

Summary of Work for the Teagle Foundation Value-Added Study:
Measuring Intellectual Development and Civic Engagement
through Value-Added Assessment

Consortium of Alma College, Augustana College, Gustavus Adolphus College,
Illinois Wesleyan University, Luther College, and Wittenberg University

July 19, 2007

In our work for the Teagle Foundation the six consortium schools are exploring skills and dispositions central to a liberal-arts education: writing, critical thinking/analytic reasoning, and civic engagement. All the consortium schools seek to educate students to become clear and effective thinkers, to communicate effectively, and to productively engage with the world while enrolled and in the years after graduation. This connection between intellectual development—particularly openness to and consideration of multiple perspectives—and awareness of larger moral and social issues has long been addressed in the literature (Perry, Belenky, Clinchy, King *et.al.*) and is an emphasis of our collaborative efforts.

Our overall goals for our collaborative assessment work include the following:

To determine the effect that our schools have on student growth in writing, critical-thinking, and civic engagement;

To explore ways to do authentic value-added assessment more effectively;

To understand better what the scores of nationally normed instruments, including the NSSE and CLA, mean for our campuses; and

To determine if any consortium schools are outliers and to examine what campus practices may lead to the greater student growth that the schools achieve.

Our work addresses the “value added” by our schools by measuring changes in the three areas referred to above from first-year to junior or senior year. We use Alexander Astin’s “Input-Environment-Output” model as a framework to delineate and control for input effects (measured by ACT/SAT scores, high-school rank, and parental income) and to help us identify environmental factors (first-year seminars, writing intensive classes, international study, etc.) that more readily foster student growth. Because the curricula and educational opportunities of the consortium schools vary, we have an opportunity to explore which approaches lead to greater student gains.

We seek to measure student growth using nationally normed survey instruments such as the National Survey of Student Engagement, the College Student Survey, and the HEDS Alumni Survey, among others. The consortium is also using two direct measures. In June of 2006 and 2007 faculty members from all six schools met and assessed representative samples of student work for writing quality and critical-thinking, using rubrics developed by the schools. In fall 2005 and spring 2006 we administered the Collegiate Learning Assessment (CLA) to first year and senior students. We have explored correlations between the scores on the NSSE, CLA, and our own faculty’s rating of student work.

Writing Assessment—Year One (2005-2006)

After a planning meeting at Augustana in August 2005 the consortium developed a draft of a writing rubric, which was refined during the academic year and used at Alma College in June 2006.

We assessed argumentative/case-building papers because they are a kind of writing common to all schools and would be appropriate for critical-thinking assessment in year two of the study. Because all schools were administering the Collegiate Learning Assessment, we sought to collect writing from every student who took the test. As a back-up we also collected a representative sample of papers from first-year students and juniors. We looked for the value added longitudinally from early in the first year to the end of the first year and cross-sectionally from first year to junior year.

In June 2006, the schools met at Alma College to assess student writing using the rubric. Each school was asked to bring 130 papers, 65 from first-year students and 65 from juniors. Guided by a chief reader with 17 years of experience in essay assessment with the College Board, two to three faculty members from each school assessed papers in terms of main idea/thesis, argument, evidence, organization, readability, conventions, and overall impression on a five-point scale. (See Appendix 1.) Each paper had a discrete random number. All other identifying information concerning school, student, and class level was masked to maintain student and teacher anonymity and to minimize the possibility that readers could make mental adjustments for the writer's year in school. All told, over 600 papers were read. Inter-rater reliability was .937 (Guzman Split Half) and .938 (Cronbach's alpha).

Means and standard deviations of the faculty members' scores were computed, and student change in writing skill was expressed as a percentile change. The first-year fall term to first-year spring term change is summarized in Table 1. By prior consortia agreement, school identities are masked.

TABLE 1: First-Year Students Fall to First-Year Students Spring
(N= 339 fall, 243 spring)

	FY Fall Paper Average	FY Spring Paper Average	Fall/Spring Difference	Pooled Standard Deviation	Effect Size	Percentile Change
	2.67	3.08	0.41	0.56	0.73	27%
	3.11	3.49	0.38	0.66	0.56	22%
	2.58	2.95	0.37	0.69	0.54	20%
	2.43	2.69	0.26	0.59	0.46	17%
	2.67	2.83	0.16	0.65	0.25	10%
	2.73	2.57	-0.16	0.62	-0.26	-10%
Average	2.69	2.85	0.24	0.63	0.38	14.5%

The change in the first year ranges from -10% to 27%. There is no national data on gains in writing skill in the first year alone. It is worth noting that three schools showed effect-size gains over .50 standard deviations. This effect size is what Ernest Pascarelli and Patrick Terenzi report in their meta-analysis of studies that were done prior to the 1990s for first-year to senior-year gains. Studies after 1990 report a .77 effect size for "English skills," a category however that includes not only "writing" but also "reading" and "literature" (Ernest Pascarelli and Patrick Terenzi, *How College Affects Students, Volume II*, 572-573).

We also calculated scores for growth in writing from first to junior year, as shown in Table 2.

TABLE 2: First-Year Students Fall to Junior Students Scores
(N= 339 first-year, 288 junior)

	FY Paper Average	JR Paper Average	JR/FY Difference	Pooled StDev	Effect Size	Percentile Change
	2.73	3.22	0.49	0.69	0.69	25%
	2.58	3.10	0.52	0.75	0.69	25%
	2.67	3.05	0.38	0.68	0.56	21%
	2.67	3.03	0.36	0.70	0.51	20%
	2.43	2.79	0.36	0.69	0.52	20%
	3.11	3.15	0.04	0.67	0.06	2%
Average	2.70	3.05	0.36	0.70	0.51	19%

Percentage change ranges from 2% to 25%. All but one school surpassed the benchmark of .50 for first to senior-year gains. Interestingly, though students at most schools continued to show gains in writing skill, the effects were modest, and some schools students' seemed to lose gains they had made in the first year. (See Table 3.) Reasons for this apparent loss will be a subject for future analysis in light of critical-thinking assessment data recently compiled. We have some evidence that the selection of papers may explain some of the anomalous scores.

TABLE 3: First-Year Students to Junior Students Percentile Change
(N= 339 first-year, 288 junior)

	Percentile Change First Year Fall to Spring	Percentile Change First Year Fall to Junior Year	Percentile Change First to Junior Year
	-10%	25%	35%
	10%	21%	11%
	20%	25%	5%
	18%	20%	2%
	27%	20%	-7%
	22%	2%	-20%
Average	14.5%	19%	4.5%

Correlations of CLA, NSSE, and Local Ratings

We ran correlations among the six institutions between the scores on the CLA and NSSE benchmark scores and the CLA and the rating of papers by the faculty. Because our sample size is only six institutions, we need to be tentative about any conclusions using data from the CLA. Our results for the correlations between the NSSE benchmark scores and the CLA Total Score are shown in Table 4 on the next page.

Except for a cluster of statistically significant correlations around Active and Collaborative Learning and Student/Faculty Interaction for the CLA Total Score and the Analytic Writing Task Score, the correlations are low or negative. Interestingly, there are no statistically significant correlations between the NSSE and the Performance Task. But even so, the correlations tend to be strongest there for Active and Collaborative Learning and Student/Faculty Interaction in the senior year.

TABLE 4: Correlation of NSSE Benchmarks with CLA Total Score and Sub-scores
(N=6)

NSSE Benchmark	CLA Total Score	CLA Analytic Writing Difference	CLA Performance Task Difference
Level of academic Challenge-FY	-.239	.095	-.479
Level of academic Challenge-SR	-.377	.652	-.002
Active & Collaborative Learning-FY	.654	.822 (.045)	.295
Active & Collaborative Learning-SR	.934 (.006)	.948 (004)	.658
Student/Faculty Interaction—FY	.733	.830 (.041)	.382
Student/Faculty Interaction—SR	.853 (.031)	.811	.628
Enriching Educational Experiences—FY	-.437	-.200	-.516
Enriching Educational Experiences—SR	.200	.222	.099
Supportive Campus Environment—FY	.097	-.174	.397
Supportive Campus Environment—SR	.106	.046	.181

We also ran correlations between CLA scores and the faculty’s rating of papers for writing quality. The first row in Table 5 correlates the change in paper scores from fall to spring for first-year students as rated by faculty using the rubric. The second row correlates change in papers scores from first year to junior year..

Table 5: Correlation of Faculty Rating of Writing Quality and CLA Scores
(N =6)

Paper Ratings	Total Score Difference	Performance Task Difference	Analytic Writing Score Difference
First-Year Fall to First-Year Spring Difference	-.161	.024	-.250
First -Year Fall to Junior Difference	.206	-.103	.426

It is worth noting that the majority of the correlations are low or negative and, with the exception of the modestly positive correlation of .426 for first-year to junior growth in writing skill and the Analytic Writing difference score. One might expect a stronger correlation with the Performance Task difference score, as this test is a more complex task akin to the argumentative papers that upper-class students were asked to do.

Critical-Thinking Assessment—Year Two (2006-2007)

Faculty from all six schools assessed critical thinking at Luther College in June of this year using a rubric developed by a faculty group during a workshop at Illinois Wesleyan University with the help of Washington State University’s Dr. Bill Condon, a lead member in a FIPSE- funded project to incorporate critical thinking in the university’s curriculum. Rather than adopt an existing rubric, the Teagle participants

decided together what qualities characterize critical thinking. Their definition emphasizes recognizing and identifying an issue or problem and seeking open-mindedly to clarify and then answer it, using valid evidence and support. Furthermore, a critical thinker acknowledges the value and limits of evidence and is explicit about the connection of conclusions to larger questions of meaning and importance. Though nothing in this definition is new, it is flexible enough to apply to student work in various disciplines. It is also one that applies to work outside the academy where critical thinkers often must address open-ended or ill-defined problems. (See Appendix 2.)

The assessment of papers was similar to that at Alma College, although because of the length of the papers, readers were not able to assess all 130 papers from each school for both cohorts. All told, the 28 faculty and administrators completed 1,441 readings of 670 papers. As measured by standard-deviation scores, the critical-thinking readers were somewhat more likely to use all levels of the rubric than the readers for writing quality at Alma did. The standard deviation of the critical-thinking ratings ranges from 1.03 to 1.15 versus 0.65 to 0.71 for the writing assessment scores. The inter-rater reliability is high with an overall Cronbach's alpha of .967.

The analysis so far has been limited to a first look at value added based on raw scores. The critical-thinking rubric has eight sub-scores and a Holistic Score. The Holistic Scores ratings correlated positively though weakly with high-school percentile rank (0.24) and with ACT Composite score (0.32) and somewhat more strongly with end-of-year college GPA (0.41). Analysis of the critical-thinking ratings showed that all the scores were highly correlated. For example, the eight sub-scores correlated with the Holistic Score between 0.85 and 0.94. The high correlations indicated that readers did not see some factors as significantly more important in rating whether the critical thinking of a student's work was of low, middle, or high quality. In the data shared in this report, the final rating for each paper is the average of the two closest scores on the Holistic Rating.

Table 6 shows the percentage gain using the Holistic Score for each school. Ratings for all the sub-scores and the Holistic Score were computed for change from first-year to junior/senior year. The difference was then used to compute an effect size, which is then expressed in the table as a percentile gain or loss.

Table 6: Change in Critical-Thinking Rating First Year to Junior/Senior Year

	Holistic First Year to JR/SR	Effect Size	Percentile Gain/Loss	ACT Composite Percentile Change FR to JR/SR Samples
	0.86	0.79	29%	24%
	0.70	0.62	23%	7%
	0.71	0.57	22%	13%
	0.51	0.45	17%	20%
	0.37	0.36	14%	18%
	-0.56	-0.48	-19%	-7%
Average	0.44	0.38	14.3%	14.8%

The range of change varies considerably from a positive 29% to a negative 19%, with a percentile gain mean of 14.8%. These scores are not yet adjusted for the influence of other factors. Note, for example, ACT Composite score, one portion of the demographic data that can help explain differences in Holistic Scores. Students from the school in line 1 demonstrated a 29% gain in critical thinking with a 24% percentile change in the ACT Composite score from first year to junior/senior year. The school in line 2 sample shows a slightly smaller critical-thinking gain of 23%, though its junior/senior sample's ACT Composite score was

only 7% higher than that of its first-year students. In short, a more precise computation of the value added must wait until the other demographic input variables are factored in. We also have yet to explore these results in light of the CLA and NSSE data.

Appendix 1: Writing Skills Rubric

Teagle Study Rubric

June 2006

	Unacceptable 1	Beginning 2	Competent 3	Skilled 4	Exemplary 5
MAIN IDEA/ THESIS	Overall position is not evident. Topic as expressed is superficial, trite, or clichéd.	Overall position is evident, but often simplistic. Topic is simplistic and one-dimensional.	Overall position is clear with a sense of developed ideas. Topic is interesting and significant, but not deeply explored in needed areas.	Overall position is clear and developed. Topic is interesting, significant, and is engaged from several angles.	Overall position is well-articulated and developed with appropriate evidence. Topic is interesting, significant and intellectually challenging with multiple facets addressed.
ARGUMENT	No argumentative structure is evident. Ideas are unconnected.	Argument structure is rudimentary. Claims are repeated rather than developed. Few objections are addressed or with misrepresentations.	Argumentative structure is evident but sometimes simplistic. Objections are addressed but formulaically.	Argumentative structure is evident. Objections are taken seriously and typically addressed fair-mindedly.	Argumentative structure is clearly evident. Objections are taken seriously and addressed fair-mindedly and with great skill.
EVIDENCE	Claims are not supported by reliable evidence from credible sources, making the paper unconvincing.	Some claims are supported by valid, reliable evidence, but support is inconsistent, making the paper less than convincing.	Claims are typically supported by valid, reliable evidence from credible sources, making the paper for the most part convincing.	Claims are almost always supported by valid, reliable sources, so that the paper is generally convincing.	Claims are supported by reliable, valid evidence from credible sources and effectively synthesized in a very convincing manner.
ORGANIZATION	Ideas appear unconnected. Several paragraphs are incoherent, lacking clear topic sentences and developed by restatement; they may contain irrelevant information. Paper shows serious lack of unity and coherence. Introduction and /or conclusion may be weak, trite, or nonexistent.	The paper is organized, though simplistically. Paragraphs are occasionally incoherent, without strong topic sentences and clear development. An introduction and conclusion are attempted but are perfunctory or formulaic. The introduction includes an overly general thesis, and the conclusion simply restates that thesis.	The introduction sets a context for the paper, states a thesis, though in a predictable way. Paragraphs are usually clear with serviceable topic sentences, development, and information. Main points are logically structured. Transitions provide coherence, but may be formulaic (first, second, third, etc.). The conclusion summarizes the paper, but does not explore implications or significance.	The introduction sets the context for the paper, states a clear thesis. Paragraphs are coherent with strong topic sentences, developing systematically so that meaning is clear. Main points are clear and logically structured. Transitions provide a sense of coherence. The conclusion summarizes the paper and makes some effort to explore implications and significance.	The introduction skillfully captures reader attention and sets the context for the paper. Paragraphs are coherent with apt topic sentences, developed so meaning is especially clear and easy to follow. The thesis is clear and effective. Main points are clear and logically structured. Transitions provide a sense of coherence. The conclusion summarizes and explores implications and significance.

READABILITY	Awkward phrasing, unskilled or inappropriate voice/tone, and unsophisticated and/or imprecise vocabulary hinder understanding.	Awkward phrasing, unskilled or inappropriate voice/tone, and unsophisticated and/or imprecise vocabulary distract from the paper's ideas.	Phrasing is generally competent; voice/tone and vocabulary are generally suitable for the paper's ideas and only occasionally work against its ideas.	Clear phrasing, appropriate shifts in voice and tone, and vocabulary enhance the paper's ideas.	Skilled phrasing, appropriate shifts in voice and tone, and apt word choice create an inviting paper.
CONVENTIONS	Numerous errors in grammar, usage, spelling and punctuation seriously impede meaning. Necessary documentation is missing.	Several errors in grammar, usage, spelling, and punctuation distract the reader and impede meaning. Problems with needed documentation exist.	Errors in grammar, usage, spelling, and punctuation are noticeable, but do not seriously impede the reader. Documentation is usually correct.	There are occasional errors in grammar, usage, spelling, and punctuation that do not impede the reader. Documentation of sources is correct.	There are very few or no mechanical errors in the paper. Documentation of sources is correct.
OVERALL IMPRESSION	The writer is unable to construct and present a significant position. Paragraphing and overall organization hinder effectiveness. Ideas are asserted rather than developed or are seriously underdeveloped. Language is substandard with errors at the sentence level. Below college-level writing	The writer presents a significant position that falls short of being convincing. Frequent though not pervasive problems at the sentence-level. Paragraphing is inconsistent. Overall organization and support are rudimentary. Remedial college level writing.	The writer presents a significant position that is generally convincing, but has some weaknesses. Paragraphs are typically organized and add to the development of ideas. Support is good, but sometimes inadequate. Organization is evident but sometimes simplistic. There are occasional, but not overly distracting, sentence-level errors, Acceptable if sometimes uninspired college work.	The writer presents a significant and thoughtful position that is for the most part convincing. Paragraphs are well organized and contribute to the development of ideas. Support is good with infrequent soft spots. Organization is clear. Sentence-level errors are infrequent. Skilled college-level writing.	The writer presents a significant and thoughtful position that is convincing and at times thought-provoking. Paragraphs are skillfully organized and add to the development of ideas. Support is sound with rare or no soft spots. Sentence-level errors are rare to non-existent. Strong college-level writing.

Appendix 2: Critical-Thinking Skills Rubric

Teagle Scoring Guide for Critical Thinking May 2007

This guide sets out several habits of mind that students must develop in order to succeed at critical thinking in their academic work. Academic work includes but is not limited to such activities as argumentation; interpretation; developing proofs, theorems, and case statements; model building; analysis; and creative projects. Faculty in various disciplines are invited to *adapt* this scoring guide to fit the contexts of their disciplines, the courses they teach, and the assignments that they present in those courses. Each habit of mind identified below is accompanied by descriptors for a range of performance exhibited in actual student work.

1. Problem: Recognizes from readings, experience, data, or observation a problem, question, or issue to address.

Emerging		Developing		Mastering	
1	2	3	4	5	6
Does not identify a problem, question, or issue or identifies an inappropriate problem, question, or issue. May be confused or represent the problem, question, or issue inaccurately. Does not establish problem's significance/relevance.		Identifies a problem, question, or issue and presents it clearly, if simply. May recognize some of the nuances, but does so inconsistently. Acknowledgment of problem's significance/relevance is too simple.		Identifies the main problem, question, or issue, as well as embedded or implicit ones; and identifies them clearly, addressing their relationships to each other. Recognizes the nuances of the problem, question, or issue, including the relevance/significance.	

2. Central/Main Idea: identifies and presents an approach and position to address the problem/issue raised.

Emerging		Developing		Mastering	
1	2	3	4	5	6
The project displays no central approach or controlling idea, or that idea remains unimportant to the work.		Presents an approach/controlling idea that addresses the issue or problem, though sometimes in an unsophisticated or simplified way.		Presents an approach/controlling idea that addresses the issue or problem raised, in a complex, sophisticated way.	

3. Perspective(s): Identifies and considers salient perspective(s), position(s), and context(s).

Emerging		Developing		Mastering	
1	2	3	4	5	6
Deals superficially with a single perspective. Even when applicable, fails to acknowledge other possible salient perspectives. Lacks a sense of fairness and open-mindedness. May not be aware of having a perspective or may not present an appropriate perspective.		Maintains a single perspective. When applicable, acknowledges other possible salient perspectives. Is mostly fair and open-minded. When appropriate to subject, student demonstrates some awareness of his or her own perspective and its influence on the approach to the task.		Skillfully conveys a single perspective and, when applicable, addresses and accommodates all other salient perspectives well. Is consistently fair and open-minded. When appropriate to subject, student shows a deep and detailed awareness of his or her own perspective and its influence on the approach to the task.	

4. Supporting Data/Evidence: Includes supporting data/evidence and assesses its quality.

Emerging		Developing		Mastering	
1	2	3	4	5	6
Provides very little data/evidence to support its position, or the data/evidence selected is low quality or irrelevant. Does not seriously assess support, to distinguish among fact, opinion, and value judgments.		Provides data/evidence to support its position; some data/evidence is low quality or irrelevant. Attempts, though sometimes mistakenly, to assess support, to distinguish among fact, opinion, and value judgments.		Provides ample evidence to support its position; almost all data is high quality and clearly relevant. Clearly assesses support, distinguishing among fact, opinion, and value judgments.	

5. Depth of thought: Deeply engages in the work.

Emerging		Developing		Mastering		
1	2	3	4	5	6	
Demonstrates little engagement with the work. The treatment remains shallow, over-simplified, and limited in focus and usefulness. Exhibits little or no ability to deal with ambiguity.		Engages the work, in places pushing the treatment to greater depth and complexity, approaching it with a spirit of exploration, or expanding the focus as needed to do the work justice. Level of complexity throughout is adequate but in need of greater development. Acknowledges ambiguity.				Engages the work fully, pushing to achieve depth and complexity, exploring and where necessary expanding the boundaries of the work. Treatment is complex, sophisticated, imaginative, and nuanced. Acknowledges and effectively manages ambiguity.

6. Reasoning: employs logic to construct a cogent argument/statement.

Emerging		Developing		Mastering		
1	2	3	4	5	6	
Work has obvious flaws in logic/analysis.		Work is generally sound, but has some flaws in logic/analysis.				Work is very sound, with no flaws or only minor flaws in logic/analysis.

7. Development: Strategically organizes and styles the work.

Emerging		Developing		Mastering		
1	2	3	4	5	6	
Employs seemingly random and/or inappropriate organization, and, where applicable, genre, and/or medium.		Makes some sensible choices of organization. Where applicable, selects an appropriate genre and/or medium.				Makes appropriate choices of organization and, where applicable, genre and/or medium. While coherent, also engages with and tests rules or boundaries of the work.

8. Conclusions/Consequences: identifies and assesses strengths and weaknesses of choices, conclusions, implications, and consequences.

Emerging		Developing		Mastering	
1	2	3	4	5	6
Fails to identify conclusions, implications, and consequences of the issue or the key relationships between the other elements of the problem, such as assumptions, contexts, data, evidence, organization, genre, medium. Seems unaware of limits of evidence and conclusions.		Identifies some conclusions, implications, and consequences and/or fails to spell out conclusions, implications, and consequences as clearly as possible. Some awareness of the limits of evidence and conclusions.		Identifies and discusses conclusions, implications, and consequences considering assumptions, context, data, evidence, organization, genre, medium. Objectively reflects upon their own assertions, including limits of evidence and conclusions.	

9. Holistic Rating. What rating would you give this work as a whole?

Emerging		Developing		Mastering	
1	2	3	4	5	6

(Absent)	Emerging (Recognizable)	(Inconsistently Competent)	Developing (unsophisticated)	(Competent but sophisticated)	Sometimes Mastering (Frequently sophisticated)
1	2	3	4	5	6

Classification of Instructional Programs Codes (CIP 2000)

2-digit code	Program
3	Natural Resources/Conservation/Environmental Studies
5	Area, Ethnic, Cultural, and Gender Studies
9	Communication Studies/Speech Communication and Rhetoric
11	Computer and Information Sciences, General
13	Education, General
14	Engineering, General
16	Foreign Languages and Literatures, General
22	Pre-Law Studies
23	English Language and Literature, General
24	Liberal Arts and Sciences/Liberal Studies
25	Library Science/Librarianship
26	Biology/Biological Sciences, General
27	Mathematics, General
30	Biological and Physical Sciences
31	Parks, Recreation and Leisure Studies
38	Philosophy and Religious Studies
40	Physical Sciences
42	Psychology, General
43	Criminal Justice/Protective Services
44	Public Administration and Social Services, General
45	Social Sciences, General
50	Visual and Performing Arts, General
51	Health Services/Allied Health/Health Sciences, General
52	Business/Commerce, Management, Marketing
54	History, General
99	General Studies or no discipline available.