8-15-2011

CLA 2007-2011 Phase 3 Longitudinal Report

UNO Office of Institutional Effectiveness
University of Nebraska at Omaha

CLA

The University of Nebraska at Omaha

CLA 2007-2011 PHASE 3 LONGITUDINAL REPORT

Your 2007-2011 Phase 3 Longitudinal Results consist of two components:

- CLA Institutional Report and Appendices
- CLA Student Data File

Report

The report introduces readers to the CLA, presents your results, and offers guidance on interpretation and next steps.

Appendices

Appendices offer more detail on CLA tasks, scoring and scaling, value-added equations, and the Student Data File.

Student Data File

Your Longitudinal Student Data File was distributed separately as a password-protected Excel file. Your Student Data File may be used to link with other data sources and to generate hypotheses for additional research.

The Collegiate Learning Assessment (CLA) is a major initiative of the Council for Aid to Education. The CLA offers a value-added, constructed-response approach to the assessment of higher-order skills, such as critical thinking and written communication. Hundreds of institutions and hundreds of thousands of students have participated in the CLA to date.

The institution—not the student—is the primary unit of analysis. The CLA is designed to measure an institution's contribution, or value added, to the development of higher-order skills. This approach allows an institution to compare its student learning results on the CLA with learning results at similarly selective institutions.

The CLA is intended to assist faculty, school administrators, and others interested in programmatic change to improve teaching and learning, particularly with respect to strengthening higher-order skills. Included in the CLA are Performance Tasks and Analytic Writing Tasks. Performance Tasks present realistic problems that require students to analyze complex materials. Several different types of materials are used that vary in credibility, relevance to the task, and other characteristics. Students' written responses to the tasks are graded to assess their abilities to think critically, reason analytically, solve problems, and write clearly and persuasively.

The CLA helps campuses follow a continuous improvement model that positions faculty as central actors in the link between assessment and teaching/learning.

The continuous improvement model requires multiple indicators beyond the CLA because no single test can serve as the benchmark for all student learning in higher education. There are, however, certain skills judged to be important by most faculty and administrators across virtually all institutions; indeed, the higher-order skills the CLA focuses on fall into this category.

The signaling quality of the CLA is important because institutions need to 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. Yet, the CLA is not about ranking institutions. Rather, it is about highlighting differences between them that can lead to improvements. The CLA is an instrument designed to contribute directly to the improvement of teaching and learning. In this respect it is in a league of its own.

Sample Sizes for Phase Comparisons

Table 2.1 shows, for each pair of phases, the number of students that have CLA scores for both phases. These numbers of students serve as the baseline for any comparison between a pair of phases. Due to changes in sample restrictions, the sample used for calculations in this report may differ slightly from the sample used in your Phase 2 longitudinal report. In this and all tables, take note of the sample sizes and exercise caution in interpreting the results if the sample size is small.

2.1 Number of Students with CLA Scores

	Phase 1 and Phase 2	Phase 2 and Phase 3	Phase 1 and Phase 3
Total CLA Score*	N/A	N/A	47
Performance Task	N/A	N/A	50
Analytic Writing Task	N/A	N/A	47
Make-an-Argument	N/A	N/A	48
Critique-an-Argument	N/A	N/A	47

* *Total Score* refers to the sample of students with both Performance Task and Analytic Writing Task scores.

2.2 Summary Statistics for Your Students Participating in the Longitudinal Assessment

Phase 1 (Fall 2007)	Number of Students*	25th Percentile Score	Mean Score	75th Percentile Score	Standard Deviation
Total CLA Score**	212	963	1034	1099	108
Performance Task	225	971	1064	1133	135
Analytic Writing Task	212	914	1002	1077	132
Make-an-Argument	214	851	971	1074	160
Critique-an-Argument	213	903	1030	1118	164
EAA***	225	1030	1155	1300	172

Phase 2 (Spring 2009)	Number of Students*	25th Percentile Score	Mean Score	75th Percentile Score	Standard Deviation
Total CLA Score**	N/A	N/A	N/A	N/A	N/A
Performance Task	N/A	N/A	N/A	N/A	N/A
Analytic Writing Task	N/A	N/A	N/A	N/A	N/A
Make-an-Argument	N/A	N/A	N/A	N/A	N/A
Critique-an-Argument	N/A	N/A	N/A	N/A	N/A
EAA***	N/A	N/A	N/A	N/A	N/A

Phase 3 (Spring 2011)	Number of Students*	25th Percentile Score	Mean Score	75th Percentile Score	Standard Deviation
Total CLA Score**	49	1116	1201	1293	127
Performance Task	50	1157	1249	1338	147
Analytic Writing Task	49	1057	1150	1247	140
Make-an-Argument	49	989	1118	1220	177
Critique-an-Argument	49	1091	1181	1293	148
EAA***	50	1140	1230	1380	177

* In this and all tables, take note of the sample sizes and exercise caution in interpreting the results if the sample size is small.

** *Total Score* refers to the sample of students with both Performance Task and Analytic Writing Task scores.

*** SAT Math + Verbal, ACT Composite, or Scholastic Level Exam (SLE) scores on the SAT scale. Hereinafter referred to as Entering Academic Ability (EAA).

Results at Your School Across Phases

Table 2.3 provides results for students at your institution that participated in Phases 1, 2, and 3. Effect sizes are the standardized difference in mean performance between a pair of phases. We also present the correlation between total scores of students at each phase; a correlation of 0 indicates no linear relationship between scores in the earlier phase and scores in the later phase, and a correlation of -1 or +1 indicates a perfect linear relationship between the scores in each phase.

2.3 Comparison of Your Students' Scores Across Phases

● ● ○	Phase 1			Phase 2		Difference*		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score Correlation	Effect Size
Total CLA Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Performance Task	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Analytic Writing Task	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Make-an-Argument	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Critique-an-Argument	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

○ ● ●	Phase 2			Phase 3		Difference*		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score Correlation	Effect Size
Total CLA Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Performance Task	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Analytic Writing Task	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Make-an-Argument	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Critique-an-Argument	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

● ○ ●	Phase 1			Phase 3		Difference*		Summary Statistics	
	Number of Students	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean Score Correlation	Effect Size
Total CLA Score	47	1044	113	1202	129	158	16	0.51	1.4
Performance Task	50	1069	126	1249	147	180	21	0.28	1.43
Analytic Writing Task	47	1015	152	1149	138	134	-14	0.42	0.88
Make-an-Argument	48	1001	171	1116	178	115	7	0.39	0.67
Critique-an-Argument	47	1029	182	1178	144	149	-38	0.25	0.82

* The Mean Score and Standard Deviation differences simply reflect the former phase values subtracted from the latter phase values.

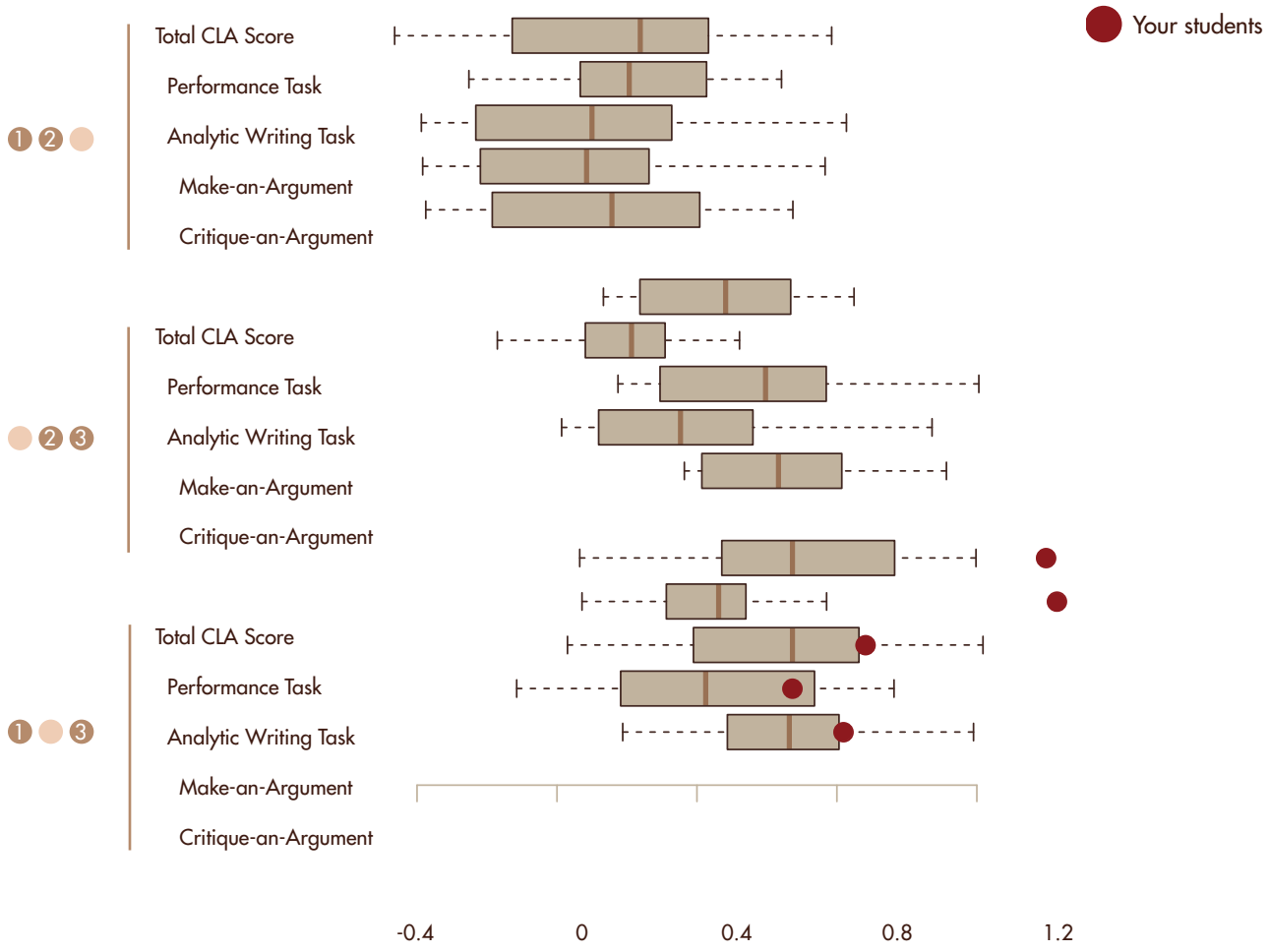
2.4 Effect Sizes and Performance Levels for Your Institution

	Phase 1 and Phase 2		Phase 2 and Phase 3		Phase 1 and Phase 3	
	Effect Size	Performance Level	Effect Size	Performance Level	Effect Size	Performance Level
Total CLA Score	N/A	N/A	N/A	N/A	1.4	Well Above
Performance Task	N/A	N/A	N/A	N/A	1.43	Well Above
Analytic Writing Task	N/A	N/A	N/A	N/A	0.88	Above
Make-an-Argument	N/A	N/A	N/A	N/A	0.67	Above
Critique-an-Argument	N/A	N/A	N/A	N/A	0.82	Above

The performance levels in Table 2.4 above indicate whether the change in your students' CLA performance was well above, above, near, below, or well below what would be expected given the distribution of effect sizes observed in the 2005–2009 CLA Lumina Longitudinal study, which serves as the reference group for this report. Specific percentile ranks for effect sizes demarcate performance level categories as shown in the table at right. The next page contains additional information about the distribution of effect sizes.

Percentile	Performance Level
90-99th	Well Above Expected
70-89th	Above Expected
30-69th	Near Expected
10-29th	Below Expected
0-9th	Well Below Expected

2.5 CLA Score Changes Between Phases (Effect Sizes)



Interpreting Score Changes

The "box and whisker" plots in Figure 2.5 show effect size distributions across CLA measures for longitudinal schools. (The next section describes the longitudinal reference cohort in detail.) The top cluster displays effect sizes between Phase 1 (first-year students) and Phase 2 (rising juniors); the middle cluster repeats this for Phases 2 and 3 (seniors); and the bottom cluster does the same for Phases 1 and 3. Your institution's effect sizes are shown in red; the numeric effect sizes are shown in Table 2.4.

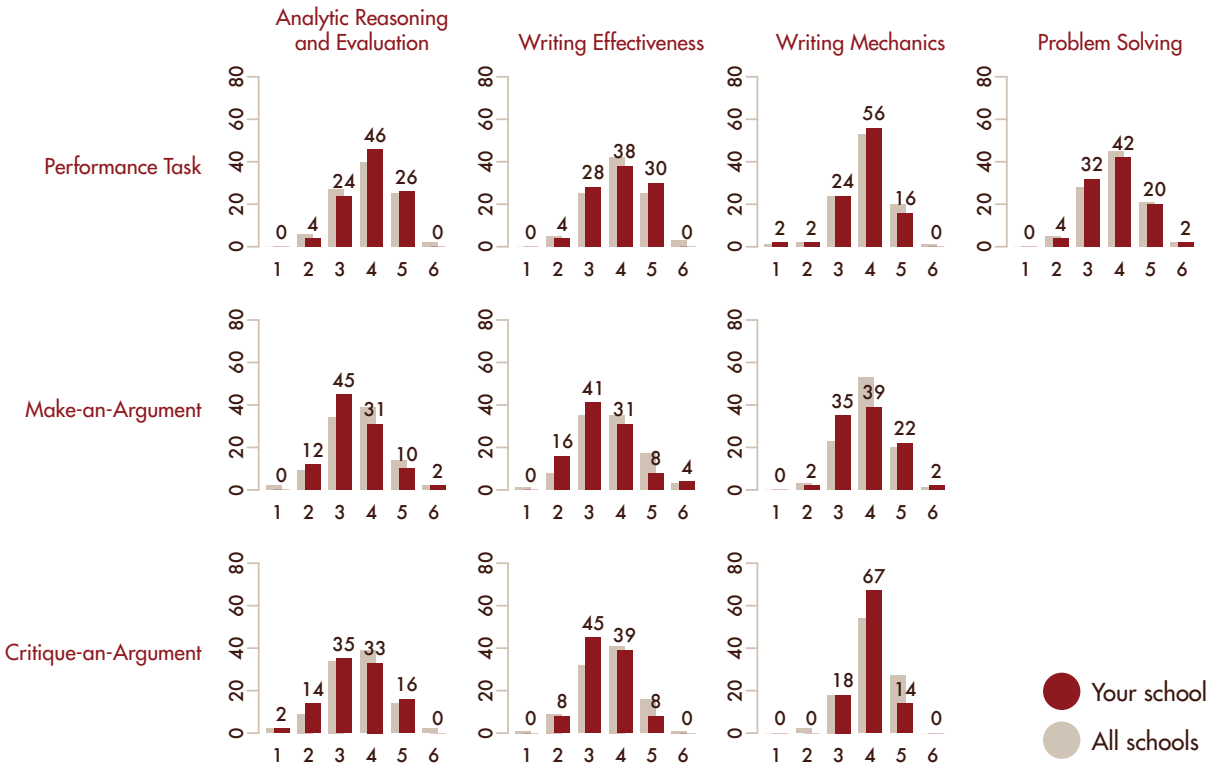
Effect sizes were calculated at a school by taking the difference in mean (or average) CLA scores of the same students from two different phases (e.g., Phase 1 and Phase 3), standardized by the standard deviation of the earlier phase. In each plot, the extreme left hand vertical bar shows the 5th percentile. The "box" itself shows the 25th (left face), 50th (internal vertical line), and 75th (right face) percentile points. The extreme right hand vertical bar shows the 95th percentile. The horizontal axis shows the effect size scale.

An effect size of 0 indicates no difference between phases. Positive effect sizes indicate that scores in later phases are higher than those of earlier phases, with larger effect sizes corresponding to larger score differences.

Subscore Distributions

Figure 2.6 displays the distribution of your students' performance in the subscore categories of Analytic Reasoning and Evaluation, Writing Effectiveness, Writing Mechanics, and Problem Solving. The numbers on the graph correspond to the percentage of *your* students that performed at each score level. The distribution of subscores across *all* schools is presented for comparative purposes. The score levels range from 1 to 6. Note that the graphs presented are not directly comparable due to potential differences in difficulty among task types and among subscores. See *Diagnostic Guidance* and *Scoring Criteria* for more details on the interpretation of subscore distributions. Table 2.7 presents the mean and standard deviation of each of the subscores across CLA task types—for your school and all schools.

2.6 Distribution of Subscores (Spring 2011)



2.7 Summary Subscore Statistics (Spring 2011)

		Analytic Reasoning and Evaluation		Writing Effectiveness		Writing Mechanics		Problem Solving	
		Your School	All Schools	Your School	All Schools	Your School	All Schools	Your School	All Schools
Performance Task	Mean	3.9	3.9	3.9	4.0	3.8	3.9	3.8	3.9
	Standard Deviation	0.8	0.8	0.9	0.9	0.8	0.7	0.9	0.8
Make-an-Argument	Mean	3.4	3.6	3.4	3.7	3.9	3.9		
	Standard Deviation	0.9	0.9	1.0	0.9	0.9	0.7		
Critique-an-Argument	Mean	3.5	3.6	3.5	3.7	4.0	4.1		
	Standard Deviation	1.0	1.0	0.8	0.8	0.6	0.7		

This section describes the longitudinal cohort of institutions that serves as the reference group for your institution's results.

During the fall 2005 - spring 2009 assessment cycles, 31 institutions tested a sufficient number of students to provide the across-school benchmark results presented in this report.

Table 3.1 shows CLA longitudinal schools grouped by Basic Carnegie Classification. The spread of schools differs slightly from that of the 1,587 four-year, not-for-profit institutions across the nation; doctorate-granting universities are somewhat overrepresented among CLA schools, while master's colleges and baccalaureate colleges are somewhat underrepresented.

3.1 Carnegie Classification of Institutional Sample

	Nation		CLA	
Carnegie Classification	Number	Percentage	Number	Percentage
Doctorate-granting Universities	275	17	12	39
Master's Colleges and Universities	619	39	9	29
Baccalaureate Colleges	693	44	10	32

Source: Carnegie Foundation for the Advancement of Teaching, Carnegie Classifications Data File, February 11, 2010.

Table 3.2 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.

3.2 School Characteristics of Institutional Sample

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.

Table 3.3 below presents summary statistics for all students including counts, means, 25th and 75th percentiles, and standard deviations for Phases 1, 2, and 3.

3.3 Student-Level Summary Statistics

Phase 1	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	9168	990	1096	1198	149
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-Argument	9627	869	1079	1167	186
EAA	11360	950	1093	1230	195

Phase 2	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	3141	1025	1141	1249	160
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-Argument	3202	1018	1120	1316	180
EAA	3329	1010	1134	1270	188

Phase 3	Number of Students	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	2290	1115	1222	1339	163
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-Argument	2317	1083	1240	1393	210
EAA	2380	1030	1153	1300	184

Table 3.4 below presents summary statistics for all students including counts, means, 25th and 75th percentiles, and standard deviations for Phases 1, 2, and 3. Note that the unit of analysis is schools, not students.

3.4 School-Level Summary Statistics

Phase 1	Number of Schools	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	47	1038	1100	1170	93
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-Argument	48	1006	1084	1153	103
EAA	48	1011	1104	1209	152

Phase 2	Number of Schools	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	32	1073	1132	1185	96
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-Argument	32	1051	1114	1160	94
EAA	32	1044	1116	1193	129

Phase 3	Number of Schools	25th Percentile	Mean Score	75th Percentile	Standard Deviation
Total CLA Score	30	1138	1202	1267	104
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-Argument	30	1152	1219	1281	111
EAA	32	1049	1126	1217	137

The information presented in your institutional report—enhanced most recently through the provision of subscores (see page 9)—is designed to help you better understand the contributions your institution is making toward your students' learning gains. However, the institutional report alone provides but a snapshot of student performance.

When combined with the other tools and services the CLA has to offer, the institutional report can become a power tool in helping you and your institution target specific areas of improvement, and effectively and authentically align teaching, learning, and assessment practices in ways that may improve institutional performance over time.

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, collaborate with their peers, and participate in professional development offerings.

Student-level CLA results are provided for you to link to other data sources (e.g., course-taking patterns, grades, portfolios, student surveys, etc.). These results are strengthened by the provision of additional scores in the areas of analytic reasoning and evaluation, writing effectiveness, writing mechanics, and problem solving to help you better pinpoint specific areas that may need improvement. Internal analyses, which you can pursue through in-depth sampling, can help you generate hypotheses for additional research.

Beyond the institution-specific results you receive, which can facilitate the alignment of teaching, learning, and assessment, the CLA fosters collaborative relationships among our participating schools. The CLA does this by encouraging the formation of consortia, hosting periodic web conferences featuring campuses doing promising work using the CLA, and sharing school-specific contact information (where permission has been granted) via our CLA contact map (*www.collegiatelearningassessment.org/contact*).

Our professional development services shift the focus from general assessment to the course-level work of faculty members. Performance Task Academies—two-day hands on training workshops—provide opportunities for faculty to receive guidance in creating their own CLA-like performance tasks, which can be used as classroom or homework assignments, curriculum devices or even local-level assessments (see: *www.claintheclassroom.org*).

Through the steps noted above we encourage institutions to move toward a continuous system of improvement stimulated by the CLA. Our programs and services—when used in combination—are designed to emphasize the notion that, in order to successfully improve higher-order skills, institutions must genuinely connect their teaching, learning, and assessment practices in authentic and effective ways.

Without your contributions, the CLA would not be on the exciting path that it is today. We look forward to your continued involvement!

Introduction

The CLA consists of a Performance Task and an Analytic Writing Task. Students are randomly assigned to take one or the other. The Analytic Writing Task includes a pair of prompts called Make-an-Argument and Critique-an-Argument.

All CLA tasks are administered online and consist of open-ended prompts that require constructed responses. There are no multiple-choice questions.

The CLA requires that students use critical thinking and written communication skills to perform cognitively demanding tasks. The integration of these skills 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 may also 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 arguments from emotional ones 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: 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; organizing the structure of the essay to maintain the flow of the argument (e.g., paragraphing, ordering of ideas and sentences within paragraphs, use of transitions); employing varied sentence structure and advanced vocabulary.

Critique-an-Argument

A Critique-an-Argument prompt asks students, in 30 minutes, to evaluate the reasoning used in an argument (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 grammatically correct, organized, well-developed, and logically sound.

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.

CLA results operate as a signaling tool of overall institutional performance on tasks that measure higher-order skills. Examining performance across CLA task types can serve as an initial diagnostic exercise. The three types of CLA tasks—Performance Task, Make-an-Argument, and Critique-an-Argument—differ in the combination of skills necessary to perform well.

The Make-an-Argument and Critique-an-Argument tasks measure Analytic Reasoning and Evaluation, Writing Effectiveness, and Writing Mechanics. The Performance Task measures Problem Solving in addition to the three aforementioned skills. Each of the skills are assessed in slightly different ways within the context of each task type. For example, in the context of the Performance Task and the Critique-an-Argument task, Analytic Reasoning and Evaluation involves interpreting, analyzing, and evaluating the quality of information. In the Make-an-Argument task, Analytic Reasoning and Evaluation involves stating a position, providing valid reasons to support the writer's position, and considering and possibly refuting alternative viewpoints.

Subscores are assigned on a scale of 1 (lowest) to 6 (highest). Subscores are not directly comparable to one another because they are not adjusted for difficulty like CLA scale scores. The subscores remain unadjusted because they are intended to facilitate criterion-referenced interpretations. For example, a "4" in Analytic Reasoning and Evaluation means that a response had certain qualities (e.g., "Identifies a few facts or ideas that support or refute all major arguments"), and any adjustment to that score would compromise the interpretation.

Still, the ability to make claims like "Our students seem to be doing better in Writing Effectiveness than in Problem Solving on the Performance Task" is clearly desirable. This can be done by comparing each subscore distribution to its corresponding reference distribution displayed in Figure 4 of your institutional report. You can support claims like the one above if you see, for example, that students are performing above average in Writing Effectiveness, but not in Problem Solving on the Performance Task.

Please examine the results presented in Figures 3.6 & 3.8 and Tables 3.7 & 3.9 in combination with the *Scoring Criteria* in the next section to explore the areas where your students may need improvement.

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.

Performance Task Scoring Criteria

	Analytic Reasoning & Evaluation Interpreting, analyzing, and evaluating the quality of information. This entails identifying information that is relevant to a problem, highlighting connected and conflicting information, detecting flaws in logic and questionable assumptions, and explaining why information is credible, unreliable, or limited.	Writing Effectiveness Constructing organized and logically cohesive arguments. Strengthening the writer's position by providing elaboration on facts or ideas (e.g., explaining how evidence bears on the problem, providing examples, and emphasizing especially convincing evidence).	Writing Mechanics Facility with the conventions of standard written English (agreement, tense, capitalization, punctuation, and spelling) and control of the English language, including syntax (sentence structure) and diction (word choice and usage).	Problem Solving Considering and weighing information from discrete sources to make decisions (draw a conclusion and/or propose a course of action) that logically follow from valid arguments, evidence, and examples. Considering the implications of decisions and suggesting additional research when appropriate.
6	• Identifies most facts or ideas that support or refute all major arguments (or salient features of all objects to be classified) presented in the Document Library. Provides analysis that goes beyond the obvious. • Demonstrates accurate understanding of a large body of information from the Document Library. • Makes several accurate claims about the quality of information.	• Organizes response in a logically cohesive way that makes it very easy to follow the writer's arguments. • Provides valid and comprehensive elaboration on facts or ideas related to each argument and clearly cites sources of information.	• Demonstrates outstanding control of grammatical conventions. • Consistently writes well-constructed, complex sentences with varied structure and length. • Displays adept use of vocabulary that is precise, advanced, and varied.	• Provides a decision and a solid rationale based on credible evidence from a variety of sources. Weighs other options, but presents the decision as best given the available evidence. When applicable: • Proposes a course of action that follows logically from the conclusion. Considers implications. • Recognizes the need for additional research. Recommends specific research that would address most unanswered questions.
5	• Identifies several facts or ideas that support or refute all major arguments (or salient features of all objects to be classified) presented in the Document Library. • Demonstrates accurate understanding of much of the Document Library content. • Makes a few accurate claims about the quality of information.	• Organizes response in a logically cohesive way that makes it fairly easy to follow the writer's arguments. • Provides valid elaboration on facts or ideas related to each argument and cites sources of information.	• Demonstrates very good control of grammatical conventions. • Consistently writes well-constructed sentences with varied structure and length. • Uses varied and sometimes advanced vocabulary that effectively communicates ideas.	• Provides a decision and a solid rationale based largely on credible evidence from multiple sources and discounts alternatives. When applicable: • Proposes a course of action that follows logically from the conclusion. May consider implications. • Recognizes the need for additional research. Suggests research that would address some unanswered questions.
4	• Identifies a few facts or ideas that support or refute all major arguments (or salient features of all objects to be classified) presented in the Document Library. • Briefly demonstrates accurate understanding of important Document Library content, but disregards some information. • Makes very few accurate claims about the quality of information.	• Organizes response in a way that makes the writer's arguments and logic of those arguments apparent but not obvious. • Provides valid elaboration on facts or ideas several times and cites sources of information.	• Demonstrates good control of grammatical conventions with few errors. • Writes well-constructed sentences with some varied structure and length. • Uses vocabulary that clearly communicates ideas but lacks variety.	• Provides a decision and credible evidence to back it up. Possibly does not account for credible, contradictory evidence. May attempt to discount alternatives. When applicable: • Proposes a course of action that follows logically from the conclusion. May briefly consider implications. • Recognizes the need for additional research. Suggests research that would address an unanswered question.
3	• Identifies a few facts or ideas that support or refute several arguments (or salient features of all objects to be classified) presented in the Document Library. • Disregards important information or makes minor misinterpretations of information. May restate information "as is." • Rarely, if ever, makes claims about the quality of information and may present some unreliable evidence as credible.	• Provides limited or somewhat unclear arguments. Presents relevant information in each response, but that information is not woven into arguments. • Provides elaboration on facts or ideas a few times, some of which is valid. Sources of information are sometimes unclear.	• Demonstrates fair control of grammatical conventions with frequent minor errors. • Writes sentences that read naturally but tend to have similar structure and length. • Uses vocabulary that communicates ideas adequately but lacks variety.	• Provides or implies a decision and some reason to favor it, but the rationale may be contradicted by unaccounted for evidence. When applicable: • Briefly proposes a course of action, but some aspects may not follow logically from the conclusion. • May recognize the need for additional research. Any suggested research tends to be vague or would not adequately address unanswered questions.
2	• Identifies very few facts or ideas that support or refute arguments (or salient features of all objects to be classified) presented in the Document Library. • Disregards or misinterprets much of the Document Library. May restate information "as is." • Does not make claims about the quality of information and presents some unreliable information as credible.	• Provides limited, invalid, overstated, or very unclear arguments. May present information in a disorganized fashion or undermine own points. • Any elaboration on facts or ideas tends to be vague, irrelevant, inaccurate, or unreliable (e.g., based entirely on writer's opinion). Sources of information are often unclear.	• Demonstrates poor control of grammatical conventions with frequent minor errors and some distracting errors. • Consistently writes sentences with similar structure and length, and some may be difficult to understand. • Uses simple vocabulary, and some vocabulary may be used inaccurately or in a way that makes meaning unclear.	• Provides or implies a decision, but very little rationale is provided or it is based heavily on unreliable evidence. When applicable: • Briefly proposes a course of action, but some aspects do not follow logically from the conclusion. • May recognize the need for additional research. Any suggested research is vague or would not adequately address unanswered questions.
1	• Does not identify facts or ideas that support or refute arguments (or salient features of all objects to be classified) presented in the Document Library or provides no evidence of analysis. • Disregards or severely misinterprets important information. • Does not make claims about the quality of evidence and bases response on unreliable information.	• Does not develop convincing arguments. Writing may be disorganized and confusing. • Does not provide elaboration on facts or ideas.	• Demonstrates minimal control of grammatical conventions with many errors that make the response difficult to read or provides insufficient evidence to judge. • Writes sentences that are repetitive or incomplete, and some are difficult to understand. • Uses simple vocabulary, and some vocabulary is used inaccurately or in a way that makes meaning unclear.	• Provides no clear decision or no valid rationale for the decision. When applicable: • Does not propose a course of action that follows logically from the conclusion. • Does not recognize the need for additional research or does not suggest research that would address unanswered questions.

Make-an-Argument Scoring Criteria

	Analytic Reasoning & Evaluation Stating a position, providing valid reasons to support the writer's position, and demonstrating an understanding of the complexity of the issue by considering and possibly refuting alternative viewpoints.	Writing Effectiveness Constructing an organized and logically cohesive argument. Strengthening the writer's position by elaborating on the reasons for that position (e.g., providing evidence, examples, and logical reasoning).	Writing Mechanics Facility with the conventions of standard written English (agreement, tense, capitalization, punctuation, and spelling) and control of the English language, including syntax (sentence structure) and diction (word choice and usage).
6	• Asserts an insightful position and provides multiple (at least 4) sound reasons to justify it. • Provides analysis that reflects a thorough consideration of the complexity of the issue. Possibly refutes major counterarguments or considers contexts integral to the issue (e.g., ethical, cultural, social, political).	• Organizes response in a logically cohesive way that makes it very easy to follow the writer's argument. • Provides valid and comprehensive elaboration on each reason for the writer's position.	• Demonstrates outstanding control of grammatical conventions. • Consistently writes well-constructed, complex sentences with varied structure and length. • Displays adept use of vocabulary that is precise, advanced, and varied.
5	• States a thoughtful position and provides multiple (at least 3) sound reasons to support it. • Provides analysis that reflects some consideration of the complexity of the issue. Possibly considers contexts integral to the issue (e.g., ethical, cultural, social, political).	• Organizes response in a logically cohesive way that makes it fairly easy to follow the writer's argument. • Provides valid elaboration on each reason for the writer's position.	• Demonstrates very good control of grammatical conventions. • Consistently writes well-constructed sentences with varied structure and length. • Uses varied and sometimes advanced vocabulary that effectively communicates ideas.
4	• States a clear position and some (2-3) sound reasons to support it. • Provides some careful analysis, but it lacks consideration of the issue's complexity.	• Organizes response in a way that makes the writer's argument and its logic apparent but not obvious. • Provides valid elaboration on reasons for the writer's position several times.	• Demonstrates good control of grammatical conventions with few errors. • Writes well-constructed sentences with some varied structure and length. • Uses vocabulary that clearly communicates ideas but lacks variety.
3	• States or implies a position and provides few (1-2) reasons to support it. • Provides some superficial analysis of the issue.	• Provides a limited or somewhat unclear argument. Presents relevant information, but that information is not woven into an argument. • Provides valid elaboration on reasons for the writer's position a few times.	• Demonstrates fair control of grammatical conventions with frequent minor errors. • Writes sentences that read naturally but tend to have similar structure and length. • Uses vocabulary that communicates ideas adequately but lacks variety.
2	• States or implies a position and provides vague or very few reasons to support it. • Provides little analysis, and that analysis may reflect an oversimplification of the issue.	• Provides limited, invalid, overstated, or very unclear argument. May present information in a disorganized fashion or undermine own points. • Any elaboration on reasons for the writer's position tend to be vague, irrelevant, inaccurate, or unreliable (e.g., based entirely on writer's opinion).	• Demonstrates poor control of grammatical conventions with frequent minor errors and some distracting errors. • Consistently writes sentences with similar structure and length, and some may be difficult to understand. • Uses simple vocabulary, and some vocabulary may be used inaccurately or in a way that makes meaning unclear.
1	• States an unclear position (if any) and fails to provide reasons to support it. • Provides very little evidence of analysis. May not understand the issue.	• Fails to develop a convincing argument. The writing may be disorganized and confusing. • Fails to provide elaboration on reasons for the writer's position.	• Demonstrates minimal control of grammatical conventions with many errors that make the response difficult to read or provides insufficient evidence to judge. • Writes sentences that are repetitive or incomplete, and some are difficult to understand. • Uses simple vocabulary, and some vocabulary is used inaccurately or in a way that makes meaning unclear.

Critique-an-Argument Scoring Criteria

	Analytic Reasoning & Evaluation Interpreting, analyzing, and evaluating the quality of information. This entails highlighting conflicting information, detecting flaws in logic and questionable assumptions, and explaining why information is credible, unreliable, or limited.	Writing Effectiveness Constructing organized and logically cohesive arguments. Strengthening the writer's position by elaborating on deficiences in the argument (e.g., providing explanations and examples).	Writing Mechanics Facility with the conventions of standard written English (agreement, tense, capitalization, punctuation, and spelling) and control of the English language, including syntax (sentence structure) and diction (word choice and usage).
6	• Demonstrates accurate understanding of the complete argument. • Identifies many (at least 5) deficiencies in the argument and provides analysis that goes beyond the obvious.	• Organizes response in a logically cohesive way that makes it very easy to follow the writer's critique. • Provides valid and comprehensive elaboration for each identified deficiency.	• Demonstrates outstanding control of grammatical conventions. • Consistently writes well-constructed, complex sentences with varied structure and length. • Displays adept use of vocabulary that is precise, advanced, and varied.
5	• Demonstrates accurate understanding of much of the argument. • Identifies many (at least 4) deficiencies in the argument.	• Organizes response in a logically cohesive way that makes it fairly easy to follow the writer's critique. • Provides valid elaboration for each identified deficiency.	• Demonstrates very good control of grammatical conventions. • Consistently writes well-constructed sentences with varied structure and length. • Uses varied and sometimes advanced vocabulary that effectively communicates ideas.
4	• Demonstrates accurate understanding of several aspects of the argument, but disregards a few. • Identifies several (at least 3) deficiencies in the argument.	• Organizes response in a way that makes the writer's critique and its logic apparent but not obvious. • Provides valid elaboration on identified deficiencies several times.	• Demonstrates good control of grammatical conventions with few errors. • Writes well-constructed sentences with some varied structure and length. • Uses vocabulary that clearly communicates ideas but lacks variety.
3	• Disregards several aspects of the argument or makes minor misinterpretations of the argument. • Identifies a few (2-3) deficiencies in the argument.	• Provides a limited or somewhat unclear critique. Presents relevant information, but that information is not woven into an argument. • Provides valid elaboration on identified deficiencies a few times.	• Demonstrates fair control of grammatical conventions with frequent minor errors. • Writes sentences that read naturally but tend to have similar structure and length. • Uses vocabulary that communicates ideas adequately but lacks variety.
2	• Disregards or misinterprets much of the information in the argument. • Identifies very few (1-2) deficiencies in the argument and may accept unreliable evidence as credible.	• Provides limited, invalid, overstated, or very unclear critique. May present information in a disorganized fashion or undermine own points. • Any elaboration on identified deficiencies tends to be vague, irrelevant, inaccurate, or unreliable (e.g., based entirely on writer's opinion).	• Demonstrates poor control of grammatical conventions with frequent minor errors and some distracting errors. • Consistently writes sentences with similar structure and length, and some may be difficult to understand. • Uses simple vocabulary, and some vocabulary may be used inaccurately or in a way that makes meaning unclear.
1	• Disregards or severely misinterprets important information in the argument. • Fails to identify deficiencies in the argument or provides no evidence of critical analysis.	• Fails to develop a convincing critique or agrees entirely with the flawed argument. The writing may be disorganized and confusing. • Fails to provide elaboration on identified deficiencies.	• Demonstrates minimal control of grammatical conventions with many errors that make the response difficult to read or provides insufficient evidence to judge. • Writes sentences that are repetitive or incomplete, and some are difficult to understand. • Uses simple vocabulary, and some vocabulary is used inaccurately or in a way that makes meaning unclear.

The CLA uses a combination of automated and human scoring. Since fall 2010, we have relied primarily on Intelligent Essay Assessor (IEA) for scoring. IEA is the automated scoring engine developed by Pearson Knowledge Technologies to evaluate the meaning of text, not just writing mechanics. Pearson has trained IEA for the CLA using a broad range of real CLA responses and scores to ensure its consistency with scores generated by human scorers.

Though the majority of scoring is handled by IEA, some responses are scored by trained human scorers. IEA identifies unusual responses, which are automatically sent to the human scoring queue. In addition, ten percent of responses are scored by both IEA and humans in order to continually evaluate the quality of scoring.

All scorer candidates undergo rigorous training in order to become certified CLA scorers. Training includes an orientation to the prompts and scoring rubrics/guides, 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.

Each response receives subscores in the categories of Analytic Reasoning and Evaluation, Writing Effectiveness, and Writing Mechanics. An additional scale, Problem Solving, is used to evaluate only the Performance Tasks. Subscores are assigned on a scale of 1 (lowest) to 6 (highest). 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 flagged for removal from results.

Because the prompts (specific tasks within each task type) differ in the possible arguments and pieces of information students can or should use in their responses, prompt-specific guidance is provided to scorers in addition to the scoring criteria that appear in the previous section.

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.

Standard ACT to SAT Crosswalk

ACT	to	SAT
36		1600
35		1560
34		1510
33		1460
32		1420
31		1380
30		1340
29		1300
28		1260
27		1220
26		1190
25		1150
24		1110
23		1070
22		1030
21		990
20		950
19		910
18		870
17		830
16		790
15		740
14		690
13		640
12		590
11		530

Source:

ACT (2008). *ACT/College Board Joint Statement*. Retrieved from http://www.act.org/aap/concordance/pdf/report.pdf

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.

In tandem with your report, we provide a CLA Student Data File, which includes variables across three categories: self-reported information from students in their CLA online profile; CLA scores and identifiers; and information provided by the registrar.

We provide student-level information for linking 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 factors related to 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. In addition, correlations between individual CLA scores and other measures would be attenuated due to unreliability.

Self-Reported Data

- Name (first, middle initial, last)
- Student ID
- E-mail address
- Date of birth
- Gender
- Race/Ethnicity
- Parent Education
- Primary and Secondary Academic Major (36 categories)
- Field of Study (6 categories; based on primary academic major)
- English as primary language
- Attended school as freshman, sophomore, junior, senior
- Local survey responses

CLA Scores and Identifiers

- For Performance Task, Analytic Writing Task, Make-an-Argument, and Critique-an-Argument (depending on the tasks taken and completeness of responses):
  - CLA scores
  - Performance Level categories (i.e., well below expected, below expected, near expected, above expected, well above expected)*
  - Percentile Rank across schools and within your school (among students in the same class year, based on score)
- Subscores in Analytic Reasoning and Evaluation, Writing Effectiveness, Writing Mechanics, and Problem Solving
- SLE score (if applicable, 1-50)
- Entering Academic Ability (EAA) score
- Unique CLA numeric identifiers
- Year, Test window (Fall or Spring), Date of test, and Time spent on test

Registrar Data

- Class Standing
- Transfer Student Status
- Program Code and Name (for classification of students into different colleges, schools, fields of study, programs, etc., if applicable)
- SAT Total (Math + Verbal)
- SAT I Math
- SAT I Verbal / Critical Reading
- SAT I Writing
- ACT Composite
- GPA (not applicable for entering students)

* The residuals that inform these levels are from an OLS regression of CLA scores on EAA scores, across all schools. Roughly 20% of students (within class) fall into each performance level.

Roger Benjamin
President & CEO

James Hundley
Executive Vice President & COO

Katharine Lyall
Chair, CAE
President Emeritus, University of Wisconsin System

Richard Atkinson
President Emeritus, University of California System

Doug Bennett
President, Earlham College

Michael Crow
President, Arizona State University

Russell C. Deyo
Vice President & General Counsel, Johnson & Johnson

Richard Foster
Managing Partner, Millbrook Management Group, LLC

Ronald Gidwitz
Chairman, GCG Partners

Eduardo Marti
Vice Chancellor for Community Colleges, CUNY

Ronald Mason
President, Southern University System

Diana Natalicio
President, University of Texas at El Paso

Charles Reed
Chancellor, California State University

Michael D. Rich
Executive Vice President, RAND Corporation

Farris W. Womack
Executive Vice President and Chief Financial Officer, Emeritus
Professor Emeritus, The University of Michigan

council for aid to education

215 lexington avenue floor 21 new york new york 10016-6023
p | 212.217.0700 f | 212.661.9766 e | cla@cae.org w | www.cae.org/cla