

Cluster Three: The Natural World  
Cluster Assessment Results and Interpretation

*Fall 2007 Assessment Results*

Carol Barry  
Ross Markle  
Cliff Wigtil  
Donna L. Sundre  
Center for Assessment and Research Studies  
James Madison University  
November 12, 2007

## Cluster 3 Assessment Report Fall 2007

### **Executive Summary**

This report provides the first results for the ninth version of the Natural World (NW-9) assessment instrument. The instrument was carefully reviewed during the summer of 2007, and several new items were added. This is the same instrument that will be marketed to many external clients as the Quantitative Reasoning (QR) and Scientific Reasoning (SR) tests. The results obtained with fall 2007 new entering first-year students were overall quite good, and the Cluster 3 Committee hopes to administer this instrument for several years to come.

### **Results in Brief:**

- A total of 1,408 entering student tests were available for this analysis; 13 (<1%) were deleted due to invalid identification or insufficient responses.
- The NW-9 was administered in two formats: paper and pencil (N= 995; 71%) and via computer (N= 413; 29%).
- Overall, the test seems to be functioning quite well. The reliabilities observed for Quantitative Reasoning ( $\alpha = .64$ ) and the total test, which is how Scientific Reasoning is conceptualized at James Madison ( $\alpha = .78$ ) are adequate for program evaluation purposes and comparable to those observed for earlier entering first-year cohorts at JMU, which tend to be lower than those obtained with sophomore samples. Though close, the reliability estimates obtained for computer-based administration were slightly lower than those for the paper and pencil format. This will be monitored.
- Entering student performances for QR were at about 62% correct and 64% correct for the total SR score. Cluster Three faculty members are in the process of creating community expectations for student performances through the Angoff procedure. Faculty members will independently provide judgments on how student who had successfully completed cluster requirements would perform on each item. Supplemental reports will compare these ratings with the performances of entering students.
- Students entering the university of Advanced Placement (AP) or Dual Enrollment college credits tended to score higher than students entering without relevant college-level credits. Increases of 4-10% in QR and SR scores were observed.
- Additional analyses produced encouraging results that provide evidence that both AP and Dual Enrollment credits contribute to higher scores on both QR and SR.

## Fall 2007, Cluster 3 Assessment Report Outline

### **Introduction**

This report describes the results of the Fall 2007 Assessment Day administration of the Natural World test version 9 (NW-9). The NW-9 assessment was designed by faculty and assessment specialists at JMU to measure the objectives of Cluster 3 (CL3), the Natural World segment of JMU's general education program. Fall 2007 was the first time the NW-9 was administered to first-year JMU students. This report contains the results of statistical analyses that describe the relationship between the CL3 program and NW-9 scores. These analyses lend support to the perception that students are learning as a result of participation in CL3 program.

### **Sample and Data collection procedures**

1,421 students participated in the Cluster 3 assessment during the Fall 2007 assessment day. The responses from 13 students were deleted because these students either omitted 50% or more of the items or did not provide a valid JMU id. The data collected in Fall 2007 is the first cycle of students to complete the Natural world test version 9 (NW-9) as first-year students. These students will complete the test again as students with 40-75 credits. The test was administered in two formats; 995 students completed the test using paper and pencil and 413 students completed the test on a computer.

### **Assessment Strategies Overview**

Assessment results are examined in multiple ways. First, scores on each of the subscales are reported for the sample overall. Scores are then analyzed by the number of cluster, AP and transfer courses taken, the relationship with cluster, AP and transfer/dual enrollment credit hours, and finally the relationship with AP test scores. Currently, scores are not examined in relation to a standard of proficiency because standards for performance on the NW-8 are still being established.

### **Reliability and Scores**

Table 1 on page 4 shows the learning objectives assessed by the NW-9 along with the average number of items correct and percent correct scores for the overall sample. Reliability information (Cronbach's coefficient alpha  $\alpha$ ) is provided for each objectives scale. Reliability values range from 0 to 1 and indicate the degree to which the scores are free from random error. In general, reliabilities above .70 are considered adequate for program evaluation or research. Scores for scales with reliabilities lower than .60 should be interpreted with caution. Reliability and scale scores are reported separately for computer-based and paper-based administrations in Appendix I. Additionally, a graph and table of the p-values (i.e., or proportion of the sample who answered each item correctly) for all items is presented in Appendix II.

### **Comparability of scores**

Because this was the first administration of the NW-9, raw scores reported in Table 1 are not directly comparable to previous cohorts of first-year students.

**Table 1. Nw-9 Scores for Fall 2007 (N = 1,408)**

<b>Cluster 3 - Learning Objectives</b>	<b>Item(s) Assessing Objective</b>	
1. Describe the methods of inquiry that lead to mathematical truth and scientific knowledge and be able to distinguish science from pseudo-science.	2, 5, 9, 14, 18, 28, 38-41, 55-57 (13 items; 19.7% of test)	M = 9.04 (69.53% correct) SD = 1.92 $\alpha = .43$
2. Use theories and models as unifying principles that help us understand natural phenomena and make predictions.	17, 20, 22, 27, 64-66 (7 items; 10.6% of test)	M = 3.91 (55.89% correct) SD = 1.36 $\alpha = .20$
3. Recognize the interdependence of applied research, basic research, and technology, and how they affect society.	1, 15, 16, 43-46 (7 items; 10.6% of test)	M = 4.09 (58.47% correct) SD = 1.64 $\alpha = .47$
4. Illustrate the interdependence between developments in science and social and ethical issues.	2, 19, 24-26, 29, 55-57 (9 items; 13.6% of test)	M = 5.50 (61.14% correct) SD = 1.44 $\alpha = .25$
5. Use graphical, symbolic, and numerical methods to analyze, organize, and interpret natural phenomenon.	4, 7, 8, 10-13, 21, 30-33, 51-53, 58-63 (21 items; 31.8% of test)	M = 12.53 (59.67% correct) SD = 3.12 $\alpha = .58$
6. Discriminate between association and causation, and identify the types of evidence used to establish causation	3, 34-37, 53, 60-63 (10 items; 15.2% of test)	M = 5.37 (53.71% correct) SD = 1.84 $\alpha = .45$
7. Formulate hypotheses, identify relevant variables, and design experiments to test hypotheses.	5, 6, 9-13, 18, 23, 28, 41, 42, 47-50, 54, 59, 60, 62, 63 (21 items; 31.8% of test)	M = 14.20 (67.60% correct) SD = 2.99 $\alpha = .59$
8. Evaluate the credibility, use, and misuse of scientific and mathematical information in scientific developments and public-policy issues.	2, 14, 24-26, 29, 38-40, 60-63 (13 items; 19.7% of test)	M = 6.99 (53.76% correct) SD = 1.89 $\alpha = .32$
Quantitative Reasoning	3, 4, 7, 8, 10-13, 21, 30-37, 51-53, 58-63 (26 items; 39.4% of test)	M = 16.09 (61.88% correct) SD = 3.73 $\alpha = .64$
Total Test	1-66	M = 42.18 (63.91% correct) SD = 7.46 $\alpha = .78$

## **I. Average Score by Number of Credit Hours**

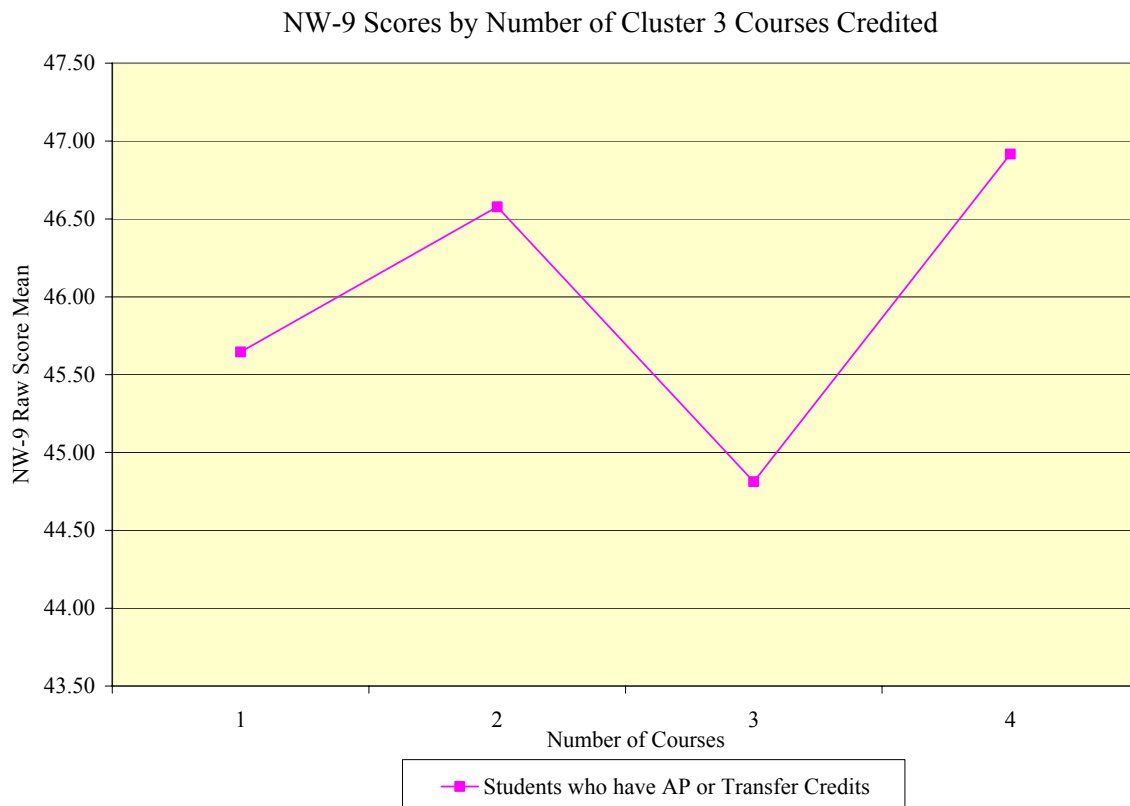
*Do students learn or develop more if they have taken more cluster-related courses?*

Table 2 presents the average scores by number of credit hours completed. Students who have completed more courses in the cluster should score or perform higher than students who have taken fewer courses in the cluster. Students may have received cluster course credit, through transfer/dual enrollment, AP tests, or even by taking courses at JMU. A visual of these results is presented in Figure 1.

**Table 2. Scores by Number of JMU Courses Taken**

# of CL 3 credits	N	Raw score Mean	Raw score SD	Percent correct Mean	Percent correct SD
<i>Students who have NO Cluster 3 Credits</i>					
All with no credits	1173	41.45	7.32	62.80	11.09
<i>Students who have AP or Transfer Credits</i>					
1	161	45.65	7.18	69.16	10.87
2	45	46.58	7.49	70.57	11.35
3	16	44.81	6.87	67.90	10.41
4 or more	12	46.92	5.99	71.09	9.08
<i>All students with Credits in Sample</i>					
1	162	45.62	7.16	69.13	10.85
2	45	46.58	7.49	70.57	11.35
3	16	44.81	6.87	67.90	10.41
4 or more	12	46.92	5.99	71.09	9.08

**Figure 1.**



## II. Relationship between NW-9 Scores and Number of Credit Hours

### Completed

*What is the relationship between number of cluster 3 credit hours completed and outcomes?*

Correlations between NW-9 scores and number of credit hours completed are presented in Table 3. A positive relationship should exist between number of cluster 3 hours completed and NW-9 scores if the content covered in the courses matches that covered on the NW test.

Correlations can range from -1 to 1. Correlations close to 0 indicate no relationship, while correlations closer to 1 indicate a very high relationship between test scores and credit hours (negative numbers would indicate those who scored high on the test had fewer credit hours).

**Table 3. Correlations between Number of Credit Hours Completed and NW-9 Raw Scores**

Credit hours	N	NW correlation	NW $r^2$	QR correlation	QR $r^2$
<i>No JMU Credit Hours</i>					
Transfer Credit Hours	85	0.21	0.04	0.21	0.04
AP Credit Hours	152	0.26	0.07	0.28	0.08

The graphs below illustrate instability in the correlations due to sampling error (Figures 2 and 3). The error bands show the 95% confidence band around the correlation estimates. Correlations with overlapping bands are not statistically different (at  $p < .05$ ).

**Figure 2.**

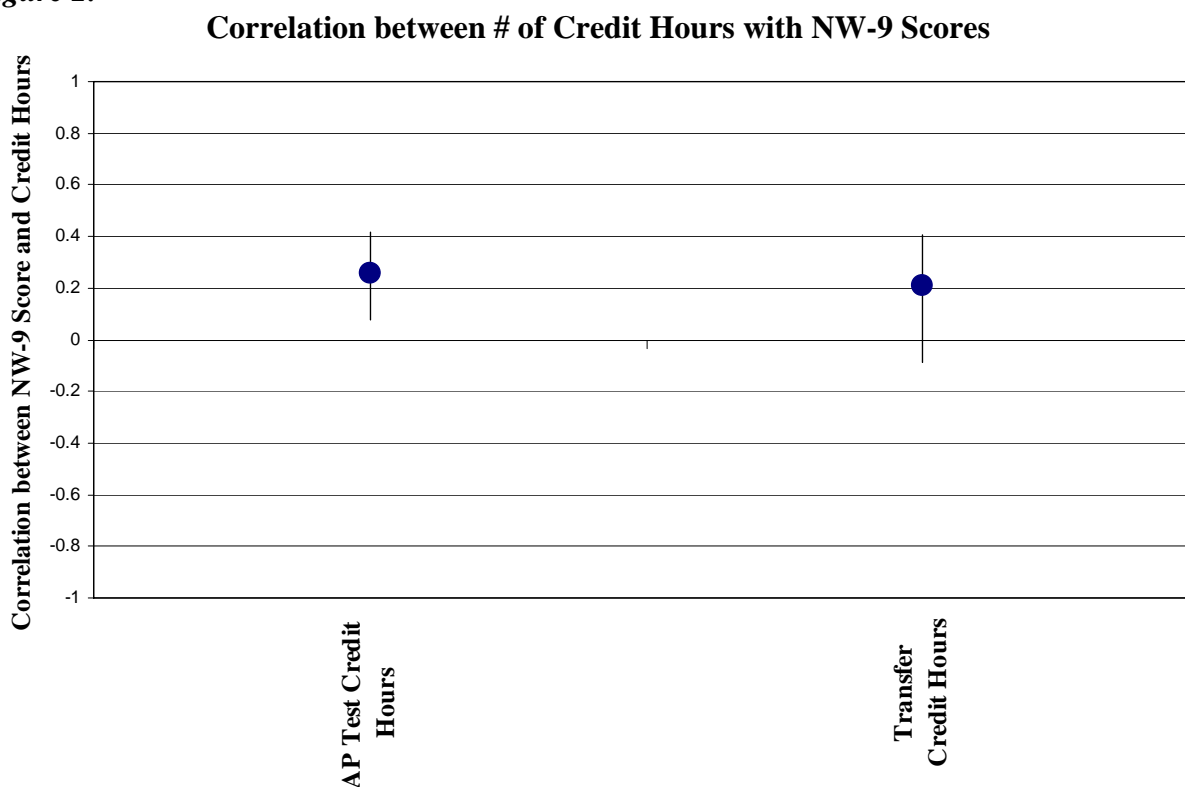
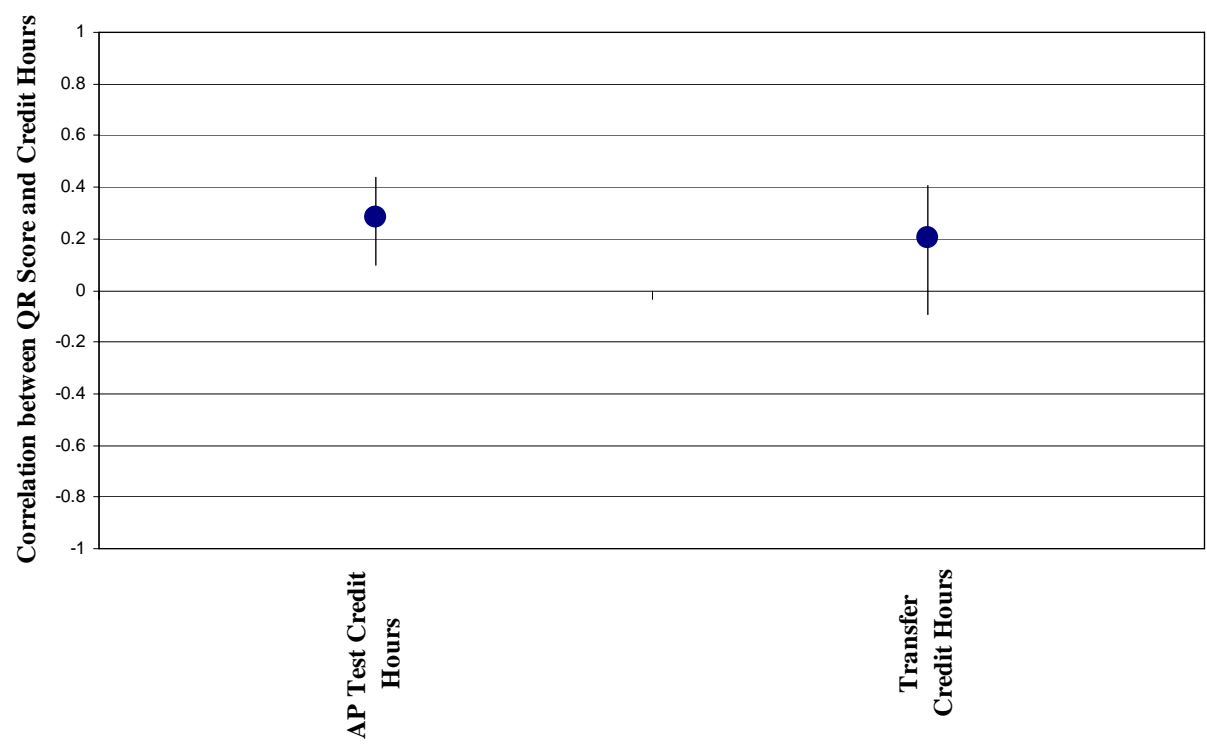


Figure 3.

### Correlation between # of Credit Hours and QR Scores



### III. Relationship between NW-9 Scores and AP Test Score

*What is the relationship between average AP test scores and outcomes?* Correlations between NW-9 scores and AP test scores are reported in Table 4. Given that students are given cluster course credit if they obtain a certain score or higher on the AP test, a positive relationship should exist if higher scores on AP tests indicate higher levels of mastery of the material. In this situation, AP test scores serve as a proxy for grades.

Again, correlations can range from -1 to 1. Correlations close to 0 indicate no relationship, while correlations closer to 1 indicate a very high relationship between test scores and AP test scores (negative numbers would indicate those who scored high on the test had lower AP test scores).

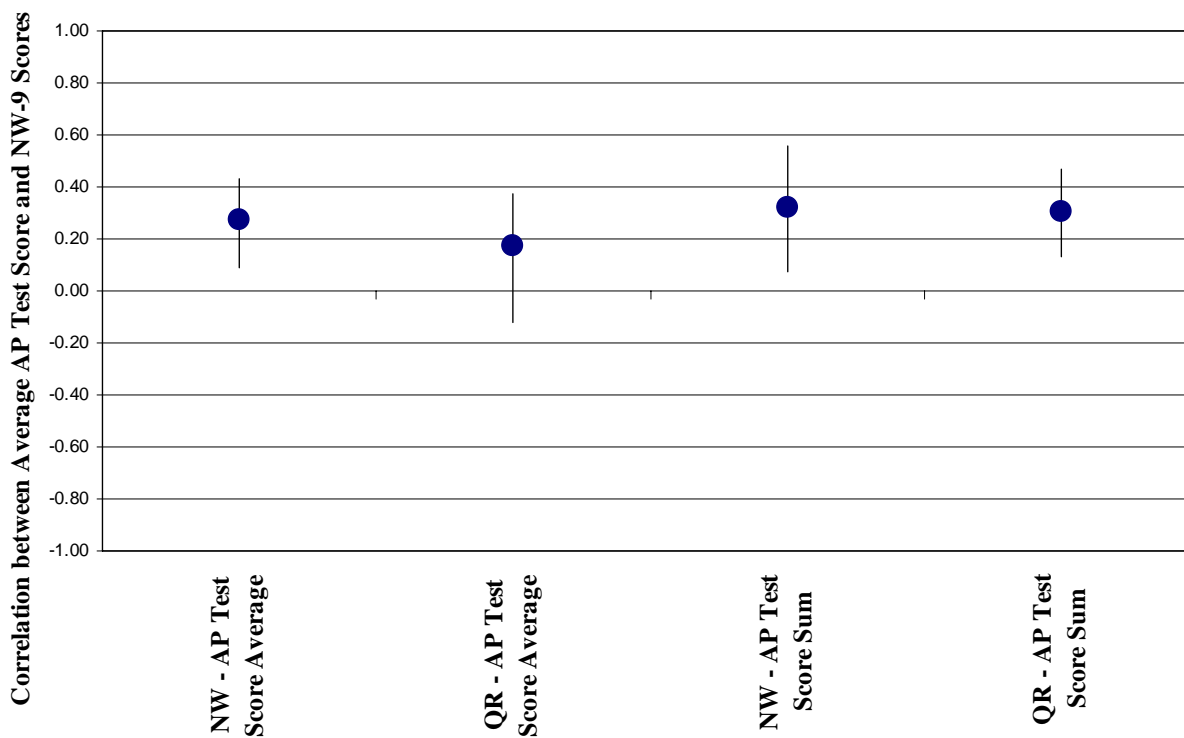
**Table 4. Correlations between Average AP test scores and NW-9 Raw Scores**

	N	NW correlation	NW $r^2$	QR correlation	QR $r^2$
AP Test Score - Average	152	0.27	0.07	0.17	0.03
AP Test Score - Sum	152	0.32	0.10	0.30	0.09

The Figure 4 below illustrates instability in the correlations due to sampling error. The error bands show the 95% confidence band around the correlation estimates. Correlations with overlapping bands are not statistically different (at  $p < .05$ ).

**Figure 4.**

#### **Correlation between AP Test scores and NW and QR Scores**



#### **IV. Meeting a Standard**

*Do students meet faculty expectations?* Students who have cluster credits when they arrive at JMU should be closer to meeting the competency or academic standard than those with no cluster credits.

This assessment strategy does not currently apply to Cluster 3 because a standard for the NW-9 has yet to be established.

## Appendix I

## Fall 2007 paper-based Natural World 9 assessment results (N = 995)

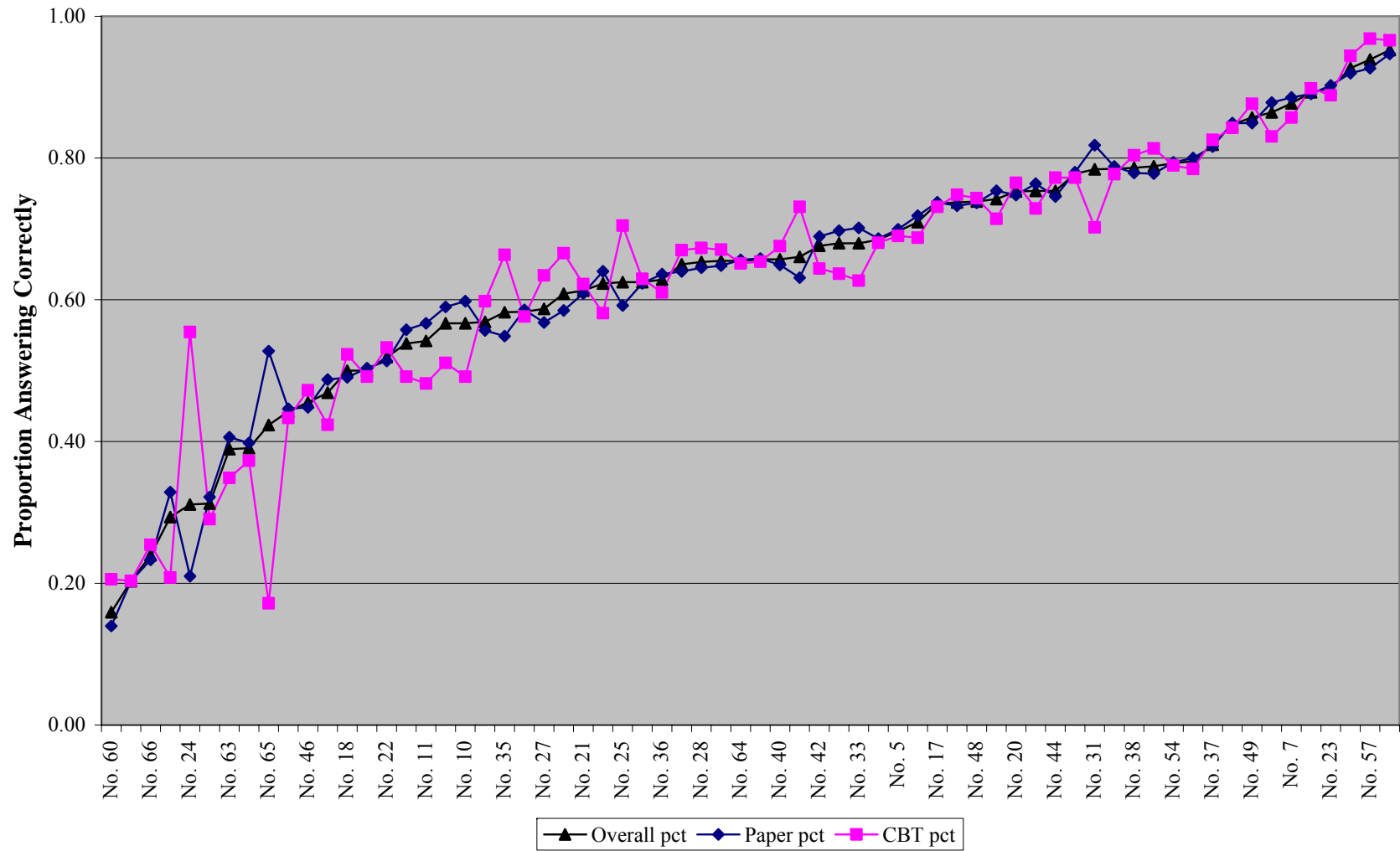
Cluster 3 - Learning Objectives	Item(s) Assessing Objective	
1. Describe the methods of inquiry that lead to mathematical truth and scientific knowledge and be able to distinguish science from pseudo-science.	2, 5, 9, 14, 18, 28, 38-41, 55-57 (13 items; 19.7% of test)	M = 9.01 (69.34% correct) SD = 1.95 $\alpha = .45$
2. Use theories and models as unifying principles that help us understand natural phenomena and make predictions.	17, 20, 22, 27, 64-66 (7 items; 10.6% of test)	M = 3.98 (56.91% correct) SD = 1.39 $\alpha = .25$
3. Recognize the interdependence of applied research, basic research, and technology, and how they affect society.	1, 15, 16, 43-46 (7 items; 10.6% of test)	M = 4.10 (58.52% correct) SD = 1.68 $\alpha = .51$
4. Illustrate the interdependence between developments in science and social and ethical issues.	2, 19, 24-26, 29, 55-57 (9 items; 13.6% of test)	M = 5.37 (59.66% correct) SD = 1.43 $\alpha = .25$
5. Use graphical, symbolic, and numerical methods to analyze, organize, and interpret natural phenomenon.	4, 7, 8, 10-13, 21, 30-33, 51-53, 58-63 (21 items; 31.8% of test)	M = 12.72 (60.59% correct) SD = 3.18 $\alpha = .60$
6. Discriminate between association and causation, and identify the types of evidence used to establish causation	3, 34-37, 53, 60-63 (10 items; 15.2% of test)	M = 5.37 (53.70% correct) SD = 1.84 $\alpha = .45$
7. Formulate hypotheses, identify relevant variables, and design experiments to test hypotheses.	5, 6, 9-13, 18, 23, 28, 41, 42, 47-50, 54, 59, 60, 62, 63 (21 items; 31.8% of test)	M = 14.25 (67.84% correct) SD = 3.09 $\alpha = .62$
8. Evaluate the credibility, use, and misuse of scientific and mathematical information in scientific developments and public-policy issues.	2, 14, 24-26, 29, 38-40, 60-63 (13 items; 19.7% of test)	M = 6.81 (52.39% correct) SD = 1.83 $\alpha = .28$
Quantitative Reasoning	3, 4, 7, 8, 10-13, 21, 30-37, 51-53, 58-63 (26 items; 39.4% of test)	M = 16.27 (62.58% correct) SD = 3.78 $\alpha = .65$
Total Test	1-66	M = 42.29 (64.07% correct) SD = 7.74 $\alpha = .79$

**Fall 2007 computer-based Natural World 9 assessment results (N = 413)**

<b>Cluster 3 - Learning Objectives</b>	<b>Item(s) Assessing Objective</b>	
1. Describe the methods of inquiry that lead to mathematical truth and scientific knowledge and be able to distinguish science from pseudo-science.	2, 5, 9, 14, 18, 28, 38-41, 55-57 (13 items; 19.7% of test)	M = 9.10 (69.99% correct) SD = 1.84 $\alpha = .39$
2. Use theories and models as unifying principles that help us understand natural phenomena and make predictions.	17, 20, 22, 27, 64-66 (7 items; 10.6% of test)	M = 3.74 (53.44% correct) SD = 1.26 $\alpha = .12$
3. Recognize the interdependence of applied research, basic research, and technology, and how they affect society.	1, 15, 16, 43-46 (7 items; 10.6% of test)	M = 4.08 (58.35% correct) SD = 1.54 $\alpha = .39$
4. Illustrate the interdependence between developments in science and social and ethical issues.	2, 19, 24-26, 29, 55-57 (9 items; 13.6% of test)	M = 5.83 (64.73% correct) SD = 1.41 $\alpha = .22$
5. Use graphical, symbolic, and numerical methods to analyze, organize, and interpret natural phenomenon.	4, 7, 8, 10-13, 21, 30-33, 51-53, 58-63 (21 items; 31.8% of test)	M = 12.07 (57.47% correct) SD = 2.94 $\alpha = .51$
6. Discriminate between association and causation, and identify the types of evidence used to establish causation	3, 34-37, 53, 60-63 (10 items; 15.2% of test)	M = 5.38 (53.75% correct) SD = 1.84 $\alpha = .47$
7. Formulate hypotheses, identify relevant variables, and design experiments to test hypotheses.	5, 6, 9-13, 18, 23, 28, 41, 42, 47-50, 54, 59, 60, 62, 63 (21 items; 31.8% of test)	M = 14.08 (67.02% correct) SD = 2.75 $\alpha = .50$
8. Evaluate the credibility, use, and misuse of scientific and mathematical information in scientific developments and public-policy issues.	2, 14, 24-26, 29, 38-40, 60-63 (13 items; 19.7% of test)	M = 7.42 (57.07% correct) SD = 1.95 $\alpha = .38$
Quantitative Reasoning	3, 4, 7, 8, 10-13, 21, 30-37, 51-53, 58-63 (26 items; 39.4% of test)	M = 15.66 (60.23% correct) SD = 3.60 $\alpha = .60$
Total Test	1-66	M = 41.92 (63.51% correct) SD = 6.76 $\alpha = .73$

Appendix II

Proportion of Students Answering Each Item Correctly



**Proportion of Students Answering Each Item Correct (Ascending Order)**

<b>Item no.:</b>	<b>Computer-Based pct</b>	<b>Paper pct</b>	<b>Overall pct</b>
No. 60	0.21	0.14	0.16
No. 2	0.20	0.20	0.20
No. 66	0.25	0.23	0.24
No. 53	0.21	0.33	0.29
No. 24	0.55	0.21	0.31
No. 61	0.29	0.32	0.31
No. 63	0.35	0.41	0.39
No. 19	0.37	0.40	0.39
No. 65	0.17	0.53	0.42
No. 43	0.43	0.45	0.44
No. 46	0.47	0.45	0.46
No. 15	0.42	0.49	0.47
No. 18	0.52	0.49	0.50
No. 39	0.49	0.50	0.50
No. 22	0.53	0.51	0.52
No. 32	0.49	0.56	0.54
No. 11	0.48	0.57	0.54
No. 4	0.51	0.59	0.57
No. 10	0.49	0.60	0.57
No. 58	0.60	0.56	0.57
No. 35	0.66	0.55	0.58
No. 6	0.58	0.59	0.58
No. 27	0.63	0.57	0.59
No. 1	0.67	0.58	0.61
No. 21	0.62	0.61	0.61
No. 41	0.58	0.64	0.62
No. 25	0.70	0.59	0.63
No. 51	0.63	0.62	0.63
No. 36	0.61	0.64	0.63
No. 13	0.67	0.64	0.65
No. 28	0.67	0.65	0.65
No. 26	0.67	0.65	0.65
No. 64	0.65	0.66	0.65

<b>Item no.:</b>	<b>Computer-Based pct</b>	<b>Paper pct</b>	<b>Overall pct</b>
No. 30	0.65	0.66	0.66
No. 40	0.68	0.65	0.66
No. 62	0.73	0.63	0.66
No. 42	0.64	0.69	0.68
No. 16	0.64	0.70	0.68
No. 33	0.63	0.70	0.68
No. 45	0.68	0.69	0.68
No. 5	0.69	0.70	0.70
No. 52	0.69	0.72	0.71
No. 17	0.73	0.74	0.74
No. 55	0.75	0.73	0.74
No. 48	0.74	0.74	0.74
No. 3	0.71	0.75	0.74
No. 20	0.77	0.75	0.75
No. 12	0.73	0.76	0.75
No. 44	0.77	0.75	0.75
No. 29	0.77	0.78	0.78
No. 31	0.70	0.82	0.78
No. 34	0.78	0.79	0.78
No. 38	0.80	0.78	0.79
No. 59	0.81	0.78	0.79
No. 54	0.79	0.79	0.79
No. 8	0.78	0.80	0.80
No. 37	0.83	0.82	0.82
No. 50	0.84	0.85	0.85
No. 49	0.88	0.85	0.86
No. 56	0.83	0.88	0.86
No. 7	0.86	0.89	0.88
No. 47	0.90	0.89	0.89
No. 23	0.89	0.90	0.90
No. 9	0.94	0.92	0.93
No. 57	0.97	0.93	0.94
No. 14	0.97	0.95	0.95