



## Office of Institutional Research

To: College Wide Assessment Team  
From: Darby Hiller  
Date: February 24, 2003  
Subject: CAAP Addendum: Item Analysis

There are three ways in which the CAAP test can help measure the general education outcome of critical thinking. First, extra item analysis can help link certain questions to the three levels of the critical thinking rubric. In this way, students' skill level (a.k.a answering a question correctly in certain skill sets) can be compared to the levels of the rubric. Second, students with greater exposure to the NMC curriculum can be compared with students with less exposure demonstrating the effectiveness of the curriculum. Third, students can be compared to other students from two-year colleges nation-wide with a standardized score. Here I provide the findings for the three different types of comparisons.

After speaking with the director of research at ACT about the data for our CAAP critical thinking test, we received all the data we need to do an extra item analysis. As mentioned in the CAAP after action report (dated January 6, 2003), there are three sets of questions in the critical thinking objective test. Seventeen questions out of 32 (53%) measure whether a student can identify the parts of an argument. Nine questions measure whether a student is able to evaluate an argument (28%). Six questions measure whether a student is able to extend an argument (19%). These three different skills roughly correspond to Level 1, 2, and 3 of our critical thinking rubric. The linkage table, previously presented in the CAAP after action report, is reprinted below.

<b>Level</b>	<b>Language in Critical Thinking Rubric</b>	<b>Language in CAAP Technical Handbook</b>
Level 1	Identifies arguments	Identifies elements of an argument
Level 2	Evaluates claims and develops conclusions	Evaluates an argument
Level 3	Demonstrates problem resolution	Extends an argument

We were able to obtain the raw data from ACT showing how students answered each question and whether the answer was right or wrong. I compiled descriptive statistics for the three sets of questions. Sixty seven percent (67%) of the test takers were able to identify the parts of an argument correctly as shown in Table 1. Fifty four percent (54%) of our students were able to correctly evaluate an argument (Table 2). Finally, 58% of our students were able to extend an argument

correctly (Table 3). (Note: these statistics assume these skills are exclusive of each other.)

<b>Table 1</b>	<b>Identifies elements of an argument</b>				
	Wrong Count	%	Right Count	%	Mean
q2v	23	19.2%	97	80.8%	0.81
q3v	30	25.0%	90	75.0%	0.75
q4v	38	31.7%	82	68.3%	0.68
q5v	29	24.2%	91	75.8%	0.76
q7v	72	60.0%	48	40.0%	0.40
q9v	27	22.5%	93	77.5%	0.78
q10v	26	21.7%	94	78.3%	0.78
q11v	16	13.3%	104	86.7%	0.87
q17v	19	15.8%	101	84.2%	0.84
q18v	25	20.8%	95	79.2%	0.79
q19v	21	17.5%	99	82.5%	0.83
q21v	46	38.3%	74	61.7%	0.62
q25v	34	28.3%	86	71.7%	0.72
q26v	44	36.7%	76	63.3%	0.63
q29v	53	44.2%	67	55.8%	0.56
q31v	82	68.3%	38	31.7%	0.32
q32v	86	71.7%	34	28.3%	0.28
<b>Sum of Right</b>			<b>1,369</b>		<b>0.67</b>

<b>Table 2</b>	<b>Evaluates an Argument</b>				
	Wrong Count	%	Right Count	%	Mean
q6v	33	27.5%	87	72.5%	0.73
q8v	46	38.3%	74	61.7%	0.62
q12v	29	24.2%	91	75.8%	0.76
q14v	49	40.8%	71	59.2%	0.59
q16v	71	59.2%	49	40.8%	0.41
q22v	64	53.3%	56	46.7%	0.47
q24v	76	63.3%	44	36.7%	0.37
q27v	54	45.0%	66	55.0%	0.55
q30v	78	65.0%	42	35.0%	0.35
<b>Sum of Right</b>			<b>580</b>		<b>0.54</b>

<b>Table 3</b>	<b>Extends an argument</b>				
	Wrong Count	%	Right Count	%	Mean
q1v	30	25.0%	90	75.0%	0.75
q13v	44	36.7%	76	63.3%	0.63
q15v	45	37.5%	75	62.5%	0.63
q20v	41	34.2%	79	65.8%	0.66
q23v	64	53.3%	56	46.7%	0.47
q28v	77	64.2%	43	35.8%	0.36
<b>Sum of Right</b>			<b>419</b>		<b>0.58</b>

The next set of statistics assumes that the skills are not exclusive of one another but build on each other. If a student was able to answer at least half of the questions in a certain skill set it was assumed that she mastered the skill. This is a liberal assumption. Only 2 people answered all the Level 1 questions right, 3 people answered all the Level 2 questions correctly, and 17 people answered all the Level 3 questions correctly. No one had a perfect score.

Of those students answering more than half of the Level 1 (identifying an argument) questions correctly (85%), 60% answered at least half of the Level 2 questions correctly and 72% answered at least half of the Level 3 questions correctly. For those able to answer at least half the Level 1 questions and the Level 2 questions correctly, 86% answered at least half the Level 3 questions correctly. For those able to answer at least half the Level 1 questions and the Level 3 questions correctly, 70% were able to answer at least half the Level 2 questions. It is clear from these statistics the weakest skill is in evaluating an argument and developing conclusions (Level 2).

For the second category of comparison, I divided the students into two groups: 1) those with 0-30 credits hours, and 2) those with 31 credits and above. Table 4 shows the distribution of those two groups.

<b>Table 4</b>	Frequency	Percent	Valid Percent
0-30 credits	70	58.3	60.3
31 credits and above	46	38.3	39.7
Total	116	96.7	100.0
Missing	4	3.3	
Total	120	100.0	

Students with greater exposure to the NMC curriculum did only slightly better than those with less exposure. Table 5 shows the percentage of students that answered at least half of the questions correctly in each skill set (skills are considered exclusive). “Freshmen” actually performed better than “Upperclassmen” in evaluating an argument (Level 2).

<b>Table 5</b>	Identify	Evaluate	Extend
0-30 credits	84.3%	54.3%	60.0%
31 credits and above	89.1%	45.7%	67.4%

These results do not indicate that students exposed to NMC curricula have gained significantly over their colleagues with fewer than 30 credits in their critical thinking skills.

Finally, as reported in the CAAP after action report, 60% of NMC’s test-takers scored higher than the national average for two-year colleges. Unfortunately, there are many other factors that could lead to this result not accounted for in the standardized scores. However, this is definitely a “feel good” statistic and something for NMC to be proud of.

I do not know if ACT will provide this level of detailed data in the future. But if so, it can help us alter the curriculum to target certain skills. Moreover, a longitudinal analysis of this data can demonstrate if changes to the curriculum are effective. If you have questions or comments about this addendum and the analyses please contact Darby Hiller, 995-1084 or [dhiller@nmc.edu](mailto:dhiller@nmc.edu).