RICHARD STOCKTON COLLEGE OF NEW JERSEY

EVIDENCE

PROGRAM ASSESSMENT FOR CONTINUOUS IMPROVEMENT

Fall 2010

VSA 101

Amy Hadley

Assessment of Peer Tutoring

Jill Gerhardt and Michael Olan

CLA 2009-2010: What Can We Learn?

Heather McGovern

Giving Our Transfer Students Due Credit:: What the CLA Tells Us About Transfer Students' Critical Thinking and Writing Skills

Heather McGovern

VSA 101

I had an opportunity this summer to attend the *2010 VSA Learning Outcomes Workshop* in Baltimore, MD. I've been involved with assessment at Stockton since attending the Assessment Institute sponsored by the Institute for Faculty Development in the summer of 2007. Attending the Outcomes Workshop this summer clarified for me the background on Stockton's continuing efforts toward gathering assessment data as well as Stockton's involvement in the Voluntary System of Accountability. In this article, I hope to clarify some of the questions that other faculty members may have about assessment initiatives at Stockton.

What's the VSA?

VSA stands for Voluntary System of Accountability. According to the VSA website, (http://www.voluntarysystem.org)the VSA is a voluntary initiative developed by the higher education community to meet the following objectives:

- •Provide a useful tool for students during the college search process

Assessment of Peer Tutoring

The enrollment and participation of female and minority students in computer fields is of great concern, and creating a female- and minority-friendly class environment is desirable. In 1985 women earned 37% of the computing degrees (NSF 2007). From that high point the trend has been downward. In 1993/1994, women earned 18% of the computing degrees in the United States and in 2006-2007 they earned 12% (Vegso, 2008).

As the number of female computing majors is going down, the need for computing professionals is going up. The U.S. Bureau of Labor Statistics (2008) released its latest projections of the fastest growing jobs. It predicts the following growth by 2016:

Database Administrators	29%
Computer Systems Software Engineers	28%
Network Administrators	27%
Computer Applications Engineers	45%
Network Systems Analysts	53%

Assessment of Peer Tutoring, continued from page 1

Computer programming and problem solving is often considered the "heart and soul of computing" (Denning and McGettrick, pg.16, 2005). It is an integral course for the computer major, and one that has been historically difficult for students, so it is especially important to provide a female and minority friendly environment. One way to facilitate this is by providing free tutoring.

> *Initially, respondents cited not knowing that a tutoring service was available as the second most common reason for not using the tutoring service. As information about the tutoring service became more widespread, students' reasons for not using it have shifted to time scheduling conflicts.*

Research demonstrates that student-student interaction contributes to students remaining in the major (Barker, 2009). Peer tutoring offers such student-student interaction within the major. This can help to develop an academic community which will support the student throughout her college career.

Peer Tutoring in Programming

Michael Olan recruited and trained the tutors, students who had successfully completed the introductory programming and problem solving courses (CSIS 2101/2102). They first tutored for these courses in Spring 2004.

Methodology

At the end of each semester, a survey was administered in the CSIS 2101/2102 courses. This study reflects data that was collected over four semesters beginning Spring 2004 and ending Fall 2005. A total of 182 questionnaires (See Appendix A) were completed, 113 from CSIS 2101 students and 69 from CSIS 2102 students.

	CSIS	MATH	BSNS	Other
CSIS 2101	47	60	0	9
CSIS 2102	66	2	1	0
Total	113	62	1	9
	63%	34%	0.5%	5%

Make-up of the Students

The distribution of responses by major was as in the first table on this page; there were several double majors. "Other " included Physics, Visual Arts, Political Science, and Hospitality. In CSIS 2101, 41.0% of the respondents were CSIS majors, and 96.0% of the CSIS 2102 respondents were CSIS majors. MATH and other majors require or encourage students to take CSIS 2101, but no other majors require CSIS 2102. Seven percent of students in CSIS 2101 and 4.3% of students in CSIS 2102 were repeating the course.

	Did not need help	Did not know about service	Time conflicts	Other
CSIS 2101	69	13	6	4
CSIS 2102	30	15	17	1

Frequency of Tutoring

Twenty-one percent of students in CSIS 2101 and 26% in CSIS 2102 used the tutoring. About half of the CSIS 2101 students who used the tutoring were CSIS majors. The average number of times that those who used the tutoring did so was 2.3 for CSIS 2101 and 2.6 for CSIS 2102. Respondents cited "did not need help" as the number one reason for not using the tutoring. Initially, respondents cited not knowing that tutoring was available as the second most common reason for not using the tutoring, but as information about the tutoring became more widespread, scheduling conflicts became the second next most common reason.

VSA 101, continued from page 1

- Demonstrate accountability and stewardship to public
- Support institutions in the measurement of educational outcomes and facilitate the identification and implementation of effective practices as part of institutional improvement efforts.

The VSA was developed in 2007 and is sponsored by the Association of Public and Land-grant Universities (APLU) and the Association of State Colleges and Universities (AASCU). Four-year public institutions, such as the Richard Stockton College of New Jersey, participate in the VSA. Other New Jersey institutions participating in VSA include TCNJ, Montclair, NJIT, Rutgers, and William Paterson. The common tool used to report data by VSA institutions is called the *College Portrait*.

After hearing the presentations as well as discussions on reliability and validity of the three instruments, I was confident that of the three instruments, the CLA was the best match for Stockton's assessment needs.

Information summarized in the *College Portrait* includes "student and campus characteristics, cost of attendance, success and progress rates, campus safety, class size, student experiences on campus, and student learning outcomes" (http://www.voluntarysystem.org). Stockton's *College Portrait* information may be viewed at: http://www.collegeportraits.org/NJ/RSCNJ.

What Kind of Data Are Reported?

If you haven't yet clicked on the link for Stockton's *College Portrait*, I suggest that you visit the site. Some information is descriptive or demographic. Other information is reported from the results of assessments and surveys. For example, information is summarized about Student Experiences at Stockton. The information includes selected results from the *National Survey of Student Engagement (NSSE)*. (You may have heard this referred to as the "Nessy.") Responses to the NSSE are summarized according to Benchmarks of Effective Educational Practice: Level of Academic Challenge, Active and Collaborative Learning, Student-Faculty Interaction, Enriching Educational Experiences, and Supportive Campus Environment.

If you've noticed that *engagement* has become a buzzword around campus, this is one of the reasons. Some of our results on the NSSE are great! For example: 88% of seniors worked with classmates on assignments outside of class. NSSE data also tells us areas in which we need to expand engagement. For example, freshmen and seniors at Stockton report that they prepare for class fewer hours on average than at other schools while our seniors reported spending more time on average than other schools' students relaxing and socializing (defined as "watching TV, partying, etc.").

The other assessment data reported on *College Portrait* involves Learning Assessment.

To participate in the VSA, an institution has to utilize one of the assessment tools approved by the VSA. There are three approved instruments: the *ETS Proficiency Profile, the Measure of Academic Proficiency and Progress*, and the *Collegiate Level Assessment*.

The VSA Instruments

The *ETS Proficiency Profile* assesses four core skill areas — critical thinking, reading, writing and mathematics (http://www.ets.org/proficiencyprofile/about). During the 2009-2010 academic year, the Assessment Committee discussed the pros and

cons of using the *ETS Proficiency Profile* at Stockton and concluded that it was probably not the best match for our needs. The *MAPP*: *Measure of Academic Proficiency and Progress* is another instrument used by some VSA institutions. According to ETS, "MAPP is an integrated test of four academic skills (critical thinking, reading, writing and mathematics), and measures these skills in three contexts: humanities, social sciences and natural sciences" (http://www.ets.org/s/mapp/pdf/5018.pdf). This instrument has not been adopted for use by Stockton and is probably inappropriate given the interdisciplinary nature of our General Studies curriculum.

The instrument that has been used for institutional assessment and VSA reporting at Stockton is the *CLA:* the *Collegiate Level Assessment*. According to CLA, the instrument "presents realistic problems that require students to analyze complex materials and determine the relevance to the task and credibility. Students' written responses to the tasks are evaluated to assess their abilities to think critically, reason analytically, solve problems and communicate clearly and cogently" (http://www.collegiatelearningassessment.org). Stockton now has several years' results indicating performance of Stockton freshmen and seniors on this instrument. Refer to previous issues of *Evidence* for discussions on past test results and the efforts of Stockton faculty who have worked on *CLA in the Classroom* projects which can be integrated into the curriculum. Other articles in this issue explore CLA results for 2009-2010 and compare transfer and native CLA results.

Learning Outcomes

Selected institutions presented research on learning outcomes at their schools as measured by one of the approved VSA instruments. After hearing the presentations and discussions on reliability and validity of the approved instruments, I was confident that of them, the *CLA* was the best match for Stockton.

One of the presentations of note was *The CLA in the Classroom at Fayetteville State University: Using Assessment to Improve Student Learning* by Brooks, Valenti, and Young. Fayetteville used *CLA in the Classroom* teaching and assessment methodologies across freshmen seminars. The researchers concluded that use of *CLA in the Classroom* supported student learning and equipped students with essential skills, influenced faculty development while gaining strong faculty support, and helped to shape the institution's academic culture. Reports on the materials and outcomes developed by Fayetteville faculty are located in their repository at: http://digitalcommons.uncfsu.edu (use the search term "CLA").

The CLA is an authentic measure of assessment, thus using CLA methodologies in the classroom is not "teaching to the test". The outcomes targeted in CLA include critical thinking, analytic reasoning, problem solving and written communication. As we develop curriculum and identify learning outcomes in both General Studies and Program courses, we need to investigate how these higher order skills are being addressed. Skills such as critical thinking, analytic reasoning, problem solving and written communication cannot be compartmentalized to a single course if authentic learning is taking place. Rather, we need to integrate these skills across the curricula and develop means of effectively measuring student learning outcomes.

Amy Hadley, *Associate Professor of SPAD*

> *Skills such as critical thinking, analytic reasoning, problem solving and written communication cannot be compartmentalized to a single course if authentic learning is taking place. Rather, we need to integrate these skills across the curricula and develop means of effectively measuring student learning outcomes.*

Assessment Peer Tutoring, continued from page 2

When ranking their reasons for using the tutoring service, CSIS 2101 and CSIS 2102 students ranked "Needed help with homework / programming projects" as their highest priority. In CSIS 2101, this was followed by "Needed help preparing for exams" and finally "Needed help in understanding course concepts". In CSIS 2102 the second and third reasons were reversed with understanding course concepts ranked over exams.

	Help with concepts	Help on homework / programming projects	Exam prep
CSIS 2101	1.9	1.1	1.7
CSIS 2102	1.25	1.4	2.9

Assessment of Tutors and Service

On average, respondents who used the tutoring service found it very useful and the tutors knowledgeable.

Useful	Not at all	Some what	Very	Ex-tremely
CSIS 2101	0	7	9	13
CSIS 2102	0	7	12	4

	Not at all	Some what	Very	Ex-tremely
CSIS 2101	0	5	9	15
CSIS 2102	0	5	13	5

They tended to use the tutoring service as individuals, and overall they were very satisfied.

	One-on-one	Group	Both
CSIS 2101	17	7	5
CSIS 2102	14	7	6

	Not at all	Some what	Very	Ex-tremely
CSIS 2101	0	5	12	12
CSIS 2102	1	7	11	4

Almost all respondents (49 of 52) who used the tutoring service said that it helped in their understanding of course material, and a significant majority (40 of 50) thought it improved their grade. Based on student responses, the tutoring service in programming was a success.

References

Barker, L., McDowell, C. and Kalahar, K. (2009) "Exploring Factors that Influence Computer Science Introductory Course Students to Persist in the Major" *SIGCSE Bulletin,* v.41, n.2, pp.282-286.

Bureau of Labor Statistics, *U.S. Departments of Labor*. (2008). www.bls.gov

Denning, P.J., and McGettrick, A. (2005) "Recentering computer science," *Communications of the ACM* (48:11) pp 15-19.

(NCWIT) National Center for Women & Information Technology (2007) "A Report on the Status of Women in Information Technology." www.ncwit.org/pdf/2007

(NSF) National Science Foundation Division of Science Resources Statistics (2007) "Women, Minorities, and Persons with Disabilities in Science and Engineering" (NSF 07-315), Arlington, VA, Feb. www.nsf.gov/statistics/wmpd/.

TED: The Editor's Desk (2009) "Women and Men in Management, professional, and related occupations, 2008"www.bls.gov/opub/ted/2009/ted_20090807.htm/

Vegso, J."CRA Taulbee Trends: Female Students & Faculty," June 2008; www.cra.org/info/taulbee/women.html

The article "Assessment of Peer Tutoring" is adapted from Gerhardt and Olan, "Peer Tutoring in Programming: Lessons Learned," *Proceedings of the 26th Annual Information Systems Education Conference,* v 26, Washington, DC, November 2009.

Appendix A: Tutoring Survey Questionnaire

MAJOR: CSIS MATH   BUSINESS   OTHER

Which course are you taking?   CSIS 2101   CSIS 2102

How many times have you taken this course before, and either dropped it or received a D, F, W or I?
    None   1   2   3 or more

If you did ***not*** use the CSIS tutoring service this semester, was it because

        did not need any outside help in this course
        did not know that there was a tutoring service
        available tutoring times did not fit my schedule
        other

(Only answer the remaining questions if you *did* use the CSIS tutoring service this semester)

How many times did you use the CSIS tutoring service this semester?

Why did you use the tutoring service (write 1 next to the most important reason, 2 next to the second most important reason, etc. --
    write 0 next to any that do not apply)

    _____ Needed help in understanding the course concepts.
    _____ Needed help completing homework assignments or programming projects.
    _____ Needed help preparing for exams.
    _____ Other (please specify)

How useful was the tutoring service.

    Not at all   Some what   Very   Extremely

Was the tutor knowledgeable about the subject?
    Not at all   Some what   Very   Extremely

Did the tutor address your problem to your satisfaction?

    Not at all   Some what   Very   Extremely

Was the tutoring you received usually: One-on-one   With a group   Both
Do you think the tutoring service helped you better understand the course material?   Yes   No
Do you think the tutoring service improved your grade in this course?   Yes   No

Please give any comments about how to improve the CSIS tutoring service.

Jill Gerhardt, Associate Professor of CSIS
Michael Olan, Professor of CSIS

CLA 2009-2010—What Can We Learn?

Overall, Stockton CLA results (see Amy Hadley's article in this issue to learn more about the CLA and why we use it at Stockton!) for 2009-2010 tell two stories. One is neutral. In a cross-sectional analysis, comparing first year students and senior students for the same year (and considering how we'd expect them to perform considering their SAT scores) , **our seniors scored "near" expected in all tasks**. In other words, their scores were as one would expect them to be given a measure of their SAT scores and time spent at a participating four year college in the U.S**. The other story is less happy—the highest percentile rank for our students is 40 (on the performance task) and the percentile rankings for the writing tasks range from 16-22.**

First, let's explore the good news. Those who can remember CLA reports over the last several years know that sometimes our students have scored much less well than we'd expect as seniors. Once, our CLA results appeared to indicate actual value added to our seniors' critical thinking abilities from their experience at Stockton. Although that was the year in which the administration distributed printed results at the Fall Faculty Conference, that year's "good" results can be almost entirely explained by first year students scoring much lower than we would have expected based on their SAT profile rather than a change in our senior scores. Overall, having our students score as we would expect is fine, but it is not call for celebration.

Indeed, deeper delving into our results is troubling. Our student scores in the performance task category, in which students write answers to case study like questions that ask them to read, synthesize, analyze, and evaluate many sources, are fairly strong. Performance task is the only area in which we are solidly in the "near" category given our confidence interval. In the other areas, we were lucky to see ourselves in "near" rather than below that, given our confidence intervals. Even in our strongest category, performance task, our percentile rank is 40; in the other categories it is below 25 . We have room for improvement.

Table one: CLA Value-Added Results for 2009-2010 (from the CLA report)

	Performance Level	Value-Added Score	Value-Added Percentile Rank	Confidence Interval Lower Bound	Confidence Interval Upper Bound
Total CLA Score	Near	-0.56	25	-1.23	0.11
Performance Task	Near	-0.13	40	-0.9	0.64
Analytic Writing Task	Near	-0.85	16	-1.57	-0.13
Make-an-Argument	Near	-0.69	21	-1.46	0.08
Critique-an-Argument	Near	-0.78	22	-1.56	0

People may desire to dismiss our CLA data because of concerns about sampling, student motivation, or test validity. Our sampling of first-year students has remained stable (we recruit whole classes of first-year seminar students from first-year seminars, selecting a distribution across GAH, GSS, GNM, and GEN courses so as to try to avoid a sample with students with a particular disciplinary bias). We've tried many different techniques for sampling seniors, including an attempt to sample native seniors randomly, but the most effective method locally has proved to be recruiting whole classes with a high native senior enrollment, again striving for classes across the disciplines to avoid disciplinary bias. Last year's spring senior sample includes seniors from senior seminars in Literature, History, and Mathematics as well as students from Auditing and other upper-level business and health science courses. Despite changes in our senior samples, our senior scores have been about as or more stable (in terms of expected score based on SAT profile) than our first-year scores.

CLA 2009-2010, continued from page 7

Table 2. Student scores on CLA, 2006-1010. Scores that indicate students did better than we'd expect in green and scores that indicate students did less well than we'd expect in red.

	First year mean SAT or EAA	Senior mean EAA (SAT or substitute measure of entering academic ability)	First year expected CLA	First year actual CLA	Senior expected CLA	Senior actual CLA
2006-07	1066	1135	1080	1131	1216	1114
2007-08	999	1092	1029	1008	1179	1104
2008-09	1066	1053	1099	1064	1183	1195
2009-10	1115	1003	1120	1124	1161	1148

Table 3. Value-added summary for seniors

	Total CLA score	Performance task	Make an argument	Critique an argument
2006-07	Well below	At	Insufficient data	Insufficient data
2007-08	below	below	below	below
2008-09	above	Well above	At	above
2009-10	Near	Near	Near	near

Student motivation, of course, is likely a factor, but it does not fully explain our results. Speak with colleagues from other schools about the CLA, and you'll be able to attest to the fact that senior motivation to take the test is challenging across the nation, not just at Stockton. If our seniors are less motivated than those elsewhere (possible), that is worrisome. To further investigate, this year we're gathering some basic information about student motivation. For last year, however, we can use time spent on the test as one measure of motivation. Mean and median time spent by first year and senior Stockton students on the CLA in 2009-2010:

Overall: 37 minutes. Range: two minutes-87 minutes.
First-year students: Range 2-78 minutes., mean 36 minutes, median 35 minutes.
Seniors: Range 3-87 minutes, mean 39 minutes, median 38 minutes.

Eleven students spent less than 10 minutes; all 11 performed well below expected . Of the 12 students who spent 10-15 minutes, most performed near expected, two below , one well below, and one above expected. Of the 21 students who spent at least an hour , 11 scored well above expected, four above expected, one near expected, one below expected, and two well below expected. Correlation analysis using a Pearson coefficient shows a correlation of .50 for time spent and performance task score and .58 for time spent and analytic writing score. Additional analysis done by IR, comparing time spent to grades in W courses, indicates that time spent is the stronger predictor of CLA writing score. We wouldn't expect a perfect correlation between a course grade/test score, because course grades measure different things (e.g., student behavior), because students might get much more support and have more time for papers in a class, and because the tasks and student attitudes towards them differ. Student motivation does explain some poor student performance. It does not explain all good or bad scores, nor does it explain why our students would be less motivated than others.

A particularly puzzling characteristic of student performance on the CLA at Stockton **is the high variability among students, much higher than that nationally.** Yes, we are a state college, but so are most of the other participants. I have no explanation to offer for this, particularly because (as you can see in the other article in this *Evidence* on transfer student performance) our high percentage of transfer students does not appear to help explain it. Nor, given the fact that student scores are "corrected" for SAT score, does another ready explanation come to me—but I am open to hypotheses!

As a member of the writing program, I find it particularly distressing and puzzling that students perform poorly in writing. Other data we're gathering about faculty pedagogy in W's or students perceptions of faculty pedagogy seem to indicate that in most areas Stockton is doing as well or better than others (one notable exception is that our students write fewer longer papers than students nationwide). Nonetheless, student performance on many direct measures of writing skill indicate repeatedly that we need to do more to develop our students' writing skills.

What to look for next year? Next year CLA will provide us with more score breakdown. Specifically, they'll let us know how our students score in critical thinking vs. grammar and mechanics on the writing tasks. This may help us know what action to take.

What action should we take? **We should all do more to help our students develop the skills tested by the CLA**, not because we want our student's CLA scores to go up (although we do) but **because we want our students, indeed, all college graduates, to have strong critical thinking and communication skills, including the ability to synthesize, analyze, and evaluate multiple sources and to communicate effectively in writing**. Note that Amy Hadley makes this same argument in the conclusion to her piece in this newsletter—she came to this conclusion independently (and I didn't bribe her!) Even if our CLA scores are invalid or if the CLA doesn't measure these skills well, we do no harm to our students if we ask them to practice these skills more frequently and provide more feedback to them to help them improve their scores.

We also can all play a role in encouraging our senior students to do their best on the test—telling them that it is important in Stockton's self-reflection and recruiting efforts. Statements on syllabi or in class for senior-heavy classes will take just a few seconds but can help reinforce that this measure is important to the community.

 If you are a skeptic of the format of the test, I challenge you to take it yourself or go to the CLA website and read sample questions. Or talk to Amy Hadley or others who've learned more about the test through Assessment Institutes. Or, even better, give a CLA-like performance task to your own students and see how they do. Many colleagues across campus have been doing so, and sometimes they find their students perform well. Other times, they are depressed at how poorly students do, even when students work hard on the task during class time for a graded assignment and have previously completed assignments that should have helped prepare them for the task. Whether students perform well or badly, faculty and students alike have more information about which skills need more development and where student skills are strong. **We will not harm students if we proceed as if the test results are valid and strive to improve student skills. We will harm students if we treat the results as invalid when they are not.**

Heather McGovern, Associate Professor of Writing and Director of the Institute for Faculty Development

CLA and Transfer students: Spring 2009 and Spring 2010

Bottom line? Based on CLA results, we have no reason to think that the critical thinking/writing skills or our native students is better or worse than that of our transfer students. This is good news and bad news: we aren't disadvantaging our transfer students. On the other hand, we aren't adding significantly more value to our native students.

Most of the time, we try not to sample transfer students in our CLA samples. However, we sampled some in Spring 2009 (when trying to complete our sample group after our experiment in random sampling failed) and intentionally sampled some in Spring 2010 so that we could use the CLA data to further investigate internal populations of interest to us (Table 1).
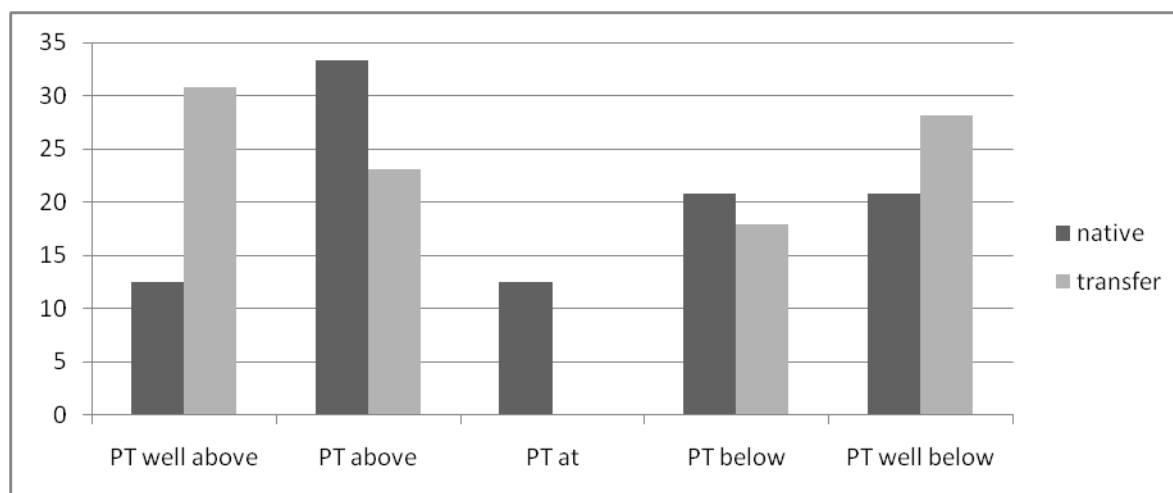
Table 1. Mean CLA raw scores by transfer and native seniors 2009-2010.

	Mean, natives	Mean, transfers	Median, natives	Median, transfers	Standard deviation, natives	Standard deviation, transfers
Performance task	**1183.532**	1110.621	**1189**	1116.5	**193.8744**	172.3003
Analytic writing	1185.167	1181.82	1183	1170	150.0887	154.3553
Make an argument	1166.245	1177.344	1151	**1171**	171.9575	171.6173
Critique an argument	**1201.521**	1185.77	**1203.5**	1159	168.7501	**197.8359**

This data includes 95 native and 128 transfer seniors.

Note that the variance in all samples is very high, and often similar across groups, except in the case of critique an argument where our transfer students showed more variance. There are no statistically significant differences between mean students scores for transfer and native students.

Another way to see the data is to look, visually, at the variation in student performance. The following chart shows comparisons of percentages of transfer and native test takers who scored "well above," "above," "at," "below," or "well below" expected when compared to how one would expect them to score relative to their Entering Academic Ability scores are based on the SAT for native students and the SLE for transfer students.

Heather McGovern, Associate Professor of Writing and Director of the Institute for Faculty Development