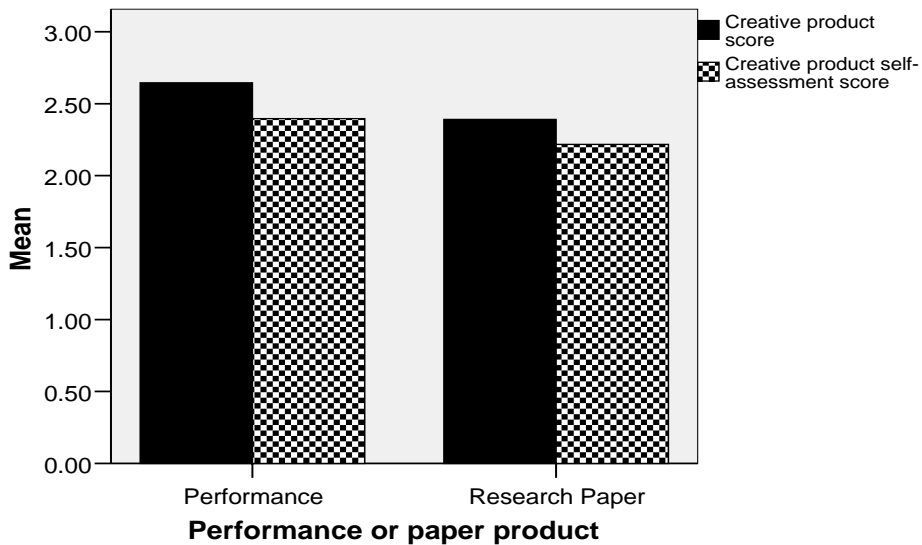
In 2006, faculty rated products higher (Mean = 2.52) than did students (Mean = 2.34). These scores represent a reversal in direction between faculty and student ratings from the prior 2 years. Figure 3 shows faculty and student ratings by year since 2001.

Figure 3. Mean Creative Product Evaluations by Year of Graduation



We classified creative products into 2 categories. Some departments require students to complete performance products. These include music recitals, gallery displays, and teaching plans. Other departments require students to complete research papers typical of professional papers in their disciplines. For 2006 creative products we evaluated faculty and student ratings by type of product. Figure 4 shows that faculty and

Figure 4. Mean Creative Product Ratings Based on Type of Product



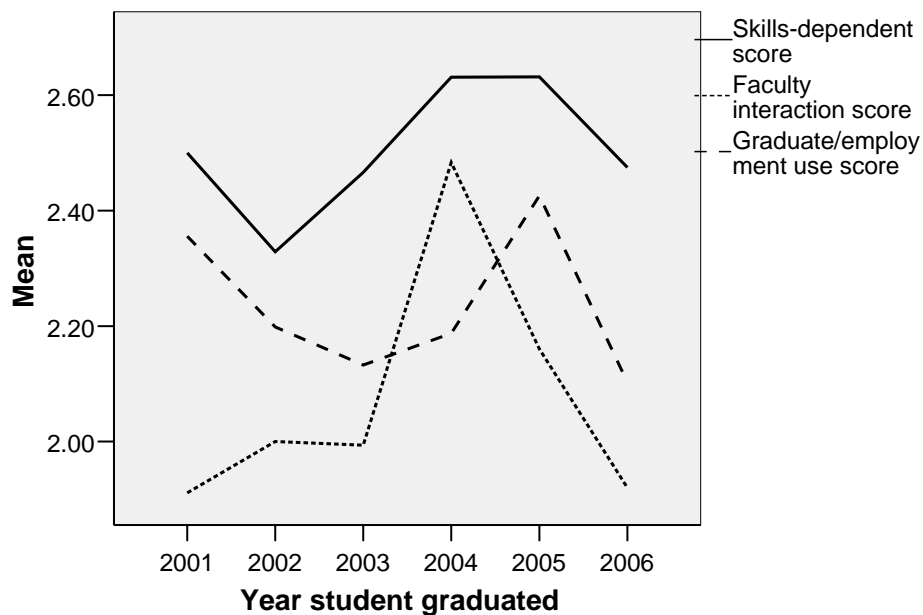
student ratings for performance products were higher than those for research papers. The difference between product types was significant for faculty ratings, $F(1,75) = 5.18, p < .05$, but not for student ratings, $F < 1$.

Departments vary in the scores they give for creative products. The average score for all departments for all years to date is 2.44. The average score by department ranges from a low of 1.72 to a high of 2.70.

We expect creative product scores to correlate with GPA in the major. An analysis of 442 students since 2001 shows a positive correlation between major GPA and creative product score, $r(440) = +.42, p < .01$. Separate correlations for each of 13 departments and the Independent Interdisciplinary majors showed non-significant correlations for 4 departments and significant positive correlations for 9 departments. The significant positive correlations ranged from $+.26$ to $+.76$.

Students answered several questions about their creative products, including whether the products reflected skills they had learned in their major, how much they had interacted with faculty in developing the product, and how likely they would be to show the product in an interview for employment or graduate school. Figure 5 shows average responses to these questions for each year. The trend for all questions before 2006 was a

Figure 5. Student ratings of creative product aspects by year of graduation



general increase in scores over the years. That trend was broken in 2006 as all scores declined below the prior 2 years. This decline may reflect different student quality and/or different instructional attention to the creative products. Students' major field test scores suggest student quality differences. However, students' general education scores suggest instructional quality differences. Departments should determine whether declines in student ratings are evident in their data. If they are, faculty should examine the rubrics they have developed for creative products, their interaction patterns with students as students develop their products, and the quality of feedback they provide to students during that development. These data may reflect some institutional trauma as a number of

faculty were preparing to leave under duress and others were concerned about administrative changes affecting the academic program over which faculty had no control. This trauma could have directly or indirectly affected students.

Figure 6 shows student ratings separated by type of creative product. Last year, faculty interaction scores declined more steeply for performance than for paper products. This year, the decline was similar for performance and paper products.

Figure 6. Student ratings of creative product aspects by year of graduation

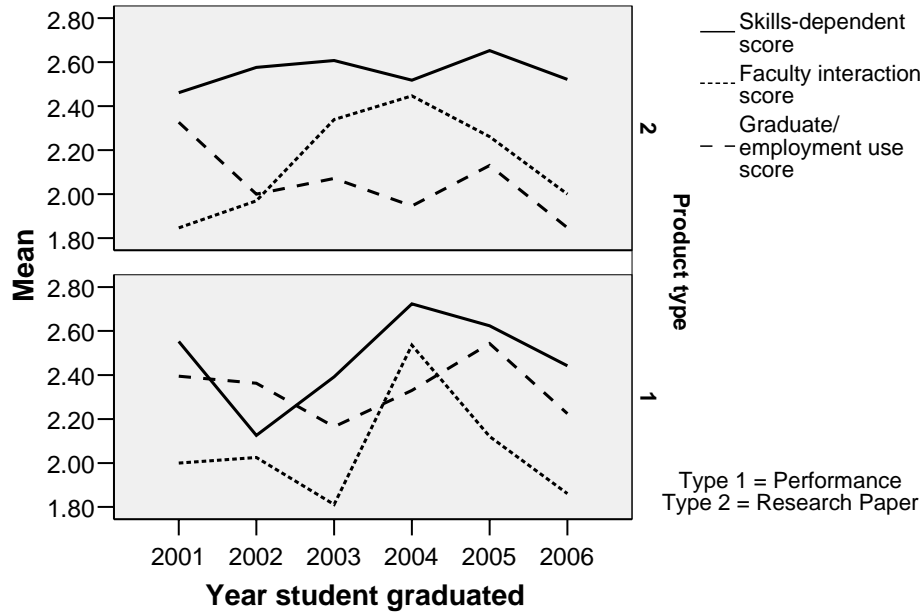


Figure 6 also shows an interesting difference between paper and performance products. Consider the difference between the lines representing student ratings of skill dependence and the graduate/employment use scores. The difference between these lines is greater for the paper than for the performance products. This suggests that students developing performance products have greater confidence in their skills than those developing papers. The large difference in lines for the paper products reflects a skills confidence gap in students. Departments requiring research papers could consider how their students might develop the confidence that the work they are doing represents professional skills.

Values

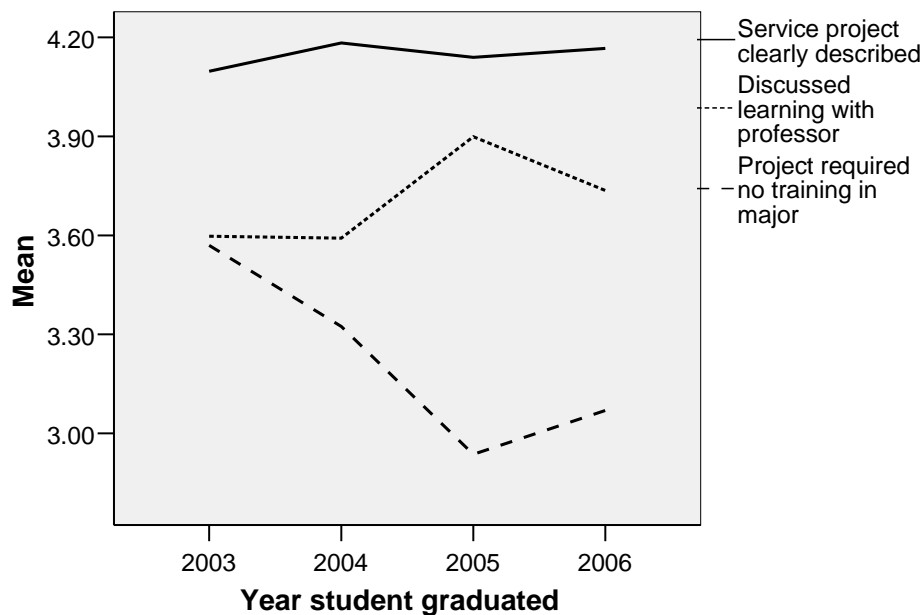
Service

The college catalog claims that every major requires its students to complete a service project. Last year we noted that the service project was not evident in catalog material for a number of departments and that some departments did not even have such projects for their students. While catalog references to a departmental service project are not required, every department should be able to document how its students complete that service project. That documentation is still not available for several departments. For other departments, the documentation on file does not match the service projects students

completed for 2006. In a few cases, departments have not ensured that all graduates completed their service projects. These are curriculum failures that should be corrected immediately. This year, students from all departments submitted service project evaluations. Thus, the integrity problems identified in last year's report have been addressed, at least in part.

Student reports of service projects confirm quality in some areas and suggest weaknesses in other areas. Students rated their service projects indicating their level of agreement on 10 questions. A "1" represents strong disagreement and a "5" indicates strong agreement. Three questions addressed instructional issues related to the service projects—whether the project was clearly described to students, whether students discussed the project with faculty, and whether the project required skills they had learned in their major. Figure 7 shows these instructional ratings by year since 2003.

Figure 7. Service project instructional scores by year student graduated

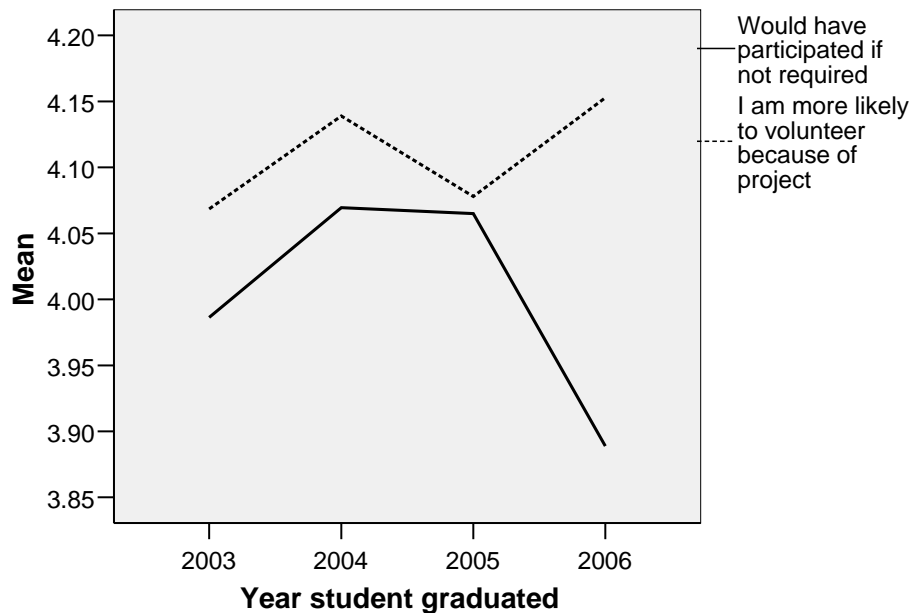


The figure reflects continuing strength in clarity of project description, strength in discussing projects with faculty, but a slight reversal of a trend in whether projects required skills students learned in their major. The bottom line in Figure 7 should be low if service projects requires students to use their learned skills. Service projects should demonstrate to students how their particular skills can be used to serve the community. Projects that simply provide community service that untrained people could perform may have community value, but they do not reflect the concern for integrating skills students are learning into a lifestyle of valuable service. The average rating just above 3 for 2006 indicates that students neither agree nor disagree that their projects require skills they learn in their major. Departments whose average scores are at or above 3 should examine their service projects to see if the project requires skills students have learned in the major.

The outcome of interest for service projects is whether students increase their commitment to service in the future. One estimate of this change in value is evident in

how students rate their willingness to volunteer for service projects. One evaluation statement addresses students' estimates of their likelihood of participating in the department's service project if it was not required. Another statement addresses the students' estimates of their willingness to participate in future service projects. Comparing these 2 lines provides some indication of change in value of service in the students due to the completed service project. Figure 8 shows these 2 lines by year of graduation.

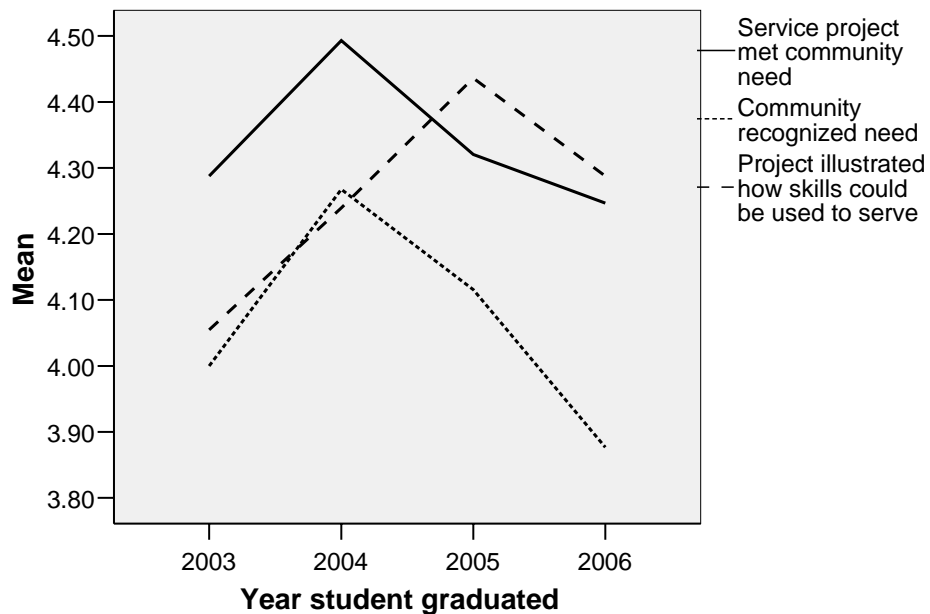
Figure 8. Service project values change by year student graduated



The data for 2006 are striking. We find the biggest difference yet between prior willingness to serve and future service likelihood. Furthermore, the students' agreement with likely service in the future was the highest score we have yet seen. These data may reflect improvements in service projects in some departments, and a developing familiarity with the service theme as more departments develop meaningful service projects.

In contrast to the value change ratings, students' ratings of the significance of their projects for meeting recognized community needs declined from last year. Figure 9 shows these ratings by year students graduated. Despite these declines, note that the agreement ratings are quite high.

Figure 9. Service project community needs items by year student graduated

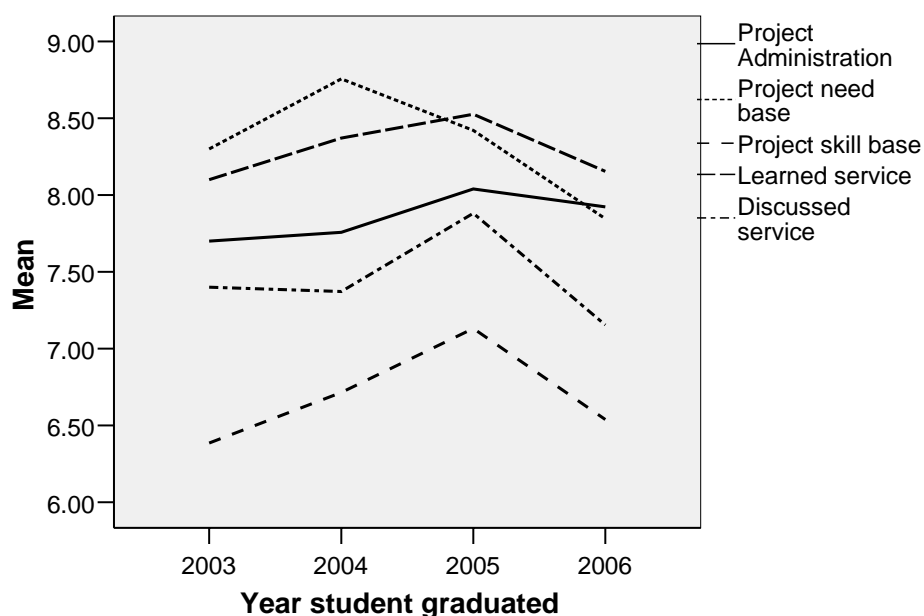


We summed scores for sets of questions to produce 5 scales, each with a maximum of 10 points. The scales are:

- ◆ Project administration—combines scores from questions about clarity of project description and faculty direction of the project.
- ◆ Need-based—combines scores from questions about whether the project was based on some community need and whether people in the community recognized this need.
- ◆ Required discipline skills—combines scores for questions about whether students used and/or learned skills related to their disciplines.
- ◆ Learned service value—combines scores for questions about learning the value of service and likelihood of volunteering in the future.
- ◆ Talked about service experience—combines scores for questions about talking with faculty and/or other students about their service experience.

Figure 10 shows student ratings by year for 4 years. All scales for 2006 declined from 2005. The declines were greatest for discussing service projects and whether projects were based on skills students learned in the major.

Figure 10. Service project scales by year student graduated



The accumulating data still suggest that Sterling College’s attention to service produces values changes in the desired direction for its students. The data also suggest that maintaining quality service programs requires ongoing attention to details of project design. Departments should attend particularly to the students’ developing professional skills when designing service projects. Faculty should also attend to how students reflect about their service projects. Some departments have exemplary programs requiring students to answer specific reflective questions. Such attention to instructional aspects of the service projects will enhance their quality.

Ethics instruction

Students completed ethics instruction evaluations for the third year. Two initial questions addressed the extent to which students received ethics instruction:

- ◆ Did you receive ethics instruction in your major?
- ◆ Are you aware of a professional ethics code in your discipline?

Students from all departments reported receiving ethics instruction. However, for four departments, not all of the students said they received ethics instruction. Furthermore, fewer than 75% of students reported awareness of an ethics code. Those unaware of a code included students from half (N = 7) of the academic departments. The 75% awareness of code rate has remained unchanged over the past 3 years. Some departments have made progress in connecting ethics instruction with professional ethics codes. Others do not use a professional code and others have used one, but have not consistently maintained their code-related instruction.

Some misunderstandings remain about why an ethics code is an important instructional component. Some departments assumed that the Assessment Committee was asking each department to find a professional code and adopt it for their students.

Departmental reports have noted that there are too many different codes for them to pick one, or that no codes match what they wish for their students. This is not the purpose of ethics code instruction. Students are preparing to enter a professional world. Our instruction should prepare students for professional responsibilities after graduation. Ethics codes exist for virtually every profession. Our students should be aware that such codes exist and that they will be held accountable to such codes. Faculty should not be satisfied with any scores indicating lack of awareness of ethics codes.

Corrective responses to the misunderstandings have occurred in several departments. Course assignments require students to find a professional code related to the profession of their interest within the discipline. Students then compare codes, looking for similarities and differences. One faculty member reported that students in his seminar found the assignment quite interesting.

Students also rated their level of agreement (*1 = strongly disagree; 5 = strongly agree*) with the following statements about ethics instruction in the department:

- ◆ The department conveyed to me the importance of ethical behavior for professionals.
- ◆ The department required me to demonstrate that I understood ethical issues.
- ◆ Because of my exposure to ethical issues in this department, I am aware of concerns I would not otherwise have known about.
- ◆ My future professional behavior will be different because I have learned ethical principles in my courses.
- ◆ How much do you agree that the activities the department chair identified represent the ethics instruction that you experienced? (Department chairs had provided a list of instructional experiences).

Table 2 shows average student responses to each question and the range of averages for all departments for 2006 graduates.

Table 2. Students' responses to questions about ethics instruction.

Question	College Mean	Range across Departments
The department conveyed . . .	4.38	3.0 – 5.0
The department required . . .	3.99	2.3 – 5.0
Because of my exposure . . .	3.79	1.7 – 5.0
My future professional . . .	3.79	2.7 – 5.0
How much do you agree . . .	4.24	2.0 – 5.0

While the means reflect general agreement with each statement, the range of scores extending to the “disagree” side of the rating scale suggests room for improvement in the nature and quality of ethics instruction. Department faculty should determine whether their students rated instruction on the low end of each mean. If so, faculty should modify how they incorporate ethics instruction into their curricular requirements.

General Education

Knowledge

Students complete the CBASE test during their senior year to provide information about their general education knowledge base. The CBASE test provides scores in 4 subject areas—English, Math, Science, and Social Sciences. We randomly assign students to one of these subject tests. All students also complete a writing sample. Each year we add scores from about 20 students to each of the 4 subject areas. For 2006, 72 students completed the CBASE test.

Because we randomly assign students to subject areas, for any year we have too few students for reliable conclusions. We aggregate data across years to evaluate general education outcomes. Table 3 shows the number of students for which we have CBASE records in each subject area. The number of scores in each subject now gives us some confidence in the conclusions we can draw.

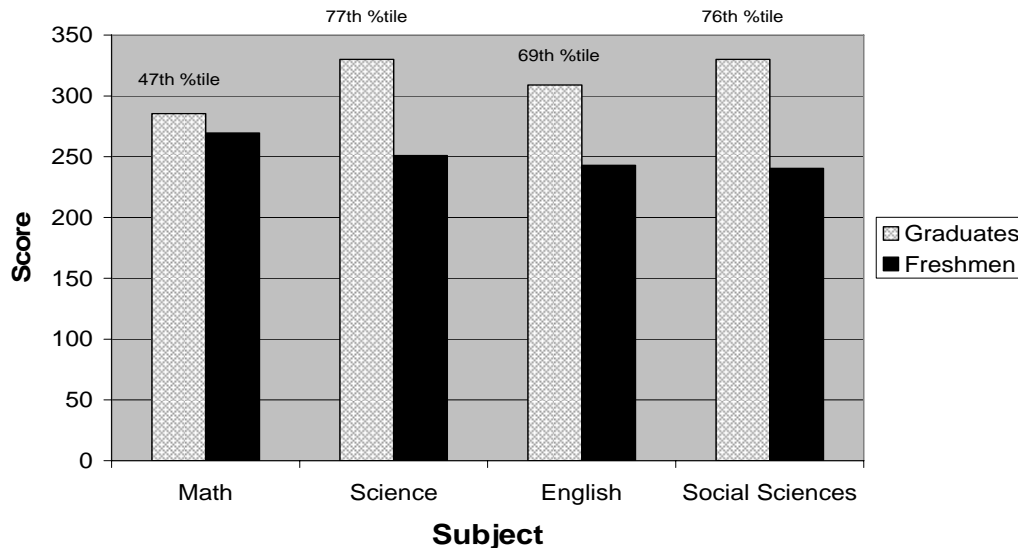
Table 3. Number of students in subject areas across all testing years.

Subject	Number
English	90
Math	94
Science	90
Social Studies	90
Total	364

A randomly selected sample of 40 freshmen also took the CBASE test, 10 in each subject area. We now have baseline data from 20 students in each subject area for comparing graduates with entering students.

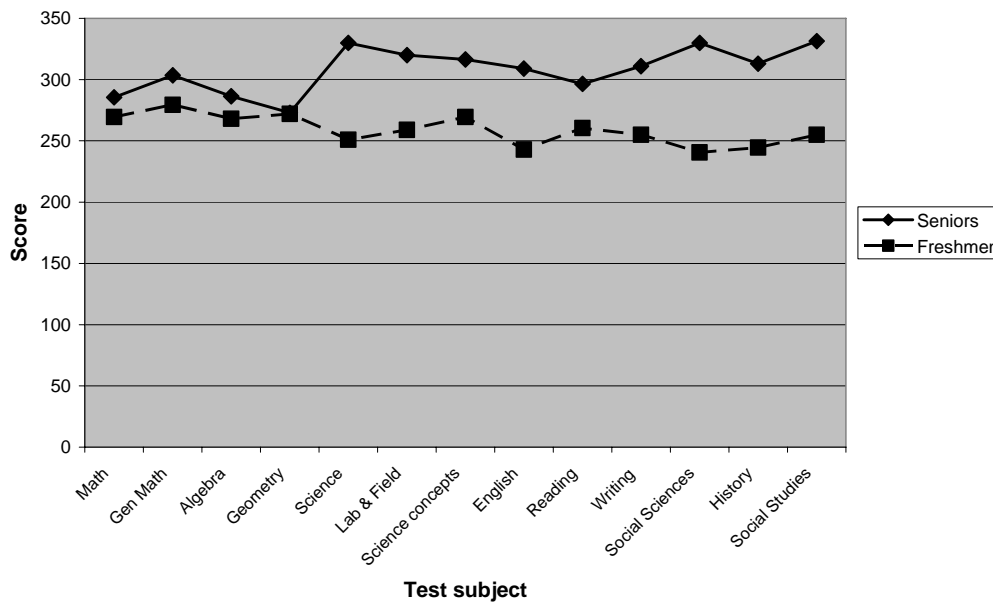
Figure 11 presents median CBASE scores and corresponding percentiles for graduates and freshmen. As in past years, these data continue to reflect general education

Figure 11. CBASE scores and percentiles



performance levels about which the college can be pleased. Figure 12 shows differences between graduates and freshmen for sub scores within the 4 main subject areas.

Figure 12. CBASE comparisons: Freshmen and seniors



CBASE interpreters suggest that any differences greater than 17 raw score points are “meaningful.” All scores for graduates, except for Math, are above the 50th percentile. The Math score is close to the 50th percentile. Furthermore, all scores, except for Math, are considerably higher for graduates than for freshmen. The English, Science, and Social Science scores around and above the 70th percentile are striking. These data still support the conclusions we first presented last year:

- ◆ Our general education program effectively presents a solid knowledge base to students in English, Sciences, and Social Sciences.
- ◆ Graduates score “meaningfully” higher in the above effective areas than do incoming freshmen.
- ◆ The difference between graduates and incoming freshmen in Math knowledge is not meaningful.
- ◆ Our graduates score above national samples on general education knowledge.
- ◆ Improving students’ knowledge of Math is the most effective way to improve general education program outcomes.

We have eagerly awaited math scores for the 2006 graduates. These scores represent students who graduated with a more rigorous general education mathematics requirement. We did not see the expected increase in math scores for these students. Math scores for 2006 graduates remained below the national median and did not differ from scores of our own prior graduates. We must recall that these scores came from a sample of our graduates, and the data are not reliable enough yet to draw any conclusions. We will continue to watch the math scores over the next several years to see if the 2006 data are repeated.

We again examined CBASE score differences between students who attended Sterling College for 4 years and those who transferred. Table 4 presents median scores

Table 4. CBASE raw scores for those attending 4 years or more or less than 4 years .

CBASE area	≥ 4 years	< 4 years	Raw score difference
English	311	288	23
Math	287	298	-11
Science	331	295	36
Social Studies	323	313	10
Writing sample	3.04	2.98	NA

updated for all students through 2006. The writing sample scores is on a different scale than the content area scores. The maximum writing sample score is 6. The sample of transfer students is still small, representing only about 20 students in each subject area. Therefore, we need to interpret these scores tentatively. Only the English and Science score differences are of a size that CBASE considers meaningful. Note that CBASE math scores are now lower for students who stayed at Sterling 4 years than for those who transferred. Do not over-interpret these sample data, but we should watch these scores carefully for several more years.

Some CBASE sub scores correspond directly with Sterling College general education objectives. Table 5 presents these sub scores and corresponding percentiles linked with general education objectives. These scores again indicate that our general education program covers knowledge components relevant to objectives.

Table 5. CBASE scores and percentiles related to GE knowledge objectives.

GE objective	CBASE sub score	Graduate median	Graduate %tile	Freshmen median
#4 Whole person development	English: Reading & Literature	305	64	253
#5 Natural world knowledge	Science: Concepts	310	69	277
#6 Social world knowledge	Social Sciences: Social Studies	322	78	264
#7 Historical patterns	Social Sciences: History	313	68	260

The freshmen sample is still small, but these data continue to indicate the largest score changes in the social science areas—the content areas in which the general education curriculum has common core requirements and an interdisciplinary course. These data strengthen the conclusion we have stated before about the possibilities of improving general education outcomes by developing more core and interdisciplinary courses.

Skills

CBASE provides indicators for the following additional general education objectives:

- ◆ #1 Information acquisition skills are included in scores related to reading, historical analysis, scientific data collection, and social science procedures.
- ◆ #2 Thinking skills are included in scores related to quantitative and critical thinking.
- ◆ #3 Communication skills are included in scores related to written communication and a writing sample.

CBASE presents these data in a different way than they present their subject scores and sub scores. For each skill we receive a statement of the percent of students who fall into each of 3 categories—high, medium, and low. A “high” score represents the top 16% of all scores. A “medium” score represents the middle 68% of scores. A “low” score represents the bottom 16% of scores. We can interpret these scores by examining whether our percentages in each category are similar to these standard percentages. Table 6 presents these percentages linked with general education skills objectives.

Table 6. CBASE scores related to GE skills objectives.

Objective #	Skill name	% low	% med	% high	Mean
1	Read accurately & critically	26	45	30	2.04
1	Read analytically	20	48	32	2.12
1	Understand writing process	19	56	24	2.05
1	Role of observing & experimenting in science	14	54	32	2.21
1	Science lab & field procedures	11	63	26	2.18
2	Use math to solve real problems	22	55	23	2.01
2	Use statistical reasoning	15	42	42	2.27
2	Evaluate algebraic expressions	29	53	18	1.89
2	Solve equations	19	69	12	1.93
2	Recognize 2-D & 3-D figures & properties	24	59	17	1.93
2	Perform calculations base on above	37	49	13	1.76
2	Interpret scientific results	3	67	29	2.26
2	Recognize social science procedures	4	51	45	2.41
3	Use standard written English	19	61	20	2.01
3	Use math notation to express ideas	14	72	14	2.00

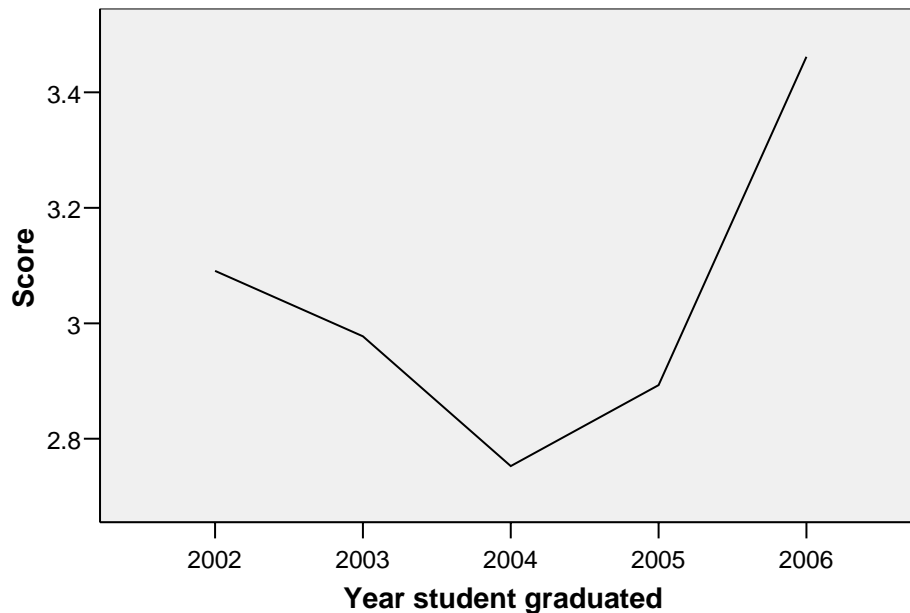
Lightly highlighted scores are those on which Sterling students performed better than the national sample. Darkly highlighted scores are those on which Sterling students performed worse than the national sample. As with content knowledge, students are strongest in social science and science skills.

On many skills, Sterling students are distributed differently than the national sample. If the first and third percentage columns are much higher than 16%, this indicates that Sterling has a larger cluster of students at either end of the skill distribution. Thus,

we have a relatively large group of students who score low on reading accurately and critically. Interestingly, we also have a large group of students who score high on this skill. The best scores are those where the “% low” is a low value and the “% high” is a high value, as for “Recognize social science procedures.”

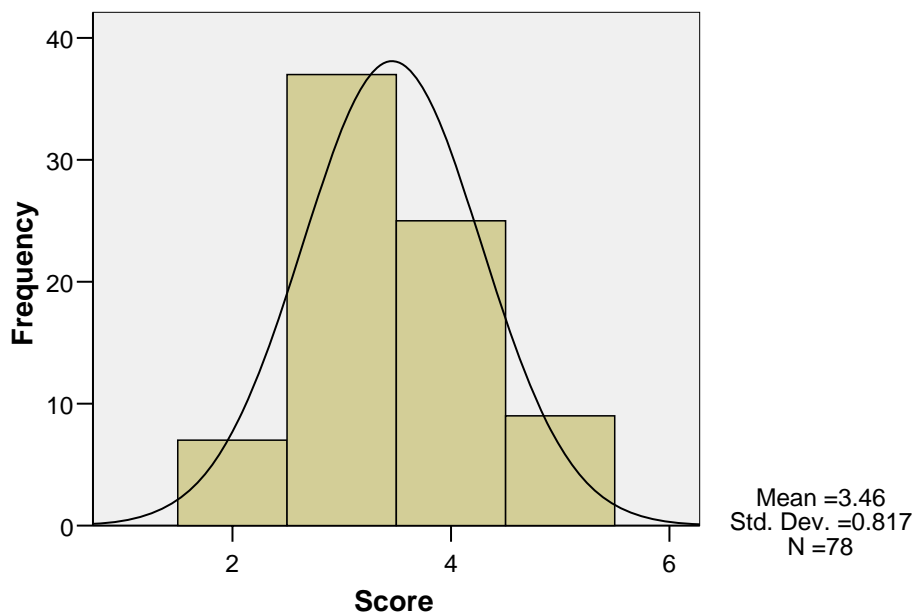
The final skills score that we report is for the writing sample. All students write an essay in response to a written prompt. That essay is scored by trained readers using clearly established criteria for good writing. Scores range from 0 (Scorer could not read the writing) to 6 (an exemplary essay in organization, paragraph structure, and support for assertions, that is generally free of grammatical and mechanical errors). Most of our students score “3”, and none have ever scored “6.” Figure 13 shows the average score for Sterling’s students by year of graduation.

Figure 13. CBASE writing sample by year student graduated



Graduates in 2006 scored better than any group of graduates since we began using the CBASE test. These data are reliable because all students completed these writing samples. Figure 14 shows the frequency distribution for Sterling students’ scores for 2006 graduates. This distribution shows fewer students scoring 2 and more scoring 4 and 5 than in previous years.

Figure 14. Writing sample distribution for 2006 graduates



While we may celebrate the increase, average scores in the 3 – 4 range still suggest that students have much room to improve their writing skills.

Success Stories

The academic community benefits by learning about the positive changes resulting from assessment in different programs. This year we again present several departmental “success stories.”

Departmental responses to assessment reports frequently include plans for program modifications. For various reasons, program changes take time to implement effectively. If the change involves curriculum revisions, the change must first be published in a new catalog and students under different catalog requirements gradually transition to the new requirement. If the change involves modifications in a class assignment, implementing that assignment effectively may take several semesters as faculty tinker with the nature of the implementation. The success stories presented below each illustrate the value of establishing a goal based on assessment data and pursuing that goal persistently over time.

New Creative Product in Education

The weak assessment cycle in our assessment program is the cycle that involves changing the method of assessing a particular outcome. We call this cycle weak because it alters methods of assessing and is not directed at changing learning structures. However, when the weak cycle works well, it has consequences for improving learning structures. Changes the Education Department made in its assessment strategies for 2006 illustrate how weak cycle changes can impact student learning—strong cycle outcomes.

For several years, the Department used actual classroom teaching activities to assess a teaching creative product. This choice for assessing student skills was very appropriate—what educators call “authentic assessment” because it assesses what the teacher really does. The Department documented the teaching products with a video of one lesson in the classroom and an evaluation rubric.

During the Department’s work to comply with state certification standards, faculty realized that they had a more comprehensive tool available for assessing student creative products. Kansas requires documentation of performance assessment in the classroom for new teachers and for continuing licensure. The Department recognized that the performance assessment standards and related rubrics were a more comprehensive measure of teacher planning and delivery than the classroom observation assessment tool they had used in past years. Therefore, the Department changed its assessment tool to match what it already had to do to document teacher performance outcomes.

Education students benefit from this assessment tool change in that they receive more specific and helpful feedback about their teaching skills. Instead of only examining performance in the classroom, the new rubric includes evaluation of the teaching context, the learning goals, the instructional design, the teacher’s integration skills, the classroom environment and dynamics, and assessments of student learning. The scoring criteria are explicit, allowing several people to evaluate the same teaching activity and determine reliability of the evaluation. The comprehensive nature of this new assessment tool makes it authentic in a much broader way than simply watching a person conduct a lesson in front of the class. The added advantage of the Department’s shift is that assessment for institutional outcome purposes matches what the Department has to produce for another function.

Ethics Instruction in Psychology

Psychology students have demonstrated compliance with professional ethical standards in research projects for years. Students first learn ethical principles of informed consent, full disclosure, freedom to discontinue participation, and confidentiality of records. Students then prepare a research proposal that addresses the ethical concerns. The Department assembles an Institutional Review Board (IRB) to evaluate the proposals students can conduct their research. The Department has documented the learning assignments and the IRB procedures from the first request for ethics assessment reports, and students have consistently reported receiving ethics instruction, being aware of ethics codes, and being required to demonstrate their understanding of ethics.

Reflecting on ethics instruction, Department faculty realized that it did not ask students to demonstrate their knowledge of professional ethics related to counseling—a major professional activity. Faculty identified courses that were required of all majors in which such content would be appropriate. The clinical courses, Theories of Personality and Abnormal Psychology fit the requirements. Department faculty then developed assignments that required students to respond to ethical issues after studying principles related to counselor/client relationships.

Early ethics instruction focused on ethical issues in research. The reflective work required in the assessment feedback cycles reminded department faculty that their ethics instruction was incomplete. New assignments are now on file to document ethics related to professional service activities.

Service project in Communication & Theatre Arts

College Catalog materials from the Communication & Theatre Arts Department reflect a continuing development of service project quality. In 2000, the catalog stated that at least one credit in a Departmental laboratory course had to include a service project. In 2002 the Department added specific service designations to laboratory experiences and began asking students to sign service project contracts. The Department worked for several years to implement its contract system for all graduates.

After the 2005 Annual Assessment Report, the Department faculty recognized some continuing weaknesses in its service project. Faculty developed a timeline for adding additional materials to the service project descriptions and creating new forms that described the project objectives and required student reflection about the value of the service experiences. They implemented these new instructional devices on or before Spring of 2006.

The new forms present a model of standard practice for good teaching. The first form, "Mission and Scope of Communication and Theatre Arts Service Projects," describes the purpose of the service project, identifies the departmental skills students could use in service, provides examples of appropriate service projects, and establishes expected activities as students complete the project. The second form, "Common Departmental Objectives for the Service Laboratory Course," was important because the department provides two partially separate learning tracks. The common objectives further clarify service project goals, and link the service project to careful planning and faculty review. A third "Response Form" asks those receiving service to respond to the quality of that service. A final form, "Student Self-Report: Questions and items for response," prompts students to reflect about their service, specifically regarding their interaction with faculty, their perception of community need for the service, the skills they practiced as they served, and their possible future participation in service. While the number of students completing the major this year was small, the responses to service project experiences by these students were outstanding.

Next Steps

Sterling College has experienced considerable faculty turnover since the current assessment plan was adopted. Furthermore, reading data does not automatically suggest what faculty might do with the data to improve student learning. After 6 years of program operation, some reflection about what faculty can and should do with the data we present seems appropriate. In this section, we present some general lessons about what the data mean and what faculty could consider as they reflect on their data to improve student learning outcomes.

Regarding compliance

Compliance problems are minimal. Students respond appropriately to timely requests to complete assessment data collection. The remaining problems involve individual departments which have not implemented effective procedures to collect the data for which they are responsible or general academic issues about double majors.

Departments are responsible for collecting and submitting the following data:

- ◆ Departmental reports

- Creative product evaluation form—each semester for all students completing their coursework that semester.
- Service project description—whenever the service project students complete is different from what the department previously described.
- Ethics instruction description form—whenever new components of ethics are included in the department’s instruction, or when faculty develop new ethics assignments. The chair should append all described assignments to the ethics instruction description form.
- ◆ Data chairs are to collect from students
 - Creative product self-assessment form
 - Service project self-assessment form

All forms described above are available at the College’s internal web site—
<http://www.sterling.edu/campus/acadaffr/facultyinfo/>

What should we do?

1. Check whether departmental reports are current. If not, revise the reports.
2. Check whether data for all graduates for 2006 were submitted. If not, determine the reasons for failed data collection and change procedures to prevent future failures.

Several general issues arise because some students earn double majors. Frequently these double majors are within a single department. Departments which offer more than one major often use a given course to meet requirements in both majors. The Assessment Committee assumes the following:

- ◆ Skills should be independently described for each major.
- ◆ Students should demonstrate the skills for every earned major.
- ◆ If students complete one creative product for two majors, that product should reflect all the skills in each major. In these cases, chairs should complete separate evaluations of the product—one for each major.
- ◆ Students could be required to complete separate products for each major. This will only be reasonable if departments clarify the nature of the distinct skills each major develops in its students.
- ◆ Changing assessment requirements (requiring separate products for each major) does not ensure that departments will address the underlying issue of clearly articulating required skills.

What should we do?

3. Through academic governance channels, faculty should determine guidelines for double majors. Guidelines should answer the following questions:
 - ◆ Can students earn two majors in one department?
 - ◆ If a department offers more than one major, should requisite knowledge and skills be separately described for each major?
 - ◆ How much overlap in knowledge and skills disqualifies offering a separate major?
 - ◆ Can a department use a single service project for students earning two majors from that department?

Clarifying these issues with sensitivity to the purposes of the creative product and service projects could strengthen the integrity of our academic program. In the process, we will resolve compliance problems related to double majors.

Regarding Knowledge

If faculty find consistent weaknesses in their students' knowledge base, they should consider potential correctives in three aspects of the instructional process—curricular design, pedagogical practices, and evaluation procedures. Curricular problems could include missing courses or missing content within existing courses. Pedagogical problems could include lack of clear objectives, poorly described assignments, or inadequate feedback practices that do not permit students to learn from their errors. Evaluation errors could include unclear evaluation criteria, or misapplication of stated rubric requirements. High grades could create a false sense of confidence about students' knowledge that lulls them into patterns of lowered effort.

What should we do?

4. If some knowledge indicator (e.g. sub score on a portion of the major field test) shows that students consistently score low in a content area the department judges as important, examine the structure and sequencing of courses that address the content. Consider changing sequencing and/or course requirements.
5. If general knowledge indicators (e.g. total score on major field test) are consistently low, evaluate pedagogical practices. Consider clarifying criteria, providing guided practice and specific feedback on student assignments.
6. If students in your department show unexpected deviations between grades and independent knowledge assessment (major field tests) consider revising your grading standards. Ensure that the evaluation instruments you use reflect the objectives and the experiences the students receive.

Regarding Skills

What should we do?

The issues faculty should consider regarding weaknesses in student demonstration of skills depend to some extent on the type of creative product students produce. For departments requiring research papers consider the following:

7. Do students have clear models of professional research papers? Students need professional models to know the goal the department seeks. Consider developing exposure over a student's years of study to increasingly complex journal articles. Many disciplines have journals that publish student papers. Also, even professional journals publish articles with varying degrees of complexity. Select articles for the general level of understanding of students in a particular class. Carefully sequenced courses in the curriculum could permit assignments of increasing complexity as the students move through the curriculum.
8. Do students have clear ideas about good writing? All faculty share responsibility for developing students' writing skills. Students should

demonstrate good knowledge of style in their creative products. An excellent writing resource for faculty and students is the small book, *Elements of Style* by Strunk and White. Our bookstore carries copies. Specific details for published articles are available through the professional journals. Teach your students to consult these, and hold them accountable to the professional standards.

9. Do your students understand the professional review process? Consider giving students feedback on paper drafts as an editor would respond to an article's author. Encourage students to send manuscripts for consideration to regional and/or student paper conferences.

Several recommendations apply to any type of creative product. These include:

10. Ensure that your expectations for the final product are clear and explicit. Many faculty develop rubrics that identify key elements and provide examples at different competence levels.
11. Ensure that students have the opportunity to receive feedback on products before final presentation.

Regarding Values

Service

Two problems plague existing service projects—some do not require students to use the skills they have learned in their discipline, and some do not require students to reflect about the value of service they can provide with their special skills.

What should we do?

12. Consider whether anyone could complete the service project you require of students without the skill training related to your discipline. If you answer positively, redesign a service project that better employs the skills of the discipline.
13. Consider what reflection about their service you require of students. If there is none, consider developing writing assignments and/or seminars that allow students to think about service in the context of their special skills and the value of service to public life and expressions of Christian identity.

Ethics

Departments vary considerably in their effective inclusion of ethics into instruction. General problems for some departments involve not understanding how professional ethics codes relate to their instruction and/or not finding effective ways for their students to demonstrate their understanding of ethical requirements.

What should we do?

14. Consider whether you can describe why it is important for students to be familiar with professional ethics codes in your discipline. If you can not, talk with your assessment liaison about the issues.
15. Consider whether you have found effective ways to include ethics instruction in your program? Identify the ethical issues your students

should know well. Ask what course content those issues most closely fit, and incorporate the ethics content into the subject matter.

16. Consider how students demonstrate their understanding of ethics. Do students complete assignments? Are those assignments related to the ethical decisions students will make as professionals in the future? If not, consider developing assignments related to ethical case studies.

Regarding General Education

Over the past 6 years faculty have not paid much attention to general education issues based on assessment data. We have changed the mathematics requirement, but faculty initiated that move independently of assessment data. Our inattention was, in part, by design. We agreed not to make changes for 3 years because of the sampled nature of the data. Those years have passed, and our inattention now might be related to structural issues. We have not had a general education department, and the General Education Committee has not been directly challenged to deal with assessment data. We have enough information now that some reasonable considerations are appropriate.

What should we do?

The group that inherits responsibility for general education should guide the faculty in considering several issues that arise from our data. These include the following:

17. Develop a systematic program for incorporating writing instruction into every major. This program might involve instructing the faculty and /or implementing specific standards for the common communication objective for every department.
18. Consider whether introductory courses for a major effectively address general education objectives? Should specific general education courses be designed around general education objectives? Common assumptions about how to construct general education may be wrong. Despite common practice, we should consider what is best for our desired student learning outcomes.

Conclusions

Sterling College faces some real threats to improving student learning processes. These threats arise both because of inertia and change. We identify four threats below in an attempt to retain a vital and effective assessment process which in turn should maintain a vital and effective instructional process.

The first threat is losing sight of our goal. We are an educational institution designed to provide effective learning processes. The threat arises from both inertia and change. From inertia, we become accustomed to repeated presentations of similar data. After the first presentation we may justify failure to respond to weaknesses on the basis of incomplete data. After the fifth presentation, we may not recognize that the amount of data now no longer justifies our non-attention. But the information is the same as last year, and last year we did not do much to change. So life goes on.

From change, we face new demands and new pressures that can easily distract us from attending to our job of improving student learning. We are changing our governance

structure. We are faced with student recruitment demands that require more of our time. We are not sure what our contracts mean. These changes easily take our attention away from understanding the purpose of assessment--to provide means for improving student learning. As faculty attention is distracted, student focus also wanes. Students may mirror the disruptions the faculty experience as they become directly aware of faculty situations, or as they simply respond to faculty who have been distracted from their basic goals. We also have many new faculty who may not understand assessment purposes or procedures. If we do not teach them in a timely way, fewer faculty each year will understand assessment purposes.

A second threat arises from the faculty turnover. Many faculty do not have a good understanding of our assessment program. Fewer than 50% of the current faculty were present when the program was developed. A full 30% are new in Fall 2006. To conduct an effective assessment program, faculty must know the program and relevant departmental history. The Assessment Committee can help with this problem. We retain all our records, and we should be aggressive in providing instruction. Faculty should also become aggressive in seeking information about past assessment reports for their department so knowledge of the program covers purpose, historic practice, and procedures.

A third threat to improving student learning outcomes through assessment is failure on the part of faculty to reflect on the data the Assessment Committee provides. Improving student learning is a data-dependent activity. Although we seldom directly measure desired outcomes, the indicators we measure give us some information about progress on achieving those outcomes. Over time, the data provide more reliable indications of student progress. While faculty in some departments examine the data carefully and make changes based on what the data reflect, faculty in other departments have made virtually no program modifications based on assessment data. Our feedback cycle is designed to stimulate reflection. However, if the data clearly indicate deficiencies, and we see no changes over the years in curriculum structure or pedagogical practices, faculty in those departments are completing assessment processes only in a perfunctory way. We have moved beyond responses to assessment that only nitpick the process. The college will experience systematic improvements in student learning outcomes as faculty take time to ask what the data mean, and what they can do to change what the data reveal are inadequate processes and outcomes.

A final threat to the effectiveness of the assessment program arises from uncertainties about faculty governance processes. We have a new Academic Dean who understands both faculty governance issues and assessment practices. The Assessment Committee urges the administration to respond positively to effective governance structures and procedures that emerge during the year. These are essential for appropriate consideration of changes assessment data may generate from departmental faculty.

Departmental data

The next pages include specific reports for each department. The Assessment Committee's file copy contains all department reports, as do the copies given to the President and the Academic Dean. Each department chair will have the departmental report for his/her department. Generally distributed copies of this annual assessment report do not contain departmental reports.