University of
Wyoming

Acknowledgements and Notes

This report presents the results of your institution across the three phases of this longitudinal study.

Additional summary findings from of this study will be published in the coming months.

Institutional Report

Appendices

Institutional Report

The Collegiate Learning Assessment (CLA) offers an authentic approach to assessment and improvement of teaching and learning in higher education. Over 400 institutions and 180,000 students have participated to date. Growing commitment on the part of higher education to assess student learning makes this a good time to review the distinguishing features of the CLA and how it connects to improving teaching and learning on your campus.

The CLA is intended primarily to assist faculty, department chairs, school administrators and others interested in programmatic change to improve teaching and learning, particularly with respect to strengthening higher order skills.

The CLA helps campuses follow a continuous improvement model that positions faculty as central actors.

CLA Education (described on page 14) does just that by focusing on curriculum and pedagogy and the link between assessment and teaching and learning.

The continuous improvement model also requires multiple assessment indicators beyond the CLA because no single test to benchmark student learning in higher education is feasible or desirable.

This, however, does not mean certain skills judged to be important by most faculty and administrators across virtually all institutions cannot be measured; indeed, the higher order skills the CLA focuses on fall into this measurable category.

The CLA presents realistic problems that require students to analyze complex materials. Several different types of materials are used that vary in relevance to the task, credibility, and other characteristics. Students' written responses to the task are graded to assess their abilities to think critically, reason analytically, solve problems, and communicate clearly and cogently.

The institution—not the student—is the initial primary unit of analysis. The CLA is designed to measure an institution's contribution, or value added, to the development of these competencies, including the effects of changes to curriculum and pedagogy.

The CLA uses detailed scoring guides to precisely and reliably evaluate student responses. It also encourages institutions to compare their student learning results on the CLA with learning at other institutions and on other assessments.

The signaling quality of the CLA is important because institutions need to benchmark (have a frame of reference for) where they stand and how much progress their students have made relative to the progress of students at other colleges. Otherwise, how do they know how well they are doing?

Yet, the CLA is not about ranking institutions. Rather, it is about highlighting differences between them that can lead to improvements in teaching and learning.

While the CLA is indeed an assessment instrument, it is deliberately designed to contribute directly to the improvement of teaching and learning. In this respect it is in a league of its own.

2: Your Results

Table 1 below provides the number of your students with both Entering Academic Ability (EAA) scores and CLA scores in different phases of the longitudinal assessment.

EAA scores represent SAT Math + Verbal, ACT Composite, or Scholastic Level Exam (SLE) scores on the SAT scale.

Per sampling restrictions of Phase 1 of the assessment, the sample of students throughout the report also have EAA scores.

1

Number of your students with CLA scores in different longitudinal phases

	Phase 1 and Phase 2	Phase 2 and Phase 3	Phase 1 and Phase 3
Performance Task	148	77	89
Analytic Writing Task	141	77	87
Make-an-Argument	143	78	88
Critique-an-Argument	144	77	87
Total Score	141	76	86
Phases referenced in columns above	F 2005	F 2005	F 2005
	S 2007	S 2007	S 2007
	S 2009	S 2009	S 2009

2: Your Results

For each longitudinal phase, Table 2 below presents summary statistics for your students including counts, means, 25th and 75th percentiles, and standard deviations.

We also calculated these statistics across all students and all schools. See Tables 7 and 8 on pages 10 and 11.

2

Summary statistics for your students participating in the longitudinal assessment

Phase 1	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	275	1014	1131	1232	152
Analytic Writing Task	263	984	1101	1196	141
Make-an-Argument	266	942	1120	1225	184
Critique-an-Agument	268	1018	1082	1167	158
Total Score	262	1033	1117	1197	115
EAA Score	272	1030	1137	1260	145

Phase 2	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	148	1046	1182	1282	179
Analytic Writing Task	146	980	1098	1196	152
Make-an-Argument	147	942	1099	1225	179
Critique-an-Agument	147	1018	1093	1167	161
Total Score	146	1041	1140	1223	134
EAA Score	147	1070	1154	1260	130

Phase 3	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	89	1093	1186	1292	141
Analytic Writing Task	89	1031	1152	1252	174
Make-an-Argument	90	959	1127	1285	218
Critique-an-Agument	89	1049	1175	1308	185
Total Score	88	1099	1169	1261	131
EAA Score	90	1030	1147	1260	138

Table 3 provides summary statistical data on your students across all three phases. We present these data in three groups to examine change across different phases.

The "effect size" column of Table 3 is particularly important. This column indicates how much change occurred between two different phases.

The larger the positive effect size, the greater the improvement. Effect sizes greater than 0.50 are generally considered "large." Negative effect sizes would indicate that the students scored higher during an earlier phase. To place your performance in context, Figure A on page 5 plots these effect sizes versus all other participating institutions.

The last column shows the correlation between students' scores in different phases. A high positive correlation indicates that the students who scored relatively highly in one phase (relative to their classmates) also tended to score relatively highly in another phase. In other words, the improvement in scores was fairly consistent across students.

3

Comparison of your students' scores in Phase 1 (Fall 2005), Phase 2 (Spring 2007), and Phase 3 (Spring 2009)

Phases 1 and 2	Phase 1		Phase 2		Difference		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Effect Size*	Mean Score Correlation
Performance Task	148	1140	154	1182	179	42	25	0.27	0.40
Analytic Writing Task	141	1113	142	1101	152	-12	10	-0.09	0.56
Make-an-Argument	143	1130	181	1103	179	-27	-2	-0.15	0.49
Critique-an-Agument	144	1094	165	1095	161	1	-4	0.01	0.41
Total Score	141	1128	113	1144	134	16	21	0.14	0.57

Phases 2 and 3	Phase 2		Phase 3		Difference		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Effect Size	Mean Score Correlation
Performance Task	77	1178	186	1193	134	15	-52	0.08	0.49
Analytic Writing Task	77	1119	144	1159	174	40	30	0.28	0.55
Make-an-Argument	78	1113	171	1133	219	20	48	0.12	0.39
Critique-an-Agument	77	1122	151	1182	186	60	35	0.40	0.50
Total Score	76	1149	137	1175	130	26	-7	0.19	0.64

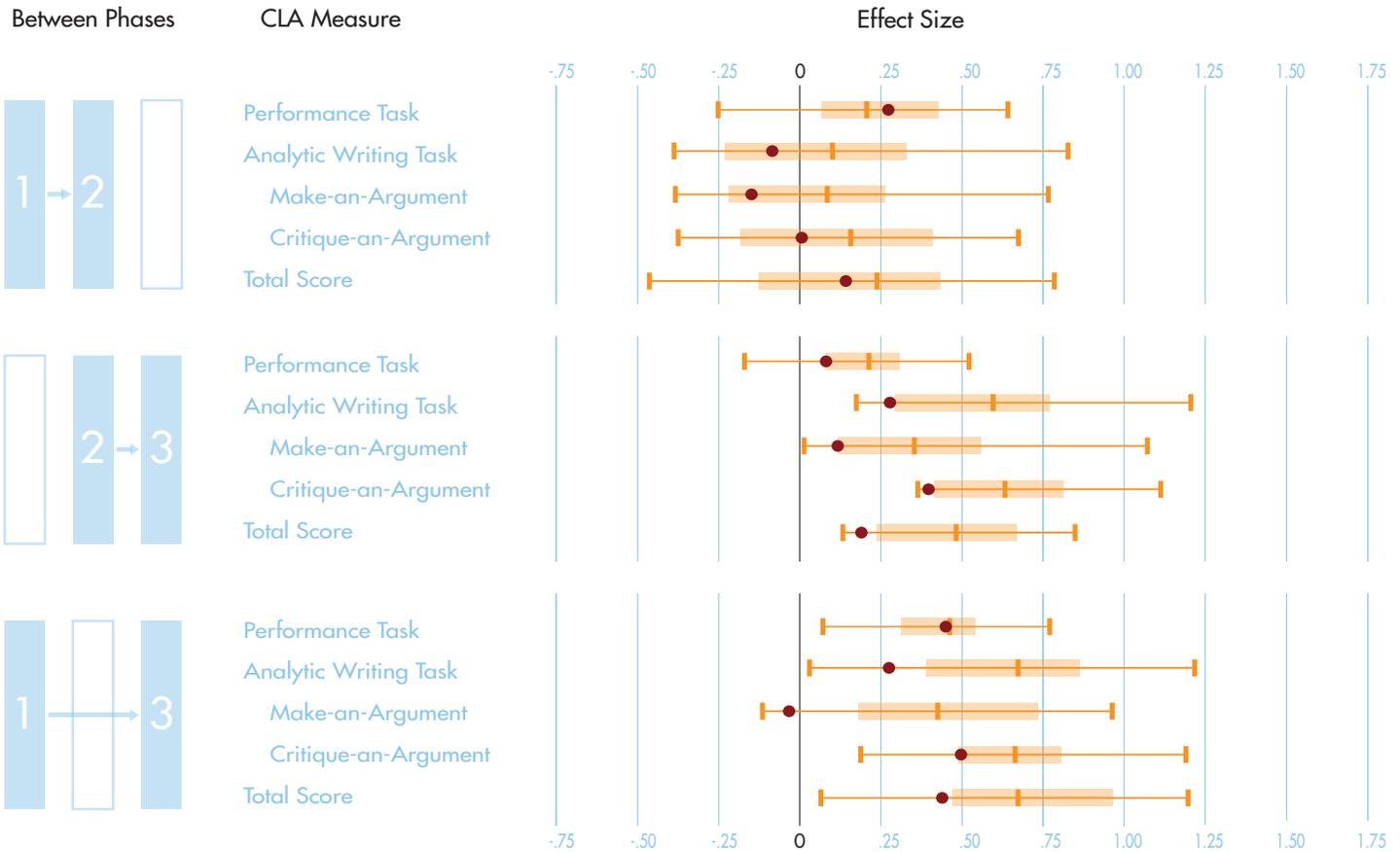| Phases 1 and 3	Phase 1		Phase 3		Difference		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Effect Size	Mean Score Correlation
Performance Task	89	1119	149	1186	141	67	-8	0.45	0.34
Analytic Writing Task	87	1116	138	1154	175	38	37	0.28	0.52
Make-an-Argument	88	1135	183	1129	220	-6	37	-0.03	0.41
Critique-an-Agument	87	1100	157	1178	186	78	29	0.50	0.31
Total Score	86	1120	114	1170	132	50	18	0.44	0.65

* A different method was used previously to calculate effect sizes, so these results differ slightly from those in the 2005–2007 Institutional Report. Previously, we divided mean differences between phases by the CLA score standard deviation of all participating students in the earlier phase, which would tend to underestimate effect sizes. Now, we divide by the standard deviation of your participating students in the earlier phase.

A Mean Score Changes between Phase 1, Phase 2, and Phase 3 by CLA Measure

Effect Size Box and Whisker Plots

Between Phases | CLA Measure | Effect Size

1 → 2
- Performance Task
- Analytic Writing Task
- Make-an-Argument
- Critique-an-Argument
- Total Score

2 → 3
- Performance Task
- Analytic Writing Task
- Make-an-Argument
- Critique-an-Argument
- Total Score

1 → 3
- Performance Task
- Analytic Writing Task
- Make-an-Argument
- Critique-an-Argument
- Total Score

Effect Size axis: -.75 -.50 -.25 0 .25 .50 .75 1.00 1.25 1.50 1.75

Key for Phases

Phase	When	Who
1	Fall 2005	First-year
2	Spring 2007	Rising Junior
3	Spring 2009	Senior

Key for Box and Whisker Plots

5th — 25th — 50th — 75th — Your School — 95th

The Figure A "box and whisker" plots show effect size distributions across CLA measures for longitudinal schools. The top cluster displays effect sizes between Phase 1 (fall 2005 first-year students) and Phase 2 (spring 2007 rising juniors); the middle cluster repeats this for Phases 2 and 3 (spring 2009 seniors); and the bottom cluster does the same for Phases 1 and 3.

Effect sizes are one way to measure change between two time points. Effect sizes were calculated at a school by taking the difference in mean (or average) CLA scores of the same students from two different phases (e.g., Phase 1 and Phase 3). This difference is then divided by the spread of scores (standard deviation) from the first phase (e.g., Phase 1) to produce an effect size.

In each plot, the extreme left hand vertical bar shows the 5th percentile. The "box" itself shows the 25th (left face), 50th (internal vertical line), and 75th (right face) percentile points. The extreme right hand vertical line shows the 95th percentile. The horizontal x-axis shows the effect size (in standard deviation units).

Circles ● identify effect sizes for your school (see Table 3 for values).

Table 4 below indicates whether students scored well above, above, at, below, or well below what would be expected given their scores as first-year students. Results are expressed in the form of value-added scores that correspond to standard errors.

Colleges with scores between -1.00 and +1.00 standard errors from their expected scores are categorized as At Expected. Institutions with scores greater than one standard error (but less than two standard errors) from their expected scores are in the Above Expected or Below Expected categories (depending on the direction of the deviation). The schools with scores greater than two standard errors from their expected scores are in the Well Above Expected or Well Below Expected categories.

Figure B on the next page displays your deviation scores in a scatterplot and documents the regression equations from which they were derived.

The two regression equations (Phase 1 to Phase 2 in blue; Phase 1 to Phase 3 in red) are each based on data from all institutions where at least 25 students received CLA total scores in both phases in the model. Therefore, the sample of institutions is not identical across the testing phases reported here, but the use of inclusive samples maximizes the strength of each equation.

4

Value added results for your school

	Deviation Score	Performance Level	Deviation Score	Performance Level
Performance Task	0.1	At	-0.1	At
Analytic Writing Task	-0.4	At	-0.9	At
Make-an-Argument	-0.5	At	-1.2	Below
Critique-an-Argument	-0.3	At	-0.7	At
Total Score	-0.1	At	-0.7	At
Phases referenced in columns above	F 2005	F 2005	F 2005	F 2005
	S 2007	S 2007	S 2007	S 2007
	S 2009	S 2009	S 2009	S 2009

B Relationship between Spring 2009, Spring 2007, and Fall 2005 CLA Performance
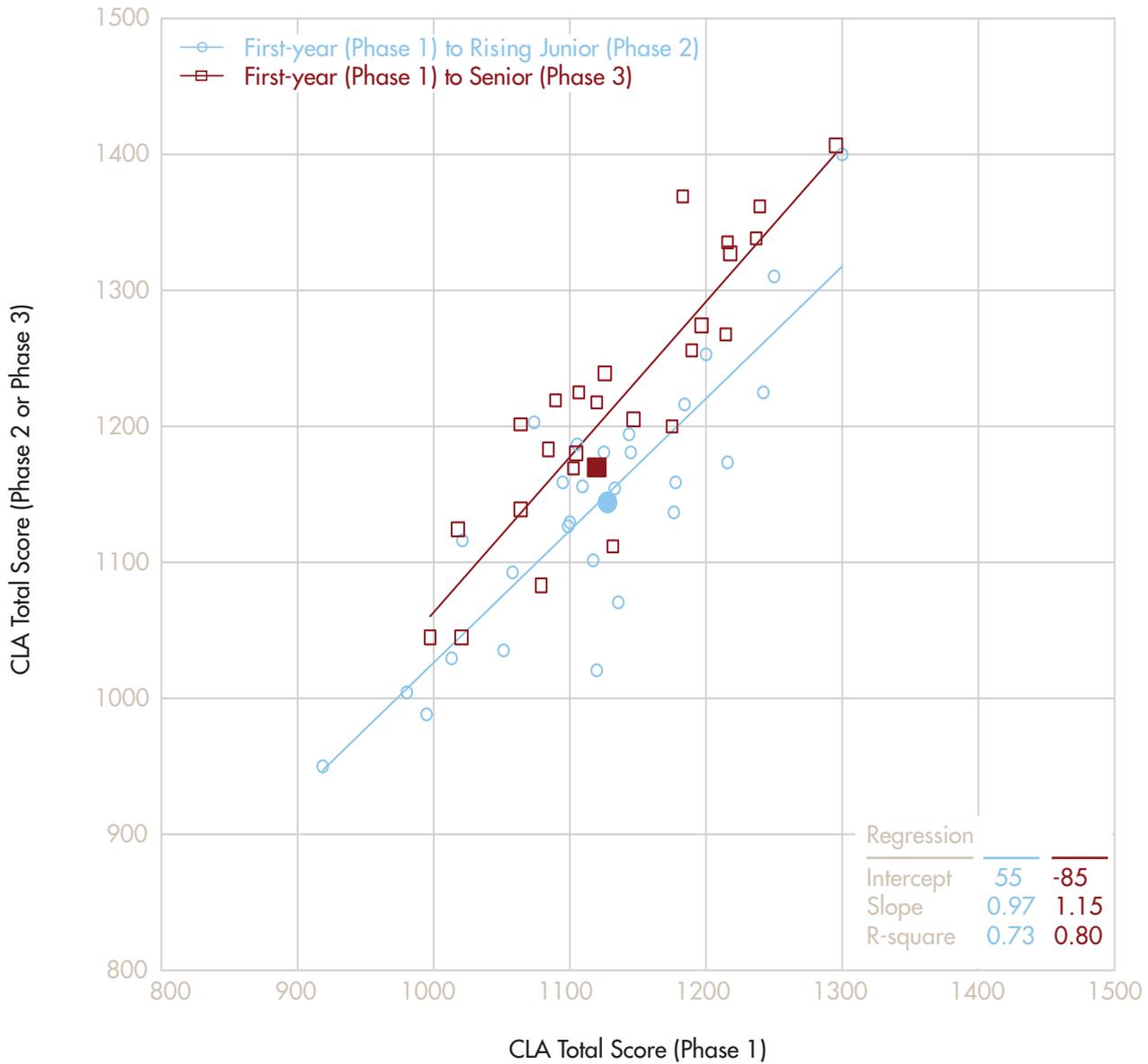
The vertical y-axis in Figure B above shows a school's mean CLA Total score for both Phase 2 (in blue) and Phase 3 (in red).

The horizontal x-axis shows a school's mean CLA Total score for Phase 1.

A school's data point is above its corresponding line if its students in Phase 2 (or Phase 3) performed better on the CLA than what would be expected relative to their Phase 1 scores.

During the spring 2009 testing cycle, 31 institutions tested a sufficient number of students to provide the school-level analyses and results presented in this report.

Table 5 shows CLA longitudinal schools grouped by Basic Carnegie Classification. The spread of schools differs slightly from that of the 1,713 four-year institutions across the nation, with Doctorate-granting Universities constituting a higher percentage among CLA schools than nationally. Accordingly, representation among both Master's Colleges and Universities as well as Baccalaureate Colleges is lower among CLA longitudinal schools than it is nationally.

5

Carnegie Classification	Nation		CLA	
	Number	Percentage	Number	Percentage
Doctorate-granting Universities	282	16%	12	39%
Master's Colleges and Universities	664	39%	9	29%
Baccalaureate Colleges	767	45%	10	32%
	1,713		31	

Source: Carnegie Foundation for the Advancement of Teaching, Carnegie Classifications Data File, June 11, 2008.

3: Longitudinal Cohort

Table 6 provides comparative statistics for colleges and universities across the nation and CLA longitudinal schools in terms of some important characteristics. These statistics suggest that CLA longitudinal schools are fairly representative of institutions nationally. Percentage public, percentage HBCU, and undergraduate student body size are exceptions.

6

School Characteristic	Nation	CLA
Percentage public	37%	52%
Percentage Historically Black College or University (HBCU)	5%	13%
Mean percentage of undergraduates receiving Pell grants	34%	31%
Mean four-year graduation rate	36%	40%
Mean six-year graduation rate	52%	61%
Mean first-year retention rate	73%	82%
Mean Barron's selectivity rating	3.4	3.8
Mean estimated median SAT score	1067	1106
Mean number of FTE undergraduate students (rounded)	4,320	9,040
Mean student-related expenditures per FTE student (rounded)	$12,365	$14,180

Source: College Results Online dataset, managed by and obtained with permission from the Education Trust, covers most 4-year Title IV-eligible higher-education institutions in the United States. Data were constructed from IPEDS and other sources. Because all schools did not report on every measure in the table, the averages and percentages may be based on slightly different denominators.

3: Longitudinal Cohort

For each longitudinal phase, Table 7 below presents summary statistics for all students including counts, means, 25th and 75th percentiles, and standard deviations.

7

Phase 1 (student level)	Student Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	11437	971	1087	1209	189
Analytic Writing Task	9221	977	1085	1196	165
Make-an-Argument	9879	942	1080	1225	188
Critique-an-Agument	9627	869	1079	1167	186
Total Score	9168	990	1096	1198	149
EAA Score	11360	950	1093	1230	195

Phase 2 (student level)	Student Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	3327	1012	1157	1277	207
Analytic Writing Task	3161	980	1119	1267	161
Make-an-Argument	3227	942	1114	1225	181
Critique-an-Agument	3202	1018	1120	1316	180
Total Score	3141	1025	1141	1249	160
EAA Score	3329	1010	1134	1270	188

Phase 3 (student level)	Student Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	2374	1093	1216	1337	194
Analytic Writing Task	2308	1091	1221	1363	186
Make-an-Argument	2318	1051	1199	1354	214
Critique-an-Agument	2317	1083	1240	1393	210
Total Score	2290	1115	1222	1339	163
EAA Score	2380	1030	1153	1300	184

3: Longitudinal Cohort

For each longitudinal phase, Table 8 below presents summary statistics for all schools including counts, means, 25th and 75th percentiles, and standard deviations.

8

Phase 1 (school level)	School Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	48	1027	1094	1168	104
Analytic Writing Task	47	1021	1089	1164	98
Make-an-Argument	48	1018	1081	1159	103
Critique-an-Agument	48	1006	1084	1153	103
Total Score	47	1038	1100	1170	93
EAA Score	48	1011	1104	1209	152

Phase 2 (school level)	School Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	32	1068	1147	1194	103
Analytic Writing Task	32	1057	1114	1170	93
Make-an-Argument	32	1062	1111	1166	94
Critique-an-Agument	32	1051	1114	1160	94
Total Score	32	1073	1132	1185	96
EAA Score	32	1044	1116	1193	129

Phase 3 (school level)	School Count	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Performance Task	31	1126	1193	1256	107
Analytic Writing Task	30	1120	1201	1267	109
Make-an-Argument	30	1095	1181	1251	109
Critique-an-Agument	30	1152	1219	1281	111
Total Score	30	1138	1202	1267	104
EAA Score	32	1049	1126	1217	137

3: Longitudinal Cohort

The institutions listed here in alphabetical order agreed to be identified as participating schools. All participated in Phase 3 of the longitudinal assessment, but some may not have tested enough students to be included in all comparative analyses.

Auburn University
Belmont University
Bowling Green State University
California State Polytechnic University-Pomona
California State University-Northridge
Carleton College
Central Michigan University
City University of New York City College
City University of New York Herbert H. Lehman College
Cleveland State University
Colorado College
Earlham College
Fayetteville State University
Grand Valley State University
Kalamazoo College
Loyola University of Chicago
Macalester College
Morgan State University
North Carolina Agricultural and Technical State University
Northern Arizona University
Pace University
Saint Xavier University
Spelman College
St. Olaf College
Syracuse University
The George Washington University
The Ohio State University
University of California, Riverside
University of Charleston
University of North Carolina at Charlotte
University of Saint Thomas
University of the Pacific
University of Wyoming
Wagner College
Winston-Salem State University
Wofford College

4: Diagnostic Guidance

CLA results operate as a signaling tool of overall institutional performance on tasks that measure higher order skills holistically. However, the three types of CLA tasks—Performance, Make-an-Argument, and Critique-an-Argument—differ slightly in the combination of skills necessary to perform well.

Indeed, some schools score significantly lower on one type than on another. Examining performance across CLA task types can serve as an initial diagnostic exercise. Specifically, cases of performance Well Below or Below Expected on a particular task type indicate that students are not demonstrating the expected level of skill at analyzing complex, realistic scenarios; writing a persuasive, analytic essay to support a position on an issue; and/or critiquing written arguments.

Performance Task

Analyzing complex, realistic scenarios

Synthesizing information from multiple sources; recognizing conflicting evidence, weighing the credibility of different sources of evidence; identifying logical fallacies, interpreting data, tables, and figures correctly; drawing reasonable and logical inferences from the available information; developing sound conclusions based on all available evidence; and utilizing the most relevant and credible evidence available to justify their conclusion.

Make-an-Argument

Writing a persuasive, analytic essay

Establishing a thesis or a position on an issue; maintaining the thesis throughout the essay; supporting the thesis with relevant and persuasive examples (e.g., from personal experience, history, art, literature, pop culture, or current events); anticipating and countering opposing arguments to the position, fully developing ideas, examples, and arguments; crafting an overall response that generates interest, provokes thought, and persuades the reader; organizing the structure of the essay (e.g., paragraphing, ordering of ideas and sentences within paragraphs); employing transitions and varied sentence structure to maintain the flow of the argument; and utilizing sophisticated grammar and vocabulary.

Critique-an-Argument

Critiquing written arguments

Identifying a variety of logical flaws or fallacies in a specific argument; explaining how or why the logical flaws affect the conclusions in that argument; and presenting their critique in a written response that is a grammatically correct, organized, well-developed, logically sound, and neutral in tone.

5: Moving Forward

We encourage institutions to examine performance across CLA tasks and communicate results across campus, link student-level CLA results with other data sources, pursue in-depth sampling, stay informed through the CLA Spotlight, and participate in CLA Education offerings.

Student-level CLA results are provided for you to link with other data sources (e.g., course-taking patterns, grades, portfolios, student satisfaction and engagement, major-specific tests, etc.).

These internal analyses can help you generate hypotheses for additional research, which you can pursue through CLA in-depth sampling in experimental areas (e.g., programs or colleges within your campus) in subsequent years or simultaneously.

We welcome and encourage your participation in the CLA Spotlight—a series of free informational web conferences. Each CLA Spotlight features campuses doing promising work using the CLA, guest-speakers from the larger world of assessment, and/or CLA staff members who provide updates or insights to CLA-related programs and projects.

CLA Education focuses on curriculum and pedagogy and embraces the crucial role that faculty play in the process of assessment.

The flagship program of CLA Education is the Performance Task Academy, which shifts the focus from general assessment to the course-level work of faculty. The Performance Task Academy provides an opportunity for faculty members to learn to diagnose their individual students' work and to receive guidance in creating their own performance tasks, which are designed to supplement the educational reform movement toward a case and problem approach in learning and teaching.

A CLA Education web site also has been formed as a clearing house for performance tasks developed by faculty. For more information, visit www.claintheclassroom.org, or contact Director of CLA Education, Dr. Marc Chun at mchun@cae.org.

Through the steps noted above we encourage institutions to move toward a continuous system of improvement in teaching and learning stimulated by the CLA. Without your contributions, the CLA would not be on the exciting path that it is today. We look forward to your continued involvement!

Appendices

Introduction

The CLA is comprised of three types of prompts within two types of task: the Performance Task and the Analytic Writing Task. Your students were asked to take all three prompts. The Analytic Writing Task includes a pair of prompts called Make-an-Argument and Critique-an-Argument.

The CLA uses direct measures of skills in which students perform cognitively demanding tasks from which quality of response is scored. All CLA measures are administered online and contain open-ended prompts that require constructed responses. There are no multiple-choice questions.

The CLA tasks require students to integrate critical thinking, analytic reasoning, problem solving, and written communication skills. The holistic integration of these skills on the CLA tasks mirrors the requirements of serious thinking and writing tasks faced in life outside of the classroom.

Performance Task

Each Performance Task requires students to use an integrated set of critical thinking, analytic reasoning, problem solving, and written communication skills to answer several open-ended questions about a hypothetical but realistic situation. In addition to directions and questions, each Performance Task also has its own document library that includes a range of information sources, such as letters, memos, summaries of research reports, newspaper articles, maps, photographs, diagrams, tables, charts, and interview notes or transcripts. Students are instructed to use these materials in preparing their answers to the Performance Task's questions within the allotted 90 minutes.

The first portion of each Performance Task contains general instructions and introductory material. The student is then presented with a split screen. On the right side of the screen is a list of the materials in the Document Library. The student selects a particular document to view by using a pull-down menu. On the left side of the screen are a question and a response box. There is no limit on how much a student can type. Upon completing a question, students then select the next question in the queue.

No two Performance Tasks assess the exact same combination of skills. Some ask students to identify and then compare and contrast the strengths and limitations of alternative hypotheses, points of view, courses of action, etc. To perform these and other tasks, students may have to weigh different types of evidence, evaluate the credibility of various documents, spot possible bias, and identify questionable or critical assumptions.

Performance Tasks also may ask students to suggest or select a course of action to resolve conflicting or competing strategies and then provide a rationale for that decision, including why it is likely to be better than one or more other approaches. For example, students may be asked to anticipate potential difficulties or hazards that are associated with different ways of dealing with a problem, including the likely short- and long-term consequences and implications of these strategies. Students may then be asked to suggest and defend one or more of these approaches. Alternatively, students may be asked to review a collection of materials or a set of options, analyze and organize them on multiple dimensions, and then defend that organization.

Performance Tasks often require students to marshal evidence from different sources; distinguish rational from emotional arguments and fact from opinion; understand data in tables and figures; deal with inadequate, ambiguous, and/or conflicting information; spot deception and holes in the arguments made by others; recognize information that is and is not relevant to the task at hand; identify additional information that would help to resolve issues; and weigh, organize, and synthesize information from several sources.

Analytic Writing Task

Students write answers to two types of essay prompts, namely: a "Make-an-Argument" question that asks them to support or reject a position on some issue; and a "Critique-an-Argument" question that asks them to evaluate the validity of an argument made by someone else. Both of these tasks measure a student's skill in articulating complex ideas, examining claims and evidence, supporting ideas with relevant reasons and examples, sustaining a coherent discussion, and using standard written English.

Make-an-Argument

A "Make-an-Argument" prompt typically presents an opinion on some issue and asks students to write, in 45 minutes, a persuasive, analytic essay to support a position on the issue. Key elements include: establishing a thesis or a position on an issue; maintaining the thesis throughout the essay; supporting the thesis with relevant and persuasive examples (e.g., from personal experience, history, art, literature, pop culture, or current events); anticipating and countering opposing arguments to the position, fully developing ideas, examples, and arguments; crafting an overall response that generates interest, provokes thought, and persuades the reader; organizing the structure of the essay (e.g., paragraphing, ordering of ideas and sentences within paragraphs); employing transitions and varied sentence structure to maintain the flow of the argument; and utilizing sophisticated grammar and vocabulary.

Critique-an-Argument

A "Critique-an-Argument" prompt asks students, in 30 minutes, to critique an argument by discussing how well reasoned they find it to be (rather than simply agreeing or disagreeing with the position presented). Key elements of the essay include: identifying a variety of logical flaws or fallacies in a specific argument; explaining how or why the logical flaws affect the conclusions in that argument; and presenting a critique in a written response that is a grammatically correct, organized, well-developed, logically sound, and neutral in tone.

A: Task Overview

Example Performance Task

You advise 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. Your document library contains the following materials:

Example Document Library

- Newspaper article about the accident

- Federal Accident Report on in-flight breakups in single-engine planes

- Internal Correspondence (Pat's e-mail to you and Sally's e-mail to Pat)

- Charts relating to SwiftAir's performance characteristics

- Excerpt from magazine article comparing SwiftAir 235 to similar planes

- Pictures and descriptions of SwiftAir Models 180 and 235

Example Questions

- Do the available data tend to support or refute the claim that the type of wing on the SwiftAir 235 leads to more in-flight breakups?

- What is the basis for your conclusion?

- What other factors might have contributed to the accident and should be taken into account?

- What is your preliminary recommendation about whether or not DynaTech should buy the plane and what is the basis for this recommendation?

Example Make-an-Argument

There is no such thing as "truth" in the media. The one true thing about the information media is that it exists only to entertain.

Example Critique-an-Argument

A well-respected professional journal with a readership that includes elementary school principals recently published the results of a two-year study on childhood obesity. (Obese individuals are usually considered to be those who are 20 percent above their recommended weight for height and age.) This study sampled 50 schoolchildren, ages 5-11, from Smith Elementary School. A fast food restaurant opened near the school just before the study began. After two years, students who remained in the sample group were more likely to be overweight––relative to the national average. Based on this study, the principal of Jones Elementary School decided to confront her school's obesity problem by opposing any fast food restaurant openings near her school.

B: Task Development

Iterative Development Process

A team of researchers and writers generate ideas for Make-an-Argument and Critique-an-Argument prompts, and Performance Task storylines, and then contribute to the development and revision of the prompts and Performance Task documents.

For Analytic Writing Tasks, multiple prompts are generated, revised and pre-piloted, and those prompts that elicit good critical thinking and writing responses during pre-piloting are further revised and submitted to more extensive piloting.

During the development of Performance Tasks, care is taken to ensure that sufficient information is provided to permit multiple reasonable solutions to the issues present in the Performance Task. Documents are crafted such that information is presented in multiple formats (e.g., tables, figures, news articles, editorials, letters, etc.).

While developing a Performance Task, a list of the intended content from each document is established and revised. This list is used to ensure that each piece of information is clearly reflected in the document and/or across documents, and to ensure that no additional pieces of information are embedded in the document that were not intended. This list serves as a draft starting point for the analytic scoring items used in the Performance Task scoring rubrics.

During revision, information is either added to documents or removed from documents to ensure that students could arrive at approximately three or four different conclusions based on a variety of evidence to back up each conclusion. Typically, some conclusions are designed to be supported better than others.

Questions for the Performance Task are also drafted and revised during the development of the documents. The questions are designed such that the initial questions prompt the student to read and attend to multiple sources of information in the documents, and later questions require the student to evaluate the documents and then use their analysis to draw conclusions and justify those conclusions.

After several rounds of revision, the most promising of the Performance Tasks and the Make-an-Argument and Critique-an-Argument prompts are selected for pre-piloting. Student responses from the pilot test are examined to identify what pieces of information are unintentionally ambiguous, what pieces of information in the documents should be removed, etc. After revision and additional pre-piloting, the best functioning tasks (i.e., those that elicit the intended types and ranges of student responses) are selected for full piloting.

During piloting, students complete both an operational task and one of the new tasks. At this point, draft scoring rubrics are revised and tested in grading the pilot responses, and final revisions are made to the tasks to ensure that the task is eliciting the types of responses intended.

C: Scoring Criteria

Introduction

This section summarizes the types of questions addressed by CLA scoring of all task types. Because each CLA task and their scoring rubrics differ, not every item listed is applicable to every task. The tasks cover different aspects of critical thinking, analytic reasoning, problem solving, and writing and in doing so can, in combination, better assess the entire domain of performance.

Assessing Critical Thinking, Analytic Reasoning and Problem Solving

Applied in combination, critical thinking, analytic reasoning and problem solving skills are required to perform well on CLA tasks. We define these skills as how well students can evaluate and analyze source information, and subsequently to draw conclusions and present an argument based upon that analysis. In scoring, we specifically consider the following items to be important aspects of these skills.

Assessing Writing

Analytic writing skills invariably depend on clarity of thought. Therefore, analytic writing and critical thinking, analytic reasoning, and problem solving are related skills sets. The CLA measures critical thinking performance by asking students to explain in writing their rationale for various conclusions. In doing so, their performance is dependent on both writing and critical thinking as integrated rather than separate skills. We evaluate writing performance using holistic scores that consider several aspects of writing depending on the task. The following are illustrations of the types of questions we address in scoring writing on the various tasks.

C: Scoring Criteria

Assessing Critical Thinking,
Analytic Reasoning and
Problem Solving

Evaluation of evidence

How well does the student assess the quality and relevance of evidence, including:

- Determining what information is or is not pertinent to the task at hand

- Distinguishing between rational claims and emotional ones, fact from opinion

- Recognizing the ways in which the evidence might be limited or compromised

- Spotting deception and holes in the arguments of others

- Considering all sources of evidence

Analysis and synthesis of evidence

How well does the student analyze and synthesize data and information, including:

- Presenting his/her own analysis of the data or information (rather than "as is")

- Committing or failing to recognize logical flaws (e.g., distinguishing correlation from causation)

- Breaking down the evidence into its component parts;

- Drawing connections between discrete sources of data and information

- Attending to contradictory, inadequate or ambiguous information

Drawing conclusions

How well does the student form a conclusion from their analysis, including:

- Constructing cogent arguments rooted in data/ information rather than speculation/opinion

- Selecting the strongest set of supporting data

- Prioritizing components of the argument

- Avoiding overstated or understated conclusions

- Identifying holes in the evidence and subsequently suggesting additional information that might resolve the issue

Acknowledging alternative explanations/viewpoints

How well does the student acknowledge additional perspectives and consider other options, including:

- Recognizing that the problem is complex with no clear answer

- Proposing other options and weighing them in the decision

- Considering all stakeholders or affected parties in suggesting a course of action

- Qualifying responses and acknowledging the need for additional information in making an absolute determination

C: Scoring Criteria

Assessing Writing

Interest

How well does the student maintain the reader's interest? Does the...

- Student use creative and engaging examples or descriptions
- Structure, syntax and organization add to the interest of their writing
- Student use colorful but relevant metaphors, similes, etc.
- Writing engage the reader
- Writing leave the reader thinking

Presentation

How clear and concise is the argument? Does the student...

- Clearly articulate the argument and the context for that argument
- Correctly and precisely use evidence to defend the argument
- Comprehensibly and coherently present evidence

Development

How effective is the structure? Does the student...

- Logically and cohesively organize the argument
- Avoid extraneous elements in the argument's development
- Present evidence in an order that contributes to a persuasive and coherent argument

Persuasiveness

How well does the student defend the argument? Does the student...

- Effectively present evidence in support of the argument
- Draw thoroughly and extensively from the available range of evidence
- Analyze the evidence in addition to simply presenting it
- Consider counterarguments and address weaknesses in his/her own argument

Mechanics

What is the quality of the student's writing?

- Is vocabulary and punctuation used correctly
- Is the student's understanding of grammar strong
- Is the sentence structure basic, or more complex and creative
- Does the student use proper transitions
- Are the paragraphs structured logically and effectively

D: Scoring Process

Score Sheet

There are two types of items that appear on a CLA score sheet: analytic and holistic. Analytic scoring items are particular to each prompt and holistic items refer to general dimensions, such as evaluation of evidence, drawing conclusions, acknowledging alternative explanations and viewpoints, and overall writing. We compute raw scores for each task by adding up all points on all items (i.e., calculating a unit-weighted sum).

Performance Task scoring is tailored to each specific prompt and includes a combination of both holistic and analytic scoring items. Though there are many types of analytic items on the Performance Task score sheets, the most common represent a list of the possible pieces of information a student could or should raise in their response. These cover the information presented in the Performance Task documents as well as information that can be deduced from comparing information across documents. The analytic items are generally given a score of 0 if the student did not use the information in their response, or 1 if they did. The number of analytic items varies by prompt.

Performance Task holistic items are scored on four or seven-point scales (i.e., 1-4 or 1-7). There are multiple holistic items per Performance Task that require graders to provide an evaluation of different aspects of critical thinking and reasoning in the student responses. These holistic items include areas such as the student's use of the most relevant information in the Performance Task, their recognition of strengths and weaknesses of various pieces of information, overall critical thinking, and overall writing.

Critique-an-Argument score sheets also include a combination of analytic and holistic scores. Critique-an-Argument analytic items are a list of possible critiques of the argument presented in the prompt. In addition, a few holistic items are used to rate the overall quality, critical thinking and writing over the entire response.

Make-an-Argument score sheets contain only holistic items scored on four or seven-point scales (i.e., 1-4 or 1-7). The holistic items include ratings for various aspects of writing (e.g., organization, mechanics, etc.) and critical thinking (e.g., reasoning and logic, sophistication and depth of treatment of the issues raised in the prompt) as well as two overall assessments of writing and critical thinking.

For all task types, blank responses or responses that are entirely unrelated to the task (e.g., writing about what they had for breakfast) are assigned a 0 and are flagged for removal from the school-level results.

D: Scoring Process

Scoring Procedure

All scorer candidates undergo rigorous training in order to become certified CLA scorers. Training includes an orientation to the prompt and score sheet, instruction on how to evaluate the scoring items, repeated practice grading a wide range of student responses, and extensive feedback and discussion after scoring each response.

After participating in training, scorers complete a reliability check where they score the same set of student responses. Scorers with low agreement or reliability (determined by comparisons of raw score means, standard deviations and correlations among the scorers) are either further coached or removed from scoring.

In fall 2008 and spring 2009, a combination of machine and human scoring was used for the Analytic Writing Task.

The CLA utilizes Pearson Knowledge Technology's Intelligent Essay Assessor program for evaluating responses to the Make-an-Argument and Critique-an-Argument prompts.

The machine scoring engine was developed and tested using scores from a broad range of responses that were previously scored by humans (often double scored). In some cases the automated scoring engine is unable to score off-topic or abnormally short/long responses. These student responses are scored by humans.

To facilitate reporting results across schools, ACT scores were converted (using the ACT-SAT crosswalk to the right) to the scale of measurement used to report SAT scores.

For institutions where a majority of students did not have ACT or SAT scores (e.g., two-year institutions and open admission schools), we make available the Scholastic Level Exam (SLE), a short-form cognitive ability measure, as part of the CLA. The SLE is produced by Wonderlic, Inc. SLE scores were converted to SAT scores using data from 1,148 students participating in spring 2006 that had both SAT and SLE scores. These converted scores (both ACT to SAT and SLE to SAT) are referred to simply as entering academic ability (EAA) scores.

Students receive a single score on a CLA task because each task assesses an integrated set of critical thinking, analytic reasoning, problem solving, and written communication skills.

Standard ACT to SAT Conversion Table

ACT	to	SAT
36		1600
35		1580
34		1520
33		1470
32		1420
31		1380
30		1340
29		1300
28		1260
27		1220
26		1180
25		1140
24		1110
23		1070
22		1030
21		990
20		950
19		910
18		870
17		830
16		780
15		740
14		680
13		620
12		560
11		500

Sources:

"Concordance Between ACT Assessment and Recentered SAT I Sum Scores" by N.J. Dorans, C.F. Lyu, M. Pommerich, and W.M. Houston (1997), College and University, 73, 24-31; "Concordance between SAT I and ACT Scores for Individual Students" by D. Schneider and N.J. Dorans, Research Notes (RN-07), College Entrance Examination Board: 1999; "Correspondences between ACT and SAT I Scores" by N.J. Dorans, College Board Research Report 99-1, College Entrance Examination Board: 1999; ETS Research Report 99-2, Educational Testing Service: 1999.

E: Scaling Procedures

Each Performance Task and Analytic Writing Task has a unique scoring rubric, and the maximum number of reader assigned raw score points differs across tasks. Consequently, a given reader-assigned raw score, such as 15 points, may be a relatively high score on one task but a low score on another task.

To adjust for such differences, reader-assigned raw scores on the different tasks are converted to a common scale of measurement. This process results in scale scores that reflect comparable levels of proficiency across tasks. For example, a given CLA scale score indicates approximately the same percentile rank regardless of the task on which it was earned. This feature of the CLA scale scores allows combining scores from different tasks to compute a school's mean scale score for each task type as well as a total average scale score across types.

A linear scale transformation is used to convert reader-assigned raw scores to scale scores. This process results in a scale score distribution with the same mean and standard deviation as the Entering Academic Ability (EAA) scores of the freshmen who took that measure. This type of scaling preserves the shape of the raw score distribution and maintains the relative standing of students. For example, the student with the highest raw score on a task will also have the highest scale score on that task, the student with the next highest raw score will be assigned the next highest scale score, and so on.

This type of scaling generally results in the highest raw score earned on a task receiving a scale score of approximately the same value as the maximum EAA score of any freshman who took that task. Similarly, the lowest raw score earned on a task would be assigned a scale score value that is approximately the same as the lowest EAA score of any freshman who took that task. On very rare occasions, a student may achieve an exceptionally high or low raw score (i.e., well above or below the other students taking that task). When this occurs, it results in assigning a student a scale score that is outside of the normal EAA range. Prior to the spring of 2007, scores were capped at 1600. Capping was discontinued starting in fall 2007.

In the past, CAE revised its scaling equations each fall. However, many institutions would like to make year-to-year comparisons (i.e., as opposed to just fall to spring). To facilitate this activity, in fall 2007 CAE began using the same scaling equations it developed for the fall 2006 administration and has done so for new tasks introduced since then. As a result of this policy, a given raw score on a task will receive the same scale score regardless of when the student took the task.

We provide a CLA Student Data File, which includes variables across three categories: student self-reported information from their CLA on-line profile; CLA scores and identifiers; and information provided/verified by the registrar.

We provide student-level information to link with other data you collect (e.g., from NSSE, CIRP, portfolios, local assessments, course-taking patterns, participation in specialized programs, etc.) to help you hypothesize about campus-specific factors related to overall institutional performance.

Student-level scores are not designed to be diagnostic at the individual level and should be considered as only one piece of evidence about a student's skills.

Self-Reported Data

- Student ID, E-mail address, and Name (first, middle initial, last)
- Age
- Gender
- Race/Ethnicity
- Primary and Secondary Academic Major (34 categories)
- Field of Study (6 categories; based on primary academic major)
- English as primary language
- Total years at school
- Attended school as Freshman, Sophomore, Junior, Senior
- CLA Local Survey Responses

CLA Scores and Identifiers

- CLA scores for Performance Task, Analytic Writing Task, Make-an-Argument, Critique-an-Argument, and Total CLA Score (depending on the number of tasks taken and completeness of responses):

  - CLA scale scores;
  - Student Performance Level categories (i.e., well below expected, below expected, at expected, above expected, well above expected) if CLA scale score and entering academic ability (EAA) scores are available;
  - Percentile Rank in the CLA (among students in the same class year; based on scale score); and
  - Percentile Rank at School (among students in the same class year; based on scale score).

- Unique CLA numeric identifiers
- Year, Administration (Fall or Spring), Type of Test (90 or 180-minute), Date of test

Registrar Data

- Class Standing
- Transfer Student Status
- Program ID and Name (for classification of students into different schools, majors, studies, programs, etc.)
- Entering Academic Ability (EAA) Score
- SAT I - Math
- SAT I - Verbal / Critical Reading
- SAT Total (Math + Verbal)
- SAT I - Writing
- SAT I - Writing (Essay sub-score)
- SAT I - Writing (Multiple-choice subscore)
- ACT - Composite
- ACT - English
- ACT - Reading
- ACT - Mathematics
- ACT - Science
- ACT - Writing
- GPA

council for aid to education