CAAP Critical Thinking Test Summary

Administered May 2011

Office of Institutional Research
July 2011

CAAP Critical Thinking Test Summary

Introduction

The Collegiate Assessment of Academic Proficiency (CAAP) is a standardized, nationally-normed assessment program from ACT that enables postsecondary institutions to assess, evaluate, and enhance student learning outcomes and general education program outcomes. There are six different CAAP tests that institutions can choose to administer: Reading, Writing Skills, Writing Essay, Mathematics, Science, and Critical Thinking.

Owens adopted the CAAP Critical Thinking test in the Spring of 2009, and it is administered every year to students enrolled in randomly-chosen sections of select 200-level courses. The test is 40 minutes in length and includes 32 items that measure students' skills in clarifying, analyzing, evaluating, and extending arguments. An argument is defined as a sequence of statements that includes a claim that one of the statements, the conclusion, follows from the other statements. The Critical Thinking Test consists of four passages that are representative of the kinds of issues commonly encountered in a postsecondary curriculum.

A passage typically presents a series of sub-arguments in support of a more general conclusion or conclusions. Each passage presents one or more arguments using a variety of formats, including case studies, debates, dialogues, overlapping positions, statistical arguments, experimental results, or editorials. Each passage is accompanied by a set of multiple-choice test items. A sample passage with test items is provided in Appendix A.

As indicated in Table 1, test questions fall within one of three content categories. Between 17 and 21 test questions (53 – 66% of questions) assess students' analysis of elements of an argument, between 5 and 9 (16 – 28%) assess students' evaluation of an argument, and 6 (19%) questions assess students' extension of an argument.

Table 1. Content Specifications Summary for the CAAP Critical Thinking Test

Content Category	Proportion of Test	Number of Items
Analysis of elements of an argument	.53–.66	17–21
Evaluation of an argument	.16–.28	5–9
Extension of an argument	.19	6
Total	**1.00**	**32**

Methodology

In May 2011, the Critical Thinking Test was administered to students enrolled in 23 randomly-selected sections of the following courses: Biology 212, English 210, English 215, English 280, Humanities 200, Psychology 220, Psychology 260, and Psychology 280. All of these classes were 200-level courses identified as having a critical thinking element in the curriculum. A total of 518 students were enrolled in these sections during the spring of 2011 and 291 completed the test, resulting in a 56.2% response rate and a 3.8% margin of error.[1] Table 2 shows self-reported demographic characteristics of test respondents.

Table 2. Demographic Characteristics of Test Respondents

Ethnicity	**#**	**%**	**Status**	**#**	**%**	**Major**	**#**	**%**
African American/Black	18	6%	Full-time	191	66%	Undecided	7	2%
Amer. Indian/Alaskan Nat.	3	1%	Part-time	79	27%	Agriculture	0	0%
White/Caucasian	216	74%	No response	21	7%	Architecture	0	0%
Mexican American/Chicano	4	1%	**Cum GPA**			Biological Sciences	3	1%
Asian/Pacific Islander	4	1%	Below 2.00	10	3%	Business	3	1%
Puerto Rican/Cuban/Hisp.	0	0%	2.01 - 2.50	42	14%	Office Management	0	0%
Filipino	0	0%	2.51 - 3.00	58	20%	Marketing & Purchasing	0	0%
Other	9	3%	3.01 - 3.50	86	30%	Communications	3	1%
Prefer not to respond	12	4%	3.51 and above	64	22%	Community Services	10	3%
No response	25	9%	No response	31	11%	Computer & Info. Sciences	1	0%
Gender			**Rank (self-reported)**			General Studies	2	1%
Male	63	22%	Freshman	60	21%	Education	36	12%
Female	212	73%	Sophomore	132	45%	Engineering	0	0%
No response	16	5%	Junior	37	13%	Fine & Applied Arts	1	0%
Age			Senior	11	4%	Foreign Languages	0	0%
18 and under	41	14%	Other	20	7%	Health Professions	136	47%
19 - 20	79	27%	No response	31	11%	Home Economics	1	0%
21 - 25	67	23%	**Enrolled at Owens as Freshmen**			Letters	5	2%
26 - 30	36	12%	Yes	212	73%	Mathematics	0	0%
31 - 39	42	14%	No	57	20%	Philosophy/Religion	0	0%
40 and older	26	9%	No response	22	8%	Physical Sciences	1	0%
No response						Social Sciences	28	10%
						Trade & Industrial	0	0%
						No response	54	19%

[1] The margin of error is a statistic that reflects the amount of sampling error in a survey's results and is based on the size of the sample (n=291) in relation to the size of the population (N=918 students enrolled in all sections of the surveyed courses during the spring of 2011). The lower the margin of error, the more confidence one can have that the data are representative of the full population. Ideally, the margin of error should be 5% or less. The way to reduce the margin of error is to increase the sample size (i.e., survey more students).

Summary of Test Results

As a general guide, differences in scores should be interpreted based on the standard deviation of the score distribution. Standard deviations are a measure of variation in the data such that 68% of scores fall within one standard deviation of the mean (60.7 $\pm$ 5.4 = 55.3 – 66.1), 95% fall within two standard deviations (60.7 $\pm$ 10.8 = 49.9 – 71.5), and 98% are within three standard deviations (60.7 $\pm$ 16.2 = 44.5 – 76.9). Consequently, differences of one standard deviation or less are considered negligible, differences between one and two standard deviations are considered moderate, and differences between two and three or more are considered substantial.

Based on the guide above, the results (displayed in Table 3) show that (a) Owens students differed very little from the two-year public national average, (b) there was very little change from 2010 to 2011, and (c) there was very little difference among different student groups.

Table 3: Average Scores by Student Group: Owens 2009-2011 and National Average

	Owens 2011		Owens 2010		Owens 2009		Current Natl. Ave.			Owens 2011		Owens 2010		Owens 2009		Current Natl. Ave.	
	Ave.	#	Ave.	#	Ave.	#	Ave.	S.D.		Ave.	#	Ave.	#	Ave.	#	Ave.	S.D.
Overall	60.2	291	60.1	107	61.0	208	60.7	5.4	**Status**								
Ethnicity									Full-time	60.0	191	60.0	84	61.0	127	NA	NA
African American/Black	57.0	18	54.0	10	59.0	12	NA	NA	Part-time	60.0	79	60.0	22	62.0	35	NA	NA
Amer. Indian/Alaskan Nat.		3				1	NA	NA	No response	59.0	21		1	61.0	46	NA	NA
White/Caucasian	61.0	216	61.0	77	62.0	129	NA	NA	**Cum GPA**								
Mexican American/Chicano		4		2	59.0	10	NA	NA	Below 2.00	58.0	10	57.0	9	58.0	15	NA	NA
Asian/Pacific Islander		4		1		1	NA	NA	2.01 - 3.00	59.0	42	58.0	15	58.0	24	NA	NA
Puerto Rican/Cuban/Hisp.				1		2	NA	NA	2.51 - 3.00	60.0	58	59.0	26	61.0	36	NA	NA
Filipino						1	NA	NA	3.01 - 3.50	60.0	86	61.0	21	62.0	44	NA	NA
Other	60.0	9		4	61.0	5	NA	NA	3.51 and above	62.0	64	63.0	27	63.0	36	NA	NA
Prefer not to respond	58.0	12	61.0	6	61.0	8	NA	NA	No response	59.0	31	59.0	9	61.0	53	NA	NA
No response	58.0	25	57.0	5	61.0	39	NA	NA	**Major**								
Gender									Undecided	57.0	7		3	61.0	8	NA	NA
Male	62.0	63	61.0	49	62.0	69	NA	NA	Agriculture				1		4	NA	NA
Female	60.0	212	60.0	56	61.0	109	NA	NA	Architecture				1		1	NA	NA
No response	58.0	16		2	61.0	30	NA	NA	Biological Sciences		3		3	63.0	5	NA	NA
Age									Business		3	58.0	7	63.0	7	NA	NA
18 and under	59.0	41	61.0	10	60.0	58	NA	NA	Office Management							NA	NA
19 - 20	60.0	79	59.0	35	61.0	53	NA	NA	Marketing & Purchasing						1	NA	NA
21 - 25	60.0	67	60.0	25	61.0	46	NA	NA	Communications		3	61.0	6		3	NA	NA
26 - 30	62.0	36	62.0	15	61.0	23	NA	NA	Community Services	60.0	10	58.0	12	61.0	6	NA	NA
31 - 39	61.0	42	60.0	9	62.0	17	NA	NA	Computer & Info. Sciences		1					NA	NA
40 and older	58.0	26	59.0	13	62.0	11	NA	NA	General Studies		2		4		1	NA	NA
No response							NA	NA	Education	61.0	36	59.0	13	63.0	8	NA	NA
Rank (self-reported)									Engineering				4		1	NA	NA
Freshman	60.0	60	58.7	39	60.0	69	60.7	5.4	Fine & Applied Arts		1		2	62.0	5	NA	NA
Sophomore	60.0	132	60.2	37	62.0	72	60.7	5.4	Foreign Languages							NA	NA
Junior	60.0	37	61.0	11	63.0	23	NA	NA	Health Professions	60.0	136	59.0	20	61.0	45	NA	NA
Senior	61.0	11	64.0	7		3	NA	NA	Home Economics		1		1			NA	NA
Other	61.0	20	62.0	11	62.0	9	NA	NA	Letters	63.0	5		1		1	NA	NA
No response	58.0	31		2	60.0	32	NA	NA	Mathematics				1			NA	NA
Enrolled as a Freshmen									Philosophy/Religion				1			NA	NA
Yes	60.0	212	60.0	77	61.0	124	NA	NA	Physical Sciences		1		4		2	NA	NA
No	61.0	57	60.0	27	62.0	36	NA	NA	Social Sciences	61.0	28	60.0	17	61.0	8	NA	NA
No response	59.0	22		3	61.0	48	NA	NA	Trade & Industrial				1			NA	NA
									No response	59.0	54	60.0	5	60.0	102	NA	NA

NOTES: ACT does not provide data for groups with a sample size less than 5. National averages are only available in aggregate and for freshmen and sophomores

The only average with a difference greater than 5.4 was that for African American students in 2010. However, because the sample size is extremely small (n=10) and because a national average for this group is not provided, this result should be interpreted with caution.

Detailed Results

For the first time since Owens started administering the Critical Thinking test, a Content Analysis Report was acquired from CAAP. The report provides information to better identify specific content areas in which Owens students are strong or weak relative to a normative group of students. This detailed report is available in Appendix B.

Table 4 shows the comparisons between Owens students (Local cohort) and normative groups upon the three measurements of the critical thinking test: analysis of arguments, evaluation of arguments, and extension of arguments. These are further broken down by student performance on the test: bottom 25%, middle 50%, and top 25%.

Differences with magnitudes less than 5%, between 5% and 10%, and greater than 10% are considered negligible, moderate, and substantial, respectively.

Table 4: Comparison of Local/Normative Groups (Local/Normative Differences in Percent Correct)

Content Category	Bottom 25%	Middle 50%	Top 25%
Analysis of Arguments	2%	1%	0%
Evaluation of Arguments	3%	8%	3%
Extension of Arguments	5%	4%	-2%

There were few differences noted among any of these measurements and subgroups, but two are notable: there was a moderate difference between local and normative extension of arguments, with the bottom 25% of OCC students scoring higher than the national group (32% compared to 27%, a difference of 5%).

There was likewise a moderate difference between the groups for the evaluation of arguments, with the middle 50% of OCC students scoring above the national group (54% compared to 46%, a difference of 8%).

All other differences were negligible as per CAAP specifications.

Conclusions, Limitations, & Questions for Future Research

Overall, results indicate that:

- Owens students do not differ meaningfully on their critical thinking skills from a national sample of students at other two-year public institutions.

- Within the detailed analysis, students in the bottom 25% scored higher than the national group on extension of arguments, indicating a moderate difference.

- Students in the middle 50% group likewise scored better than the national group in evaluation of arguments, also a moderate level of difference.

- There was no real change in students' critical thinking skills from 2009 to 2011.

- There are no meaningful differences in critical thinking skills among different student groups of Owens students.

However, there are several limitations to the data that should be considered in the interpretation and use of results:

- Differences in scores from the national average, from year-to-year, and between groups are statistically very small and well within the normal range. Only differences of 5 points or more are unlikely to be due to chance and are of practical significance.

- Sample sizes for various student groups are very small and national averages for different student groups are not available; therefore, data by student demographic characteristics should be interpreted very cautiously.

Appendix A

Sample Passage 1

Senator Favor proposed a bill in the state legislature that would allow pharmacists to prescribe medications for minor illnesses, without authorization from a physician (i.e., a "prescription"). In support of her proposal, Favor argued:

Doctors have had a monopoly on authorizing the use of prescription medicines for too long. This has caused consumers of this state to incur unnecessary expense for their minor ailments. Often, physicians will require patients with minor complaints to go through an expensive office visit before the physician will authorize the purchase of the most effective medicines available to the sick.

Consumers are tired of paying for these unnecessary visits. At a recent political rally in Johnson County, I spoke to a number of my constituents and a majority of them confirmed my belief that this burdensome, expensive, and unnecessary practice is widespread in our state. One man with whom I spoke said that his doctor required him to spend $80 on an office visit for an uncommon skin problem which he discovered could be cured with a $2 tube of prescription cortisone lotion.

Anyone who has had to wait in a crowded doctor's office recently will be all too familiar with the "routine": after an hour in the lobby and a half-hour in the examining room, a physician rushes in, takes a quick look at you, glances at your chart and writes out a prescription. To keep up with the dizzying pace of "health care," physicians rely more and more upon prescriptions, and less and less upon careful examination, inquiry, and bedside manner.

Physicians make too much money for the services they render. If "fast food" health care is all we are offered, we might as well get it at a good price. This bill, if passed into law, would greatly decrease unnecessary medical expenses and provide relief to the sick: people who need all the help they can get in these trying economic times. I urge you to vote for this bill.

After Senator Favor's speech, Senator Counter stood to present an opposing position, stating:

Senator Favor does a great injustice to the physicians of this state in generalizing from her own health care experiences. If physicians' offices are crowded, they are crowded for reasons that are different from those suggested by Senator Favor. With high operating costs, difficulties in collecting medical bills, and exponential increases in the costs of malpractice insurance, physicians are lucky to keep their heads above water. In order to do so, they must make their practices more efficient, relying upon nurses and laboratories to do some of the patient screening.

No one disputes the fact that medical expenses are soaring. But, there are issues at stake which are more important than money—we must consider the quality of health care. Pharmacists are not trained to diagnose illnesses. Incorrect diagnoses by pharmacists could lead to extended illness or even death for an innocent customer. If we permit such diagnoses, we will be personally responsible for those illnesses and deaths.

Furthermore, since pharmacies make most of their money by selling prescription drugs, it would be unwise to allow pharmacists to prescribe. A sick person who has not seen a physician might go into a drugstore for aspirin and come out with narcotics!

Finally, with the skyrocketing cost of insurance, it would not be profitable for pharmacists to open themselves up to malpractice suits for mis-prescribing drugs. It is difficult enough for physicians with established practices to make it; few pharmacists would be willing to take on this financial risk. I recommend that you vote against this bill.

Sample Items for Passage 1

1. Favor's "unofficial poll" of her constituents at the Johnson County political rally would be more persuasive as evidence for her contentions if the group of people to whom she spoke had:
   I. been randomly selected.
   II. represented a broad spectrum of the population: young and old, white and non-white, male and female, etc.
   III. not included an unusually large number of pharmacists.

A. I only
B. II only
C. III only
D. I, II, and III

2. In her example of the man who paid $80 for an office visit to treat an uncommon skin problem, Favor seems to assume, but probably should not, that:
   A. the man would have discovered this cure without the doctor's diagnosis.
   B. two dollars is the average price of the cortisone lotion.
   C. eighty dollars is the average price for an office visit of this kind.
   D. cortisone lotion is effective on all rashes.

3. Counter's concern that a sick person who has not seen a physician might go into a drugstore for aspirin and come out with narcotics is probably unfounded because:
   A. sick persons often send others to get their drugs.
   B. narcotics are not normally prescribed for "minor ailments."
   C. most people do not buy aspirin at the drugstore.
   D. most people who need narcotics go to a physician to get them.

4. It is obvious from Favor's speech that she believes which of the following?
   A. Most prescriptions are unnecessary.
   B. Senator Counter will oppose the bill.
   C. If the bill is passed into law, it will greatly reduce the cost of all medical treatment.
   D. If the bill is passed, the average costs for treatment of minor ailments would be reduced significantly.

5. It is clear from Senator Counter's speech that he believes:
   A. physicians are not having difficult economic times.
   B. Favor's description of the crowded physician's office is not completely inaccurate.
   C. the cost of malpractice insurance is not growing at an accelerated pace.
   D. the quality of health care will not diminish if pharmacists are allowed to prescribe drugs.

Appendix B

CAAP

Content Analysis Report

Owens Community College

Institution Code: 3323
Institution Type: 2-Year
Subgroup: none
Test Date: Spring 2011

Normative Group: National 2-year Colleges

Introduction

This report provides information intended to help postsecondary institutions better identify specific content areas in which their students are strong or weak relative to a normative group of students.

Results within this report are generated after several crucial steps:	For each CAAP test that your students have taken:
1. Students in both the local and normative groups (or, cohorts) are classified into proficiency groups according to whether scale scores for a given CAAP test were in the top 25%, middle 50%, or bottom 25% of their cohort.	1. Results for content categories are presented separately for the normative group and local cohorts of students. Note that the normative group represents students who have tested using a specific CAAP test form within the past three years. Local and normative group percents enable you to compare your students' performance to that of users in the normative group.
2. For each CAAP test that a given student has taken (e.g., Mathematics, Reading, Writing Skills, Science, or Critical Thinking), item responses (correct or incorrect) are used to compute the percentage of items answered correctly within each content category.	2. In this report's figures, proficiency group results are presented adjacent to one another. This enables you to compare content category performance for your students among levels of proficiency.
3. For each student proficiency group, students' percent-correct scores on items within content categories are then averaged, resulting in the percent of items correctly answered. This information is presented in this report's figures.	3. Results reflect major content categories within each CAAP test. For instance,the CAAP Reading test is composed of Referring Skills and Reasoning Skills content categories. Through a comparison of the percent of items correct across content categories, you can determine your students' relative strengths and weaknesses for each specific content category.

Critical Thinking

The CAAP Critical Thinking Test is a 32-item test that measures students' skills in analyzing, evaluating, and extending arguments. An argument is defined as a sequence of statements that includes a claim that one of the statements, the conclusion, follows from the other statements. The test presents arguments using a variety of formats, including case studies, debates, dialogues, overlapping positions, statistical arguments, experimental results, and editorials. Arguments are embedded in issues that are likely to be encountered in a postsecondary curriculum.

Executive Summary

This report contains results from CAAP Critical Thinking form 12-A. Valid scores were obtained for 291 students. Table C-1 belows shows differences in percentages between the local and normative groups. Differences with magnitudes less than 5%, between 5% and 10%, and greater than 10% are considered negligible, moderate, and substantial, respectively. Negative differences indicate areas where local students had more difficulty with content category items than did the normative group, whereas positive differences indicate that local students found items easier than did the normative group.

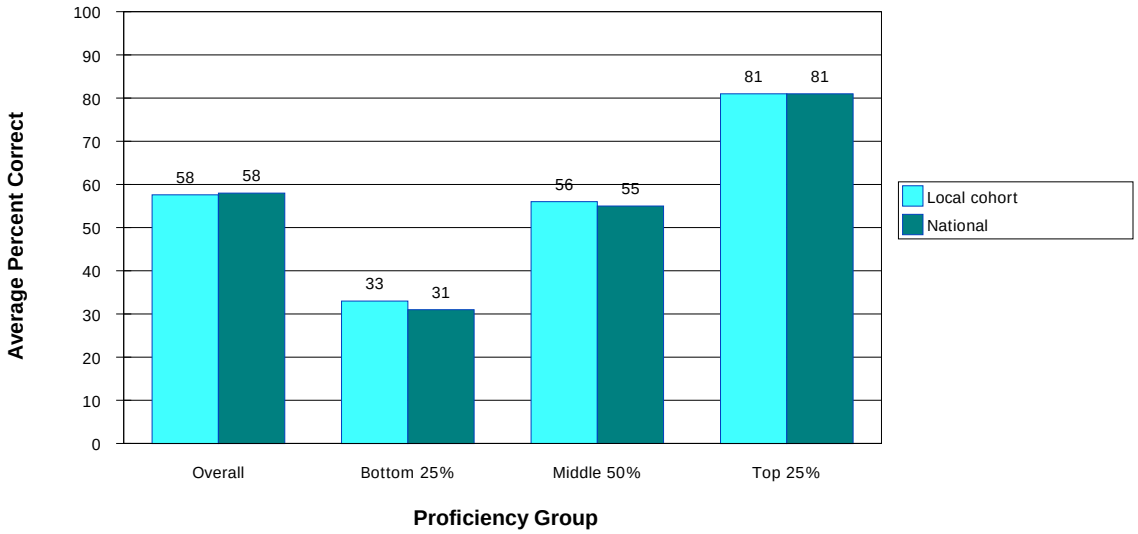
N

Table C-1: Critical Thinking Comparison Highlights

	Local-Normative Group Differences in Percent Correct		
Content Category	Bottom 25%	Middle 50%	Top 25%
Analysis of Arguments	2%	1%	0%
Evaluation of Arguments	3%	8%	3%
Extension of Arguments	5%	4%	- 2%

Critical Thinking
Content Area: Analysis of Arguments

Figure C-1: Average Percent of Correct Answers in Analysis of Arguments



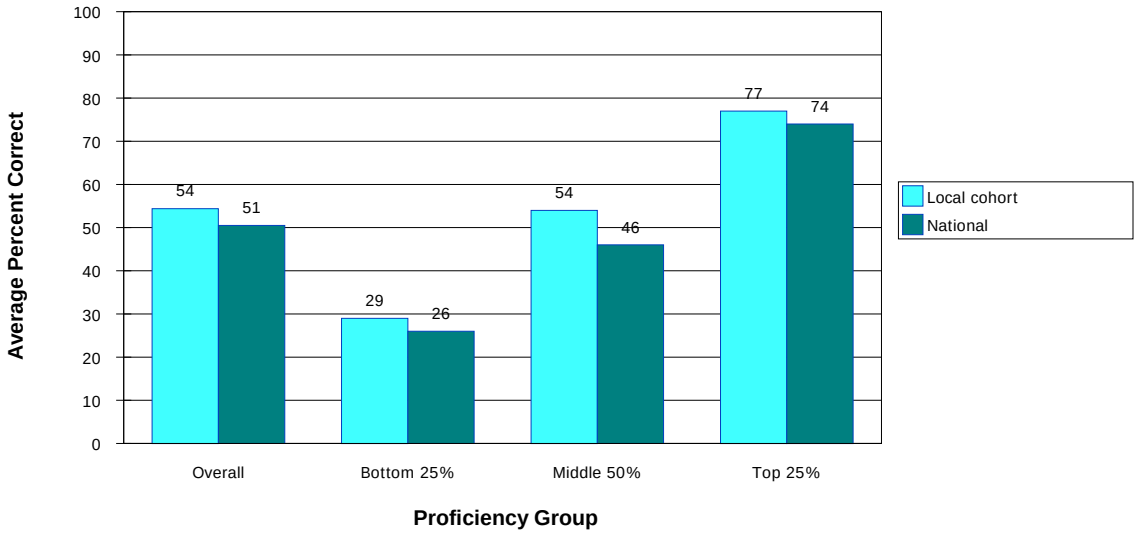
Interpretation Guide

The results in Figure C-1 are based on your students' responses to the Analysis of Arguments items on the CAAP Critical Thinking test. There are 20 items in this content category, constituting approximately 63% of the Critical Thinking test. Items in this category assess the students' ability to identify essential elements of an argument, including hypotheses, premises, and conclusions, and also their ability to identify logical fallacies, exaggerated claims, unstated assumptions, analogies, and multiple points of view. Students are also tested regarding their ability to analyze the structure of arguments, including their ability to distinguish between statements of fact and opinion, to make judgments about equivalent and nonequivalent statements, and to recognize inductive and deductive arguments and supported and unsupported claims. Also tested is students' ability to recognize patterns and sequences of arguments, including their ability to see relationships of premises, subarguments, and subconclusions to the overall argument.

Overall results and results for the bottom 25%, middle 50%, and top 25% of total CAAP Critical Thinking scores for the local and normative groups are provided. This is to enable you to compare bottom, middle, and top students in the local cohort to the bottom, middle, and top students in the normative group. Any differences less than 5%, between 5% and 10%, and greater than 10% are considered negligible, moderate, and substantial, respectively.

Critical Thinking
Content Area: Evaluation of Arguments

Figure C-2: Average Percent of Correct Answers in Evaluation of Arguments



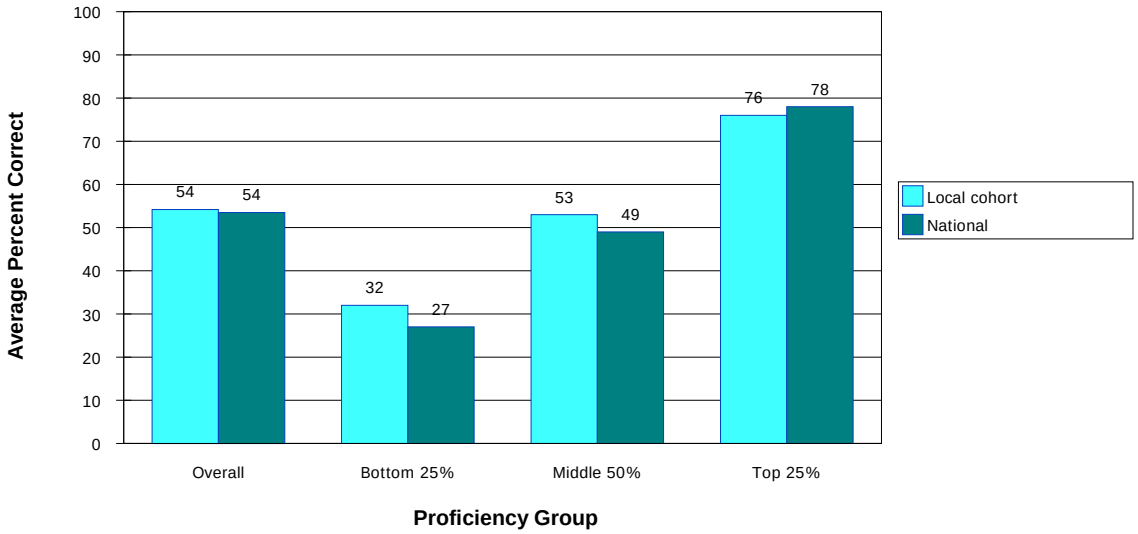
Interpretation Guide

The results in Figure C-2 are based on your students' responses to the Evaluation of Arguments items on the CAAP Critical Thinking test. There are six items in this content category, constituting approximately 19% of the Critical Thinking test. Items in this category assess the students' ability to evaluate information on the basis of its consistency, relevance, and accuracy, and to make judgments about its sufficiency. In addition, students are assessed in their ability to evaluate replies to arguments on the basis of their intent, appropriateness, and strength.

Overall results and results for the bottom 25%, middle 50%, and top 25% of total CAAP Critical Thinking scores for the local and normative groups are provided. This is to enable you to compare bottom, middle, and top students in the local cohort to the bottom, middle, and top students in the normative group. Any differences less than 5%, between 5% and 10%, and greater than 10% are considered negligible, moderate, and substantial, respectively.

Critical Thinking
Content Area: Extension of Arguments

Figure C-3: Average Percent of Correct Answers in Extension of Arguments



Interpretation Guide

The results in Figure C-3 are based on your students' responses to the Extension of Arguments items on the CAAP Critical Thinking test. There are six items in this content category, constituting approximately 19% of the Critical Thinking test. Items in this category assess students' skills in using given premises to reach related conclusions and in recognizing the scope of application of arguments. Students' ability to develop or recognize arguments that are based on analogies is also assessed. Some items in this category also assess students' understanding of how modifications to an argument can strengthen or weaken the argument or resolve conflicts within the argument.

Overall results and results for the bottom 25%, middle 50%, and top 25% of total CAAP Critical Thinking scores for the local and normative groups are provided. This is to enable you to compare bottom, middle, and top students in the local cohort to the bottom, middle, and top students in the normative group. Any differences less than 5%, between 5% and 10%, and greater than 10% are considered negligible, moderate, and substantial, respectively.