What did we learn from the CLA?

Neal Christopherson, Office of Institutional Research
August, 2007

Whitman College administered the Collegiate Learning Assessment (CLA) on a one-time trial basis during the 2006-07 academic year. The results provide an interesting comparison with other schools, a unique angle not provided by the survey data we usually use for peer comparison. This report will first examine the CLA in the context of current assessment debates in higher education, provide an overview of the CLA itself, and then present the results for Whitman College. The use of our CLA results will be discussed at the end of this report.

The CLA in the Context of Current Higher Education Debates

In recent years, parents and lawmakers have expressed a growing concern about what they perceive to be a hollowing of higher education, a belief that undergraduate education is not as good as it could, or should be. Whether these concerns are warranted is open to debate, and is not of concern for this report. Rather, it is important to acknowledge the existence of such concerns, as they are creating ripples in the landscape of higher education.

According to a report issued by Secretary of Education Margaret Spellings, the way to improve higher education is through greater transparency, and through a focus on "value-added" assessment. By focusing on "transparency," Spellings has indicated it is no longer adequate to merely use assessment of individual students to improve learning on a single campus, but rather there is now an interest in data that can be used to make comparisons of student learning across campuses. Employers, taxpayers, and, (as a result?) legislators, want to know how institutions compare in the amount of learning (value-added) they foster among their students. Spellings argues that the results of student learning assessments, including value-added measurements that indicate how students' skills have improved over time, should be made available to prospective students and employers, and reported in the aggregate publicly. This focus on value-added assessment encourages an emphasis on measures that can be compared across institutions.

Fueling this fire is research from Pascarella and Terranzini's landmark study *How College Affects Students*. They argue that students "grow" in college no matter what school they attend, and where they go doesn't make much difference. Students improve about the same amount in any college setting, and any end-point differences we observe across schools are simply due to starting-point differences. On page 592 of *How College Affects Students* they write:

> "[T]here is little consistent evidence to indicate that college selectivity, prestige, or educational resources have any important net impact on students in such areas as learning, cognitive and intellectual development, other psychosocial changes, the development of principled moral reasoning, or shifts in other attitudes and values. Nearly all the variance in learning and cognitive outcomes is attributable to individual aptitude differences among students attending different colleges. Only a small and perhaps trivial part is uniquely due to the quality of the college attended."

In other words, no matter where students go, they all learn about the same amount.

The Collegiate Learning Assessment fits nicely into this picture. The CLA addresses this new value-added focus, and is seen by some as a magic bullet that will appease the accreditors, government, employers, parents, and taxpayers without compromising institutional integrity. The CLA is led by Richard Hersh, former president of Hobart and William Smith College and Trinity College (CT). He started the CLA because he wanted to be able to tell people (primarily parents) why small private colleges are worth the extra money. Some reasons might seem obvious: smaller classes, more interaction with faculty, a more supportive campus community, etc. What if someone asks if students learn more at small colleges? Hersh found he had difficulty answering this question, in part because he couldn't give a good answer to this question. For example, how do we know what students are learning? Can we promise parents that, all things being equal, their children will learn more at Whitman College than at the University of Washington? Why should they pay an extra $25,000 per year if their kids aren't going to learn any more here than there? And, indeed, the CLA claims that some schools will be able to answer this question positively. One of five schools who use the CLA can use the results to argue that, yes, our students learn more than students at other schools.

About the CLA

The CLA does *not* claim to measure general intellectual ability and intelligence, nor does it measure content or understanding in a particular discipline. It does not claim to measure the ability, progress, or "value-added" *for any individual student*. Richard Hersh publicly argues that a 90-minute writing exam cannot possibly capture the abilities of any individual student, let alone summarize the value-added by four years of college education.

Rather, the unit of analysis in CLA results is the *institution*, the student body in aggregate. The CLA claims to provide a summative measure at the institutional level, an assessment of the value-added by a school's programs (taken as a whole) with respect to certain learning outcomes. The CLA attempts to measure the level of competency in critical thinking, analytic reasoning, problem solving, and writing. Results are meant to alert faculty and administrators to whether the level of performance attained by their students (as a whole) are better, worse, or about what would be expected given the ability level of incoming students.

The CLA consists of two different types of writing exams: the Performance Task and the Analytic Writing Task. Whitman students were randomly assigned one or the other. The Performance Task provides students with a mini-library of documents, such as letters, memos, summaries of research reports, newspaper articles, photographs, diagrams, tables, charts, and interview notes or transcripts. Students are asked to identify the strengths and limitations of alternative hypotheses, points of view, and courses of action. An example:

You are the assistant to Pat Williams, the president of DynaTech, a company that makes precision electronic instruments and navigational equipment. Sally Evans, a member of DynaTech's sales force, recommended that DynaTech buy a small private plane (a SwiftAir 235) that she and other members of the sales force could use to visit customers. Pat was about to approve the purchase when there was an accident involving a SwiftAir 235. You are provided with the following documentation:

1. Newspaper articles about the accident
2. Federal Accident Report on in-flight breakups in single engine planes

3. Pat's e-mail to you & Sally's e-mail to Pat
4. Charts on SwiftAir's performance characteristics
5. Amateur Pilot article comparing SwiftAir 235 to similar planes
6. Pictures and description of SwiftAir Models 180 and 235

Please prepare a memo that addresses several questions, including what data support or refute the claim that the type of wing on the SwiftAir 235 leads to more in-flight breakups, what other factors might have contributed to the accident and should be taken into account, and your overall recommendation about whether or not DynaTech should purchase the plane.

The document library will always contain information that is spurious, silly, correlation but not causation, and other pitfalls. Students have to weigh the value of different sources and make a well-reasoned, evidence-based argument. My impression, looking over the shoulders of students working on a performance task, was that students were guided through a series of questions to help them write their memo, or whatever the task calls for.

The Analytic Writing Task requires two essays: a forty-five-minute "Make an argument," in which students either support or reject a position on some issue; and a thirty-minute "Break an argument," in which they consider the validity of someone else's reasoning. "Make an argument" asks students to react to an opinion. For example:

> Public figures such as actors, politicians, and athletes should expect people to be interested in their private lives. When they seek a public role, they should expect that they will lose at least some of their privacy.

Students can address the issue from any perspective they wish, as long as they provide support for their views. "Break an argument" might present students with a passage like this:

> The following is from an editorial in the Midvale Observer, a local newspaper. "Ever since the 1950's, when television sets began to appear in the average home, the rate of crimes committed by teenagers in the country of Alta has steadily increased. This increase in teenage crime parallels the increase in violence shown on television. According to several national studies, even very young children who watch a great number of television shows featuring violent scenes display more violent behavior within their home environment than do children who do not watch violent shows. Furthermore, in a survey conducted by the Observer, over 90 percent of the respondents were parents who indicated that primetime television programs between 7 and 9 p.m. should show less violence. Therefore, in order to lower the rate of teenage crime in Alta, television viewers should demand that television programmers reduce the amount of violence shown during prime time."

Rather than agree or disagree with the position, students must discuss how well reasoned they find the argument by considering the soundness of its logic.

I suspect most faculty at Whitman would agree that our seniors, no matter what their major, should be able to do pretty well on all of these exercises, and our seniors should be able to do significantly better than our First-Year students. In addition, a random group of our seniors can probably do better (as a whole) than a random group of seniors at Eastern Washington University. These exercises require critical thinking, analytic reasoning, and good writing skills,

which are all general abilities where we want, and expect, our students to excel. If our seniors couldn't do well on this exam, or couldn't outperform students at a regional public university, we'd probably want to know.

What the CLA Does and Does Not Measure

CLA scores are reported to the institution in the aggregate. Hersh has made it clear he does not believe the CLA is a good measure of individual student competence, but, by focusing on the institution in the aggregate, he argues that it *is* a good measure of the value added by the school's instructional and other programs. Its goal is to provide information that will help colleges determine how much students are improving, and whether that improvement is in line with the gains of comparable students at other institutions.

The CLA compares value-added learning across schools by using a school's input characteristics (such as SAT scores) to predict an expected average CLA score. Is the actual average CLA score from a given school near the expected score, given the quality of students who attend that school? If our first-year students have an average SAT score of 1330, they should aggregately do better on the CLA than a sample of first-year students at a school where the average SAT is 1100. Similarly, seniors at our school should score better than seniors at the 1100 SAT school, simply based on how smart they were when they arrived.

And indeed, average CLA scores for an institution are highly correlated with average SAT scores, and the CLA is able to create a very accurate predictive model for CLA scores (R^2 = ~.75). Where the CLA claims to measure "value-added," and where it claims to show which schools have students who "learn more than other schools" is in the difference between the predicted and actual scores. For example, if seniors at a particular school scored significantly better than predicted, this indicates more value-added than other schools. In this way, comparisons between schools are not based on how high students score, but how much better-than-expected students score. This allows for more meaningful comparisons.

So, for Whitman College, it is not enough for our seniors to simply score better than our first-year students, because this happens at every other college. However, if our seniors score significantly higher than predicted, we're doing well, and we can make the claim that the education we provide is superior to that at other schools. If our seniors score at expected, we can't necessarily make the claim that our students have learned much more than students at an 1100 SAT school, we can only argue that we did better on the CLA because our students were smarter when they got here. (Or, of course, we can argue that the instrument is flawed and not worth much).

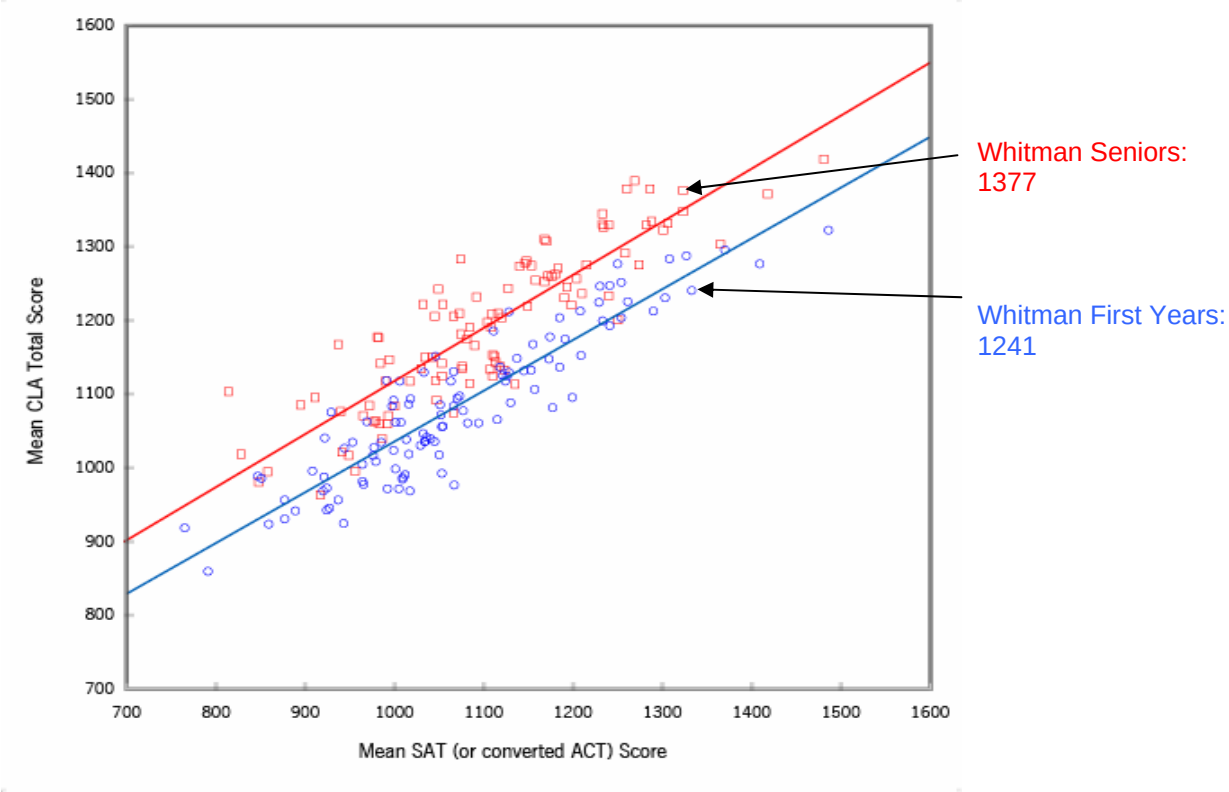
Whitman College CLA Results

A random sample of first-year students and seniors were invited to take the CLA. Students were paid a stipend of $25 after completing the exam. Interestingly, we had a more difficult time finding willing first-year students, and had to draw a second random sample to get the necessary number of students. The CLA is completed on-line, and after logging in about half the students were randomly assigned to the performance task, and about half were randomly assigned to the analytic writing task.

So, how did we do? After taking into account our students' incoming academic abilities, our first years and seniors both scored about where they were expected to score. **Figure 1** shows

first year students and senior scores in relation to their expected scores. Each blue circle represents the average SAT-CLA scores for first-year students at a particular school, each red square represents the average SAT-CLA scores for seniors at a school. The gap between first-years and seniors is important, but more important is the distance seniors score from the prediction line. Black arrows point to the Whitman scores.

Figure 1: Expected and Actual CLA Scores



While our students out-performed nearly all other schools, we scored very close to where we were expected to score. Here is a summary of our results:

- Based on the average SAT score (1333) of the 70 freshmen who completed the exam, their expected average CLA score was 1264. Our freshmen scored 1241, which is considered to be At Expected (*-23 CLA scale points and -0.6 standard errors from the line*). While the average was slightly lower than expected, it was not significantly lower. In other words, we could reasonably expect a different sample of our first year students to score right about the expected score.

- Based on the average SAT score (1323) of the 73 seniors who completed the exam, their expected average CLA score was 1351. Our seniors scored 1377, which is At Expected (*26 CLA scale points and 0.5 standard errors from the line*). While the average was slightly higher than expected, it was not significantly higher. In other words, we could reasonably expect a different sample of our seniors to score right about the expected score.

- Based on the average SAT scores of our freshmen and seniors, we would expect a difference of 87 CLA scale points. The difference between how our seniors and freshmen scored was 136 points. On this measure, we performed better than 80 percent of institutions, and did significantly better than our expected value-added (*49 CLA scale points and 1.1 standard error units from expected*).

Interestingly, our seniors scored much better on the Performance Task, and much better than students at other schools, compared to their performance on the Analytic Task. In particular, seniors were weakest on the "make-an-argument" section of the Analytic Task. The difference between first year and senior scores was smallest on the "make-an-argument" section of the exam. **Table 1** summarizes mean differences between Whitman Senior scores and the mean scores of all seniors taking the CLA to show the relative strength of our performance task scores.

Table 1: Mean CLA Scores for Whitman Seniors and All Other Seniors

	Whitman Seniors	All CLA Seniors	Difference
Performance Task	1431	1195	236
Analytic Writing Task	1323	1224	99
Make-an-Argument	1290	1197	93
Break-an-Argument	1354	1237	117
SAT Score	1323	1104	219

Next Steps

According to the Collegiate Learning Assessment, Whitman students improve their critical thinking, analytic reasoning, problem solving, during their time in college. We can make this claim for two reasons. First, our seniors scored much better than our first year students, and both groups of students scored about where they would be expected to score based on their pre-college academic credentials. Second, while seniors would be expected to out-perform freshman at any respectable 4-year college or university, the difference between Whitman senior and first-year scores was greater than expected, and better than 80% of other schools participating in the CLA.

At this point, I would like to pose two questions regarding the use of our results. 1) Are we satisfied with our CLA results? According to the CLA, we did pretty well, but we aren't one of their top "value-added" schools. 2) To what extent do we want to explore correlates to CLA scores? For example, we can compare average CLA scores across academic divisions and see if certain aspects of our academic program improve critical thinking, analytic reasoning, problem solving, and writing.

These are both indirect ways of asking how much stock we want to put in the results of this exam. There is obviously a danger of putting too much importance on these results, because, to use a cliché, the amount of learning that takes place at Whitman cannot be summed up in a single 90 minute exam. However, we have the results, and taking some time to explore their implications might prove to be a fruitful exercise. How we answer the two questions above will help us take (or not take) the next step.