

**Cluster Three: The Natural World
Cluster Assessment Results and Interpretation**

Spring 2002 Assessment

Prepared by the Cluster Three Committee

June 2003

Based on a report prepared by:

Donna L. Sunde,

Kit Murphy,

Keston Fulcher,

Robin RiCharde, and

Susan Davis

**Center for Assessment and Research Studies
James Madison University**

Executive Summary

This document presents the recent results of assessment in Cluster Three and the Cluster Three Committee's interpretation of those results. Spring 2002 is the third administration of the Natural World Test (NW-5) to a sophomore/junior sample of students. The instrument provides two scores: Total and a Quantitative Reasoning subscore. The following findings highlight this report:

- The current version of the assessment instrument for Cluster Three (NW-5) is fairly reliable (Cronbach's $\alpha = .77$). The test is less reliable for entering students (Cronbach's $\alpha = .65$). Reliability for individual objectives remains problematic. The reliability estimates appear stable over time.
- For sophomores and juniors with 45-70 credit hours taking the exam in Spring 2002, the average overall score was 58.1%. These results are not significantly different than the two earlier sophomore samples; however, all three sophomore samples scored significantly higher than both of the two earlier entering student samples.
- As with results from earlier years, scores on the Total Natural World test and the Quantitative Reasoning subscore increased significantly with increasing numbers of Cluster Three courses completed.
- Multiple regression analyses reveal that Cluster Three related AP and JMU courses both significantly predict Total and Quantitative Reasoning scores. In contrast, Cluster related transfer credits do not. Of additional interest, cumulative credit hours across subject areas negatively predict Natural World Total and Quantitative Reasoning scores. In other words, test scores are not enhanced via academic maturation via course taking.
- Over 90% of correlations between course grades and scores on NW-5 and Quantitative Reasoning were positive.
- A repeated measures multivariate analysis of students completing their Cluster Three coursework revealed significantly higher overall posttest scores with no interaction by package. In other words, for package completers, Cluster Three performances were significantly higher on their sophomore posttest than their first-year entering pretest, but no significant differences were observed when the packages students completed were compared. This is a very powerful finding.
- Taken together, these results indicate that Cluster Three courses increase student mastery of Cluster objectives. However, several issues remain, and this report presents several recommendations for improving the ability of Cluster Three assessment to help Cluster faculty increase student mastery of objectives.

I. Introduction

The Cluster Three Committee maintains that the main function of assessment is to help faculty understand the impact of their teaching on the learning of students in the Cluster. Our ultimate goal is to use assessment to improve the courses we teach, and thereby increase student mastery of Cluster objectives. Although constituents outside the Cluster and the University may also use assessment results for a variety of reasons, this report reflects our desire to use assessment to help us improve our teaching, and as a result, student learning.

One key question of assessment is whether Cluster courses help students master the objectives of the Cluster. The best test of this hypothesis would be to conduct an experiment in which students are randomly allocated to two groups: those who take Cluster Three courses, and those who do not. We would predict that those students taking Cluster Three courses should, when tested with an appropriate assessment instrument, outperform students not taking Cluster Three courses. Clearly, such a study is impossible given that all students must complete Cluster Three to graduate.

Instead, we must adopt a nonexperimental approach. It is possible to compare students who have and have not taken Cluster Three courses, but these students may differ in variables other than whether they have completed the Cluster. For example, students who have not taken Cluster Three courses are likely to be freshman and hence at a different stage of intellectual maturity than students who have completed the Cluster. The strategy we adopt to deal with this interpretational problem is to conduct several analyses, each of which tests a different prediction of the hypothesis that Cluster Three courses help students master the objectives of the Cluster. Results addressing each prediction are subject to alternative interpretations, but if all predictions are supported, then our confidence in the hypothesis increases. The five predictions tested in this report are:

1. Comparison of First-year with Sophomore Students: Students with sophomore level status should outperform entering students on the Natural World and the Quantitative Reasoning assessment test.
2. Effect of Number of Courses: Assessment scores should increase with increasing number of courses completed.
3. Correlations Between Scores and Course Grades: There should be a positive correlation between test scores and the grades received by students in Cluster Three courses.
4. Package Completion Comparison: Students who have completed coursework in the Cluster should score higher than students who have not.
5. Pre/Post Comparison: Students completing Cluster Three courses should score higher on the sophomore post-test than they did as entering first-year students.

In addition to testing these predictions for the Cluster as a whole, we present assessment information for individual packages and courses to assist faculty in evaluating their own packages and courses.

II. Assessment Instrument

A. Nature and administration of instrument

The Natural World assessment instrument consists of 50 objective answer questions, each of which is worth 2 points. Students have one hour to complete the 50 questions. This exam is administered to randomly chosen freshman during orientation and to randomly chosen individuals with 41-70 credit hours (sophomores and juniors) in February. Most of the results contained in this report concern the Spring 20021 test, which tests sophomores and juniors. Data from other years and groups are included as appropriate.

B. Versions of the test

There have been five versions of the Natural World objective test. The current form, NW-5, has 50 questions designed to measure learning in 12 of the 17 objectives in Cluster Three.

C. Reliability of test scores

Reliability, as estimated using Cronbach's alpha (α), measures the degree to which the items in a test measure the same thing. More technically, it is related to the correlations between performances on items. High reliability means that correlations among items are high and hence the items are measuring the same underlying construct. Alphas of 0.6 are considered acceptable, and alphas of 0.8 or higher are considered quite good. If reliability falls below 0.6, then the items on the test are measuring different aspects of knowledge, and the overall test score is not a very meaningful measure of what is being tested.

The reliability of the Natural World test has steadily improved with each revision (Table 1). Due to the summer 1999 efforts, two new test forms (NW-3 and NW-4) were piloted in fall 1999. The best items from those administrations were selected to form NW-5. As can be seen, scores generated using the current Natural World test show the best reliability to date, ranging from .65-.67 for entering students and from .75-.77 for sophomore/junior samples. These reliability estimates appear quite stable.

Table 1. Reliability (α = Cronbach's alpha) of the Natural World assessment instrument

Administration date: Form	First year students	Sophomores-Juniors
Spring 1998: NW-2		$\alpha = .61$
Fall 1998: NW-2 Spring 1999: NW-2	$\alpha = .49$	$\alpha = .69$
Fall 1999: NW-3 & 4 Spring 2000: NW-5	$\alpha = .58$	$\alpha = .75$
Fall 2000: NW-5 Spring 2001: NW-5	$\alpha = .65$	$\alpha = .75$
Fall 2001: NW-5 Spring 2002: NW-5	$\alpha = .67$	$\alpha = .77$

III. Results, Interpretations, and Additional Analyses Needed

In this section, we present six major analyses that address the effectiveness of Cluster Three courses. We present the results of the analyses, provide our interpretation of these results, and suggest additional analyses that might help us better understand the results.

A. Performance on Individual Objectives

Results

Table 2 (pp. 6-8) provides summaries of Spring 2002 student performance on the Natural World assessment test for the 979 second-semester sophomores and first-semester juniors who were randomly selected to take the instrument. The Quantitative Reasoning subscore includes those items that address the ability of students to reason quantitatively. Combining the items that address Cluster Three objectives (9, 10, 11, 13, 14, and 16) associated with quantitative literacy and reasoning generated this score.

The percent correct for each of the objectives assessed ranged from a low of 40% to a high of 70% correct. The average overall score was 58.1%. Reliabilities for individual objectives were fairly low. The reliability estimate for the Quantitative Reasoning subscore (0.64) and Learning Goals I and II (0.60 and 0.66, respectively) were moderate, but the reliability for Learning Goal III was low (0.44). Review of the table also indicates that several objectives are not well represented on the NW-5. More work remains to be conducted in the development of additional items to better represent all of the Cluster Three course objectives.

Table 2. Performance by Spring 2002 students on each objective (n=979).
M = mean, s = standard deviation, α = reliability

LEARNING GOAL I: To understand the role of mathematics and science in interpreting the natural world and to understand the importance of these disciplines in society.

Objectives	Test Items Assessing Each Objective
1. Identify the characteristics that distinguish mathematics and science from each other and from other ways of obtaining knowledge.	Items 11,18, 23, 24 (4 items; 8% of the total test) M = 5.6 (70% correct) s = 1.8 α = .23
2. Demonstrate an understanding of the nature of mathematical truth and scientific knowledge.	Items 7, 27, 39, 40, 45 (5 items; 10% of the total test) M = 6.4 (64% correct) s = 2.0 α = .26
3. Demonstrate an understanding of the role of theories in science as unifying principles that explain observations and make predictions.	Items 1, 14, 30, 33 (4 items; 8% of the total test) M = 4.6 (58% correct) s = 2.0 α = .19
4. Demonstrate an understanding of the role of aesthetics in mathematics.	No items
5. Demonstrate an understanding of the interdependence of applied research, basic research, and technology, and how each can benefit society.	Items 3-5, 12, 15 (5 items; 10% of the total test) M = 6.7 (67% correct) s = 2.7 α = .51
6. Illustrate how developments in mathematics and science can raise important ethical issues.	No items
7. Illustrate the interdependence of social forces and the fields of mathematics and science.	No items
Total for Learning Goal I	18 items (36% of total test) M = 23.3 (65% correct) s = 5.5 α = .60

LEARNING GOAL II: To apply mathematical and scientific concepts to understand the natural world.

Objectives	Test Items Assessing Each Objective
8. Design and execute experiments to solve problems or test hypotheses.	Item 46 (1 item; 2% of the total test) M = 0.8 (40% correct) s = .97
9. Obtain, organize, analyze, interpret, and present data.	Items 34-38 (5 items; 10% of the total test) M = 5.3 (53% correct) s = 2.4 $\alpha = .32$
10. Interpret data and statistical arguments presented in tables and graphical displays.	Items 8-10, 16, 17 (5 items; 10% of the total test) M = 5.0 (50% correct) s = 2.6 $\alpha = .43$
11. Discriminate between association and causation, and identify the types of evidence used to establish causation.	Items 13, 41-44 (5 items; 10% of the total test) M = 5.9 (59% correct) s = 2.3 $\alpha = .32$
12. Formulate an hypothesis and identify relevant variables necessary to test that hypothesis.	No items
13. Use mathematics as an abstract language to analyze natural phenomena.	Items 2, 6, 19 (3 items; 6% of the total test) M = 3.9 (65% correct) s = 1.4 $\alpha = .08$
14. Use graphical, symbolic, and numerical methods to make mathematical and statistical arguments.	Items 20-22, 25, 26 (5 items; 10% of the total test) M = 5.8 (58% correct) s = 2.2 $\alpha = .21$
15. Demonstrate an understanding of theories and models that describe natural phenomena.	Items 47 – 50 (4 items; 8% of the total test) M = 4.0 (50% correct) s = 2.1 $\alpha = .46$
Total for Learning Goal II	28 items (56% of total test) M = 30.8 (55% correct) s = 7.8 $\alpha = .66$

LEARNING GOAL III: To evaluate mathematical and scientific arguments at a level commonly encountered by informed citizens.

Objectives	Test Items Assessing Each Objective
16. Evaluate the use of scientific, mathematical and statistical arguments in the analysis of public policy issues.	Items 28, 29, 31, 32 (4 items; 8% of the total test) M = 4.0 (50% correct) s = 2.3 $\alpha = .44$
17. Locate, evaluate, and select reliable resources to learn about scientific developments.	No Items
Total for Learning Goal III	4 items (8% of total test) M = 4.0 (50% correct) s = 2.3 $\alpha = .44$

QUANTITATIVE REASONING AND OVERALL TEST SCORES

Objectives	Test Items Assessing Each Objective
Quantitative Reasoning	Items 2, 6, 8-10, 13, 16, 17, 19-22, 25, 26, 28, 29, 31, 32, 34-38, 41-44 (27 items: 54% of the total test) M=29.9 (55% correct) s = 7.7 $\alpha = .64$
<i>Total for Entire Examination Spring 2002 n = 801 sophomores and juniors</i>	50 items (100% of test) M =58.1 (58.1% correct) s = 12.9 $\alpha = .77$

Interpretation

The averages for the overall test (58.1%) and the quantitative reasoning score (55%) are not especially high. However, two factors suggest that it may be premature to conclude from these scores that Cluster Three students have not mastered the Cluster objectives. First, there are a significant number of students contributing to the average who either have not completed their Cluster Three requirements or have transfer or AP credit for one or more Cluster Three courses, so this average is not the best indication of the performance of students completing the Cluster. Second, Cluster Three faculty have not yet determined what percentage score constitutes an acceptable score; such a process (e.g., a “bookmarking” process) may indicate that that observed average score is acceptable, given the difficulty of the test.

Ideally, the performance of students with respect to individual objectives could assist faculty in Cluster Three courses in altering their courses to improve performance with respect to those objectives. Theoretically, scores for each objective could be calculated for individual packages, and faculty in those packages could use scores for their package to inform them about how well objectives are mastered by students in the package. However, the low reliability for assessment of individual objectives (see Table 2) precludes use of these scores for this purpose; therefore, all analyses presented below will involve only the Total Test score and the Quantitative Reasoning subscore.

Additional Analyses

To understand better the level of performance of students who have completed their Cluster Three requirements, the scores for these students will be reported separately. Table 3 (pp. 10-12) provides the parallel results for 227 sophomore/junior students identified as having completed their Cluster Three coursework. Results for this subset of participants reveal very slight increases in overall scores by cluster objective, learning goals, the Quantitative Reasoning subscore, and the Total test score. It should be noted that many students included in the overall analyses have some Cluster Three courses completed. Additional analyses will shed more light on the impact of Cluster Three course completion on Natural World test performances.

Table 3. Performance by Spring 2002 package completers on each objective (n=227).
M = mean, s = standard deviation, α = reliability

LEARNING GOAL I: To understand the role of mathematics and science in interpreting the natural world and to understand the importance of these disciplines in society.

Objectives	Test Items Assessing Each Objective
1. Identify the characteristics that distinguish mathematics and science from each other and from other ways of obtaining knowledge.	Items 11,18, 23, 24 (4 items; 8% of the total test) M = 5.8 (73% correct) s = 1.8 α = .24
2. Demonstrate an understanding of the nature of mathematical truth and scientific knowledge.	Items 7, 27, 39, 40, 45 (5 items; 10% of the total test) M = 6.5 (65% correct) s = 1.9 α = .13
3. Demonstrate an understanding of the role of theories in science as unifying principles that explain observations and make predictions.	Items 1, 14, 30, 33 (4 items; 8% of the total test) M = 4.5 (56% correct) s = 2.1 α = .24
4. Demonstrate an understanding of the role of aesthetics in mathematics.	No items
5. Demonstrate an understanding of the interdependence of applied research, basic research, and technology, and how each can benefit society.	Items 3- 5, 12, 15 (5 items; 10% of the total test) M = 7.0 (70% correct) s = 2.5 α = .41
6. Illustrate how developments in mathematics and science can raise important ethical issues.	No items
7. Illustrate the interdependence of social forces and the fields of mathematics and science.	No items
Total for Learning Goal I	18 items (36% of total test) M = 23.9 (75% correct) s = 5.3 α = .55

LEARNING GOAL II: To apply mathematical and scientific concepts to understand the natural world.

Objectives	Test Items Assessing Each Objective
8. Design and execute experiments to solve problems or test hypotheses.	Item 46 (1 item; 2% of the total test) M = 0.8 (40% correct) s = .97
9. Obtain, organize, analyze, interpret, and present data.	Items 34-38 (5 items; 10% of the total test) M = 5.8 (58% correct) s = 2.4 $\alpha = .35$
10. Interpret data and statistical arguments presented in tables and graphical displays.	Items 8-10, 16, 17 (5 items; 10% of the total test) M = 5.3 (50% correct) s = 2.6 $\alpha = .42$
11. Discriminate between association and causation, and identify the types of evidence used to establish causation.	Items 13, 41-44 (5 items; 10% of the total test) M = 6.0 (60% correct) s = 2.4 $\alpha = .38$
12. Formulate an hypothesis and identify relevant variables necessary to test that hypothesis.	No items
13. Use mathematics as an abstract language to analyze natural phenomena.	Items 2, 6, 19 (3 items; 6% of the total test) M = 4.0 (66% correct) s = 1.5 $\alpha = .17$
14. Use graphical, symbolic, and numerical methods to make mathematical and statistical arguments.	Items 20-22, 25, 26 (5 items; 10% of the total test) M = 6.1 (61% correct) s = 2.2 $\alpha = .26$
15. Demonstrate an understanding of theories and models that describe natural phenomena.	Items 47– 50 (4 items; 8% of the total test) M = 4.2 (53% correct) s = 2.3 $\alpha = .49$
Total for Learning Goal II	28 items (56% of total test) M = 32.4 (58% correct) s = 8.2 $\alpha = .69$

LEARNING GOAL III: To evaluate mathematical and scientific arguments at a level commonly encountered by informed citizens.

Objectives	Test Items Assessing Each Objective
16. Evaluate the use of scientific, mathematical and statistical arguments in the analysis of public policy issues.	Items 28, 29, 31, 32 (4 items; 8% of the total test) M = 2.8 (35% correct) s = .81 $\alpha = .41$
17. Locate, evaluate, and select reliable resources to learn about scientific developments.	No Items
Total for Learning Goal III	4 items (8% of total test) M = 4.2 (53% correct) s = 2.3 $\alpha = .41$

QUANTITATIVE REASONING AND OVERALL TEST SCORES

Objectives	Test Items Assessing Each Objective
Quantitative Reasoning	Items 2, 6, 8-10, 13, 16, 17, 19-22, 25, 26, 28, 29, 31, 32, 34-38, 41-44 (27 items: 54% of the total test) M=31.3 (58% correct) s = 7.9 $\alpha = .66$
<i>Total for Entire Examination Spring 2002</i> <i>N = 227 package completers</i>	50 items (100% of test) M =60.4 (60.4% correct) s = 12.8 $\alpha = .77$

B. Comparison of First-Year with Sophomores and Juniors

If Cluster Three has an impact on student learning, then those students who have taken Cluster Three courses should score higher than those who have not. The best test of this prediction is to compare pre- and post-tests; i.e., to test students both before and after they have taken Cluster Three courses. The spring 2002 cohort of students completing NW-5 is the first group upon which a repeated measures analysis could be conducted. Many of the students who took the NW-5 test also took the same version of the test when they entered the University in fall 2000 as freshmen. This analysis will be conducted in the pre/post-test comparison section. At this time, we will compare the

performances of entering first-year students from the fall 2000 and 2001 administrations with sophomore/junior samples from spring 2000, 2001, and 2002 testing sessions.

Results

Separate analysis of variance procedures were used to compare the performance of the five groups on the NW-5 Total Test score and on the Quantitative Reasoning subscore. Significant overall F-tests were obtained for the total NW-5 test ($F_{4,4395} = 31.49$; $p < 0.0001$), as well as for Quantitative Reasoning ($F_{4,4395} = 27.34$; $p < 0.0001$). A Tukey test was used for follow-up comparisons of group performance on the NW-5 and Quantitative Reasoning. All three sophomore/junior groups scored significantly higher on the NW-5 and the Quantitative Reasoning subtest than did either of the first-year groups (Table 4). There were no significant differences between the three-sophomore/junior groups or the two freshman samples.

Table 4. Average overall test score and quantitative reasoning subscore for freshman and sophomore/juniors.

Group of Students	N	NW-5 Test Score	Quantitative Reasoning Subscore
Freshman (Fall 2000)	993	53.7 ^a	27.7 ^a
Freshman (Fall2001)	746	54.9 ^a	28.4 ^a
Soph/juniors (Spring 2000)	881	58.1 ^b	30.1 ^b
Soph/juniors (Spring 2001)	979	58.4 ^b	30.1 ^b
Soph/juniors (Spring 2002)	801	58.1 ^b	29.9 ^b

^bSuperscript letters indicate groups with significant differences ($p < 0.05$) based on Tukey multiple comparisons procedure.

Interpretation

The significantly higher scores of sophomore/juniors are consistent with a positive impact of Cluster Three courses. Alternatively, the increase could be due to general student maturation, to other courses students are taking at JMU, or to the loss from the University of those students with the lowest aptitude for science and mathematics.

Although the increases were statistically significant, the magnitudes of the increases in the overall score and the quantitative reasoning subscore were relatively small. One interpretation of this small increase is that Cluster Three courses do not greatly increase mastery of the objectives. However, several confounding factors make it difficult to determine the validity of this interpretation. First, the sophomore/junior group includes students who have not completed their Cluster Three requirements; these students score lower than students who have completed the Cluster (see below). Second, it appears that first-year students are more motivated than sophomore/juniors to do well on the assessment test. The motivation scores reported by entering students were higher than those reported by sophomore/juniors; these differences were statistically significant. The likely effect of this difference in motivation would be to decrease slightly the difference in NW-5 scores between freshman and sophomore/juniors. Third, there

may be differences in general aptitude for science and mathematics between the freshman of Fall 2000 and 2001 and the sophomore/juniors of 2000, 2001, and 2002 that may contribute to the observed difference.

Additional analyses

The Spring 2002 semester provided us the first opportunity to conduct repeated measures analyses of NW-5 test scores for students that entered the University in Fall 2000. Additional analyses to be conducted (see below) will help rule out general intellectual maturation as an alternative explanation for the observed increase in test scores.

C. Effect of Number of Cluster Three Courses Completed on Test Scores

If Cluster Three courses enhance students' mastery of Cluster Three learning objectives, then student scores should increase with increasing numbers of Cluster Three courses completed. In testing this prediction, we present two analyses. In the first analysis, the number of courses includes all Cluster-related courses, including those awarded through transfer and AP credit. In the second analysis, students with any transfer or AP credits were excluded, so that the effect of JMU courses alone could be assessed. In this analysis, the number of courses indicates the number of JMU courses taken.

These analyses examined sophomore/juniors taking the Spring 2002 test. The total pool of students available for these analyses ($N = 785$) is smaller than the pool used above ($N = 801$) because some students failed to fill in correctly their ID numbers, which precluded linkage of their NW-5 scores to their course histories. However, this loss was very small due to careful procedures to provide missing IDs during the scanning of answer sheets.

Results

Analyses including transfer and AP credits – Total Natural World scores and Quantitative Reasoning subscores show a gradual increase with increasing numbers of Cluster-related courses taken (Table 5). These overall increases were statistically significant for both measures. For the Total Test score, $F_{5,779} = 3.86$, $p = .0018$ with students taking four Cluster-related courses scoring higher on the Total Test score when compared to students with only one course completed. On the Quantitative Reasoning subscore, $F_{5,779} = 4.57$, $p = .0004$, again scores show a consistent and gradual increase with greater course completions. Table 5 superscripts indicate several significant differences on the Quantitative Reasoning subscore consistent with the hypothesis that increased coursework impacts learning achievement.

Table 5. Average Scores on NW-5 as a function of the number of Cluster-related courses taken either at JMU or elsewhere

Cluster-related Courses	N	Total Test Score	Quantitative Reasoning Score
None	16	52.2	25.1 ^{ac}
One	131	55.4 ^a	28.6 ^a
Two	201	57.4	29.1 ^a
Three	251	58.6	30.4
Four	145	60.7 ^b	31.5 ^b
Five or more	41	61.4	31.8 ^d

Superscript letters indicate groups with significant differences based on Tukey multiple comparisons procedure. For example, students who had completed four classes scored significantly higher ($p < 0.05$) on the Total Test than did students with one class. Analyses conducted with the Quantitative Reasoning scores revealed that students who completed four Cluster 3 classes performed significantly better than students with zero, one, or two courses completed. An additional significant difference was observed between students completing five courses or more and those with no coursework.

Similar results were obtained when only mathematics courses were examined. Total Test and Quantitative Reasoning scores increased with increasing numbers of mathematics courses taken (Table 6). These increases were both statistically significant (Total score: $F_{2,782} = 4.60$, $p = 0.0104$; Quantitative Reasoning: $F_{2,782} = 6.89$, $p = 0.0011$), with students taking two Cluster-related mathematics courses scoring higher on the Total Test score than students without a mathematics course completion. For the Quantitative Reasoning score students completing one or two courses performed significantly higher than students with no Cluster-related mathematics courses.

Table 6. Average Scores on NW-5 as a function of the number of Cluster-related mathematics courses taken either at JMU or elsewhere

Mathematics courses taken at JMU or elsewhere	N	Total Test Score	Quantitative Reasoning Score
None	62	54.2 ^a	27.1 ^a
One	529	58.0	29.8 ^b
Two	194	59.8 ^b	31.2 ^b

Superscript letters indicate groups with significant differences ($p < 0.05$) based on Tukey multiple comparisons procedure.

Analyses excluding students with transfer and AP credits – Total scores and Quantitative Reasoning scores show a gradual and consistent increase with increasing numbers of Cluster Three courses taken at JMU (Table 7). These increases were statistically significant for both measures (Total scores: $F_{4, 545} = 3.05$, $p = 0.0166$; Quantitative Reasoning: $F_{4, 545} = 3.31$, $p = 0.0107$), with students taking four or more Cluster Three courses scoring higher on the Total test than students completing only one course. On the Quantitative Reasoning subscore students with four or more JMU course completions scored significantly higher than students with no Cluster Three courses.

Table 7. Average scores on NW-5 as a function of the number of Cluster Three courses taken at JMU

Cluster-related Courses taken only at JMU	N	Total Test Score	Quantitative Reasoning Score
None	16	52.2	25.1 ^a
One	108	54.9 ^a	28.4
Two	161	57.8	29.2
Three	169	58.4	30.2
Four or more	96	60.1 ^b	31.0 ^b

Superscript letters indicate groups with significant differences ($p < 0.05$) based on Tukey multiple comparisons procedure.

Similar results were obtained when only mathematics courses were examined. Total and Quantitative Reasoning subscores show a consistent and gradual increase with increasing numbers of mathematics courses taken (Table 8). These increases were statistically significant for Natural World Total test scores ($F_{2,547} = 4.11$, $p = .0170$) and for Quantitative Reasoning subscores ($F_{2,547} = 5.18$, $p = 0.0059$). For both analyses students taking one or two Cluster Three mathematics courses scored significantly higher on the Total test score and statistically higher on the Quantitative Reasoning subscore than students with no Cluster Three mathematics courses completed.

Table 8. Average scores on NW-5 as a function of the number of Cluster Three mathematics courses taken at JMU

Mathematics Courses Taken at JMU or elsewhere	N	Total Test Score	Quantitative Reasoning Score
None	43	52.4 ^a	26.0 ^a
One	403	57.9 ^b	29.7 ^b
Two	104	58.7 ^b	30.2 ^b

Superscript letters indicate groups with significant differences ($p < 0.05$) based on Tukey multiple comparisons procedure.

Interpretation

The observed increases in NW-5 scores with increasing numbers of Cluster Three courses are consistent with the hypothesis that Cluster Three courses enhance student mastery of the Cluster Objectives. These results are intriguing since the total sophomore/junior sample tested will inevitably include students who have not taken any Cluster-Three related classes. These results suggest that the relationship between course number and test scores may be a more sensitive measure of the effect of Cluster Three courses than is the comparison of freshman with sophomore/juniors. Our results thus far provide evidence that sophomores score higher than entering students, and that increased course exposures whether taken at JMU or through other credit bearing opportunities are related to higher scores on both the Natural World total and the Quantitative Reasoning subtest. However, the relative impact of these opportunities was not addressed, nor was the hypothesis that overall student maturation might be the prime predictor of achievement on our Natural World assessment instruments.

D. Prediction of Natural World Assessment Test Scores

To better assess the relative contribution of JMU courses, Advanced Placement (AP), transfer courses, and to assess the impact of general maturation and courses outside the Cluster, a new analysis was conducted to predict NW-5 test scores using four predictor variables: number of JMU Cluster Three courses completed, number of Cluster-related AP, number of transfer courses completed, and the total number of credit hours completed at the time of the assessment. It was hoped this analysis would allow an assessment of the independent importance of JMU courses, AP, transfer courses, and general intellectual maturation, as measured by total credit hours might have on performance on the NW-5 test.

A multiple regression analysis was conducted to ascertain the contributions of the number of mathematics and science transfer credits, Advanced Placement credits in science and mathematics, JMU science and mathematics courses, and cumulative credit hours toward predicting performance on the Natural World assessment test. The four predictors were entered as a block. The overall relationship was significant $F_{4, 780} = 20.15$, $p < .001$, with $R^2 = .09$. Table 9 provides the regression results. The t-test follow-up results indicate whether each predictor in the model significantly contributed to prediction of the Natural World Total score. These results indicate that science and mathematics transfer courses did not contribute significantly to the prediction of Natural

World total test scores. However, the award of Advanced Placement credit hours and JMU course taking do significantly predict Natural World scores. In addition, the cumulative credit hours for all coursework appear to be a significant **negative** predictor of student performance. In other words, Natural World test scores are not enhanced simply via maturation through course taking; only relevant course exposures predict performance. The finding that Cluster Three related transfer courses do not predict Natural World test scores is intriguing and worthy of further study.

Table 9. Natural World test total model R^2 with B coefficients, Beta weights, and t-test results (N=785)
 $R^2 = .09$

Variable entered	B coefficient	Beta Coefficient	t-test
Transfer courses	.6967	.0474	1.166
AP courses	8.2188	.2629	7.561*
JMU courses	1.6291	.1562	4.065*
Total credit hours	-.2470	-.1435	-3.915*

* $p < .05$

A parallel analysis was conducted for the prediction of Quantitative Reasoning scores that resulted in very similar findings. The model accounted for 10% of the variance in Quantitative Reasoning scores. These results are presented in Table 10. The overall model was significant, $F(4, 780) = 21.74$, $p < .0001$, with $R^2 = .10$. Again, transfer credit hours in mathematics and science did not significantly predict test scores, but the number of Advanced Placement science and mathematics courses and JMU Cluster Three courses completed did. Again, the total number of credit hours accrued was a significant negative predictor. These results suggest that general collegiate maturation does not enhance Quantitative Reasoning scores; in fact it is a negative predictor of Quantitative Reasoning test performance. However Cluster Three relevant Advanced Placement and JMU coursework are significant and positive predictors of Quantitative Reasoning test performances.

Table 10. Quantitative Reasoning test total model R^2 with B coefficients, Beta weights, and t-test results (N=785)
 $R^2 = .10$

Variable entered	B coefficient	Beta Coefficient	t-test
Transfer courses	.3911	.0447	1.104
AP courses	5.3555	.2879	8.309*
JMU courses	.9298	.1498	3.914*
Total credit hours	-.1170	-.1143	-3.131*

* $p < .05$

E. Correlations of Test Scores with Cluster-Related Course Grades.

If Cluster Three courses help students master the Cluster objectives, then the grades assigned in Cluster Three courses should correlate with the scores of students on the NW-5. These correlations are expected to be positive, but they are not expected to exceed 0.50 because course grades are earned via a variety of skills, competencies, and procedures. Further, Cluster Three packages consist of three or more courses, and most courses generally cover material in addition to the objectives of Cluster Three.

Two analyses were conducted to test for positive correlations between scores and grades. First, correlations were calculated for individual courses for both the Total Test score and for the Quantitative Reasoning subscore. Results for both scores are presented for the Spring 2000, 2001, and 2002 assessments. Correlations are not presented for some cohorts taking ANTH 196, GEOL 110, GSCI 116, GSCI 121, GSCI 122, PHYS 140, PHYS 220, and PHYS 240 because fewer than 20 students were assessed in each of these courses. Statistical significance of individual correlations was not assessed because statistical significance is strongly dependent on sample size and sample sizes varied greatly among courses. Statistical significance of the correlations was assessed in a second analysis aimed at determining whether the average correlation across all courses was significantly greater than zero. Because the distribution of correlations was not normal, a nonparametric test (Wilcoxon signed ranks test) was used. If Cluster Three courses enhance mastery of the objectives, then the median correlation between course grades and test scores should be greater than zero. Separate analyses were conducted for the Total Test score and Quantitative Reasoning subscore correlations.

Results

For the six Cluster Three mathematics courses (Table 11), 15 (94%) of the 16 correlations between test scores and course grades were positive. Review of correlations between Quantitative Reasoning test scores and Cluster Three mathematics course grades resulted in identical results; 15 (94%) of the correlations were positive. The average correlation between test scores and course grades were significantly greater than zero for both the mathematics courses with Total Test scores and for the Quantitative Reasoning subscore.

Table 11. Correlations of test scores with mathematics course grades. Numbers in parentheses represent sample sizes.

Math Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
MATH 103	.37 (62)	.29 (53)	.26 (44)	.23 (62)	.26 (53)	.18 (44)
MATH 107	-.23 (21)	.02 (52)	.23 (73)	-.36 (21)	.09 (52)	.20 (73)
MATH 205	.12 (259)	.17 (272)	.29 (208)	.12 (259)	.17 (272)	.26 (208)
MATH 220	.14 (245)	.24 (320)	.19 (293)	.12 (245)	.21 (320)	.20 (293)
MATH 235	.18 (67)	.39 (74)	.26 (45)	.13 (67)	.35 (74)	.26 (45)
GMATH 115	-- (19)	.09 (24)	-- (20)	-- (19)	.22 (24)	-- (20)

A set of analyses for the spring 2002 correlations was also conducted. For the five spring 2002 mathematics courses for which correlations were available, all 10 correlations (100%) were positive. For the mathematics courses the median correlation with the Natural World Total test was $r = 0.24$, $p = .043$, $n=5$. For the Quantitative Reasoning subscore the median correlation with math course grades was $r = .24$, $p=.043$, $n=5$.

For the 14 Cluster Three science courses (Table 12), 37 (95%) of the 39 correlations with the Natural World Total score were positive. Correlations observed between Cluster Three science courses and Quantitative Reasoning test scores were also predominantly positive (95%). The average correlation between test scores and science related course grades were significantly greater than zero for both the Natural World Total and the Quantitative Reasoning score. These results are reported by Cluster Three package courses in Table 12.

A parallel set of analyses using only the spring 2002 correlations was also conducted. For the 12 Cluster Three science courses, 22 (92%) of the 24 correlations were positive. The median correlations were significantly different than zero when tested using the Wilcoxon signed ranks test. For the science courses the median correlation of Natural World Total test scores with course grades was $r = .22$, $p = .0029$, and $n= 12$. The median correlation between science course grades and Quantitative Reasoning subscores was $r = .25$, $p = .0029$, $n = 12$. As a follow-up the differences between math and science course median correlations was calculated and the two values were not significantly different from one another. In other words, the extent of relationship between correlations with Natural World and Quantitative Reasoning

subtest scores for science and mathematics courses were not found to be significantly different.

Table 12. Correlations of test scores with Cluster Three science courses, listed by package.

Package A Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
GSCI 101	.28 (352)	.24 (370)	.20 (252)	.27 (352)	.20 (370)	.17 (252)
GSCI 102	.13 (130)	.29 (107)	.20 (69)	.05 (130)	.23 (107)	.25 (69)
GSCI 103	.45 (91)	.28 (76)	.37 (57)	.49 (91)	.24 (76)	.25 (57)
GSCI 104	.15 (128)	.09 (164)	.15 (109)	.11 (128)	.04 (164)	.16 (109)

Package B Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
GISAT 112	.30 (107)	.26 (88)	.51 (85)	.32 (107)	.29 (88)	.56 (85)
GISAT 141	.13 (105)	.24 (98)	.47 (87)	.17 (105)	.25 (98)	.44 (87)
GISAT 113	.18 (101)	.24 (98)	.43 (69)	.16 (101)	.25 (98)	.43 (69)

Package C Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
CHEM 120	.02 (57)	.23 (60)	.24 (48)	.08 (57)	.29 (60)	.28 (48)
BIO 270	.34 (31)	.33 (49)	.11 (42)	.17 (31)	.26 (49)	.13 (42)

Package D Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
CHEM 131	.37 (100)	.40 (95)	.09 (74)	.40 (100)	.40 (95)	.18 (74)
BIO 130	.39 (56)	.31 (54)	-.05 (34)	.35 (56)	.32 (54)	-.07 (34)

Package E Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
GSCI 115	.54 (23)	.32 (26)	.35 (26)	.43 (23)	.33 (26)	.46 (26)
GSCI 116	-- (17)	-.23 (21)	-- (20)	-- (17)	-.10 (21)	-- (20)

Package F Courses	Total Test Score			Quantitative Reasoning Score		
	Spring 2000	Spring 2001	Spring 2002	Spring 2000	Spring 2001	Spring 2002
GSCI 121	.31 (24)	.26 (40)	-- (18)	.32 (24)	.26 (40)	-- (18)

Interpretation

The finding that there were positive correlations between course grades and NW-5 scores for the cluster related courses supports the hypothesis that Cluster Three courses enhance student mastery of objectives. Few correlations exceeded 0.50, but strong correlations were not expected given that grades in courses often reflect learning of material in addition to Cluster Three objectives.

Several cautions must be exercised in interpreting the correlations above for individual courses. First, relatively large sample sizes are needed to detect statistically significant correlations of the magnitude expected for these analyses. For example, a sample size of 43 is required to determine that a correlation of 0.30 is statistically significant; a correlation of this magnitude would be considered meaningful in this context. Thus, the correlations reported for GSCI 121, for example, are not significantly different than zero, although their magnitude is similar to what is predicted if this course has a positive impact on mastery of Cluster objectives. This is why these correlations are not tested for significance.

There are several reasons why there may not be a moderate positive correlation between grades in a course and performance on the NW-5 test, other than the hypothesis that a course does not enhance student mastery of Cluster objectives. First, a course may cover enough material in addition to the Cluster Three objectives that course grades are not determined primarily by mastery of Cluster Three objectives. Second, lack of assessment of some Cluster Three objectives may mean that the NW-5 did not adequately assess the Cluster Three objectives covered in a specific course.

Additional Analyses

For courses that did not exhibit moderate positive correlations, it may be useful for course faculty to explore the above-mentioned hypotheses to determine if either applies. Examination of correlations between specific items on NW-5 and specific course assessment that are both related to the same Cluster Three objectives may

provide a more appropriate analysis in cases where course content and NW-5 assessment do not overlap extensively.

F. Differences among Packages

To determine whether student performance on the NW-5 varied by package, a one-way analysis of variance was conducted using NW-5 scores for students who had completed a package. There were significant overall differences among packages in Total Test scores ($F_{5,214}=2.40$, $p = 0.04$) and in Quantitative Reasoning (Quantitative Reasoning $F_{5,214}=2.66$, $p=.02$). However, when multiple comparisons were conducted to determine whether individual groups were significantly different from others, no significant differences were obtained.

Table13. Natural World Total Score Performances and Quantitative Reasoning Total Score Performance By Package Completed.

Cluster Three Package Completed	N	NW-5 Mean	QUANT Mean
A: General Science & Mathematics	53	59.1	30.0
B: Living Systems & the Environment: From Cell to Globe	63	63.3	33.1
C: Health Issues	28	55.9	28.6
D: Advanced Science & Mathematics	54	63.8	33.6
E: Humans & the Earth Environment	18	61.9	32.2
F: Light & Sound: Science & Perceptions	4	59.3	31.5

No significant differences among the groups were observed using Tukey multiple comparison tests.

Interpretation

Sample sizes for some packages (e.g., E and F) are quite small, which negatively impacts the statistical power with which significant differences can be detected. Therefore, it remains unknown whether these packages scored significantly higher or lower than other packages. It would be useful for faculty to review the scores of the students completing their particular package in relation to other packages to assess whether the difference observed is of practical or educational significance to them. However, the scores of package completers may not be the best measure of a package's effectiveness in increasing student mastery of objectives; it is likely that self-selection into particular packages might influence test performances. A better way to assess the impact of packages would be to examine the increase in student scores in a pre- and post-test design. This analysis follows.

G. Pre-Post Test Comparisons for Package Completers

A pre- and post-test comparison was conducted using individuals that had completed all Cluster Three requirements and for whom scores for both the pre-test and the post-test were obtained. For this analysis, a Cluster Three package was included only if 20 or more subjects were obtained. A repeated measures multivariate analysis revealed significantly higher overall posttest scores, but no interaction effect by package for both the NW-5 Total test score ($F_{4, 142} = 3.52, p = .009$) and for the Quantitative Reasoning subscore ($F_{4, 142} = 2.92, p = .023$). In other words, students performed significantly better on the Natural World after completing their coursework. In addition, no significant differences were observed when the packages students completed were compared. These results were obtained for both the Total test score, and for the Quantitative Reasoning subtest.

These are very powerful results that the Cluster Three committee has waited a long time to test. Table 14 provides the pre and post-test scores and standard deviations on the NW-5 Test and the Quantitative Reasoning subtest at the beginning of their freshmen year and at the midpoint of their sophomore/junior year.

Table 14. Change over Time for Cluster Completers Total Score and Quantitative Reasoning Subscore Means (Standard Deviations) Spring 2002

Level of Analysis	N	Fall 2000 Group (Pre) Total	Spring 2002 Group (Post) Total	Fall 2000 Group (Pre) Quantitative Reasoning	Spring 2002 Group (Post) Quantitative Reasoning
All Package Completers	148	56.8 (10.2)	62.7 (11.9)	29.5 (6.7)	32.7 (7.4)
Package A	29	52.5 (12.0)	61.3 (9.9)	27.0 (7.1)	32.2 (6.2)
Package B	45	58.5 (8.7)	64.0 (12.9)	31.0 (3.2)	33.2 (8.2)
Package C	20	52.5 (9.1)	56.9 (11.6)	26.6 (5.4)	29.5 (7.6)
Package D	42	59.6 (10.3)	64.5 (12.8)	30.3 (7.0)	33.6 (7.9)
*Package E	12				
*Package F	1				

*Means are not reported for groups smaller than 20.

Differences in Test Performance by Number of Package Courses Completed

Data in Section C examined whether performance on NW-5 increased with Cluster Three courses as a whole. In this section, the results for individual packages are presented. These analyses have the potential to help package faculty determine the degree to which courses in their packages help students master Cluster objectives. In

the analyses below, package courses completed include AP and transfer credit courses, in addition to JMU Cluster Three courses. Unlike the analysis presented above combining all packages, the package analyses below do not include students who have taken no Cluster Three courses because it is impossible to assign students to packages with certainty until they have taken at least one course.

Results

Package A – Test scores tended to increase with increasing numbers of courses taken (Table 15; Total test scores: $F_{3,252} = .34$, $p = 0.80$; Quantitative Reasoning scores: $F_{3,252} = .53$, $p = 0.66$). However, none of the observed means were significantly different from one another. Small sample sizes may contribute to the null finding. We await the interpretation of the faculty for interpretation of these results.

Table 15. Test Performances across Package A Course Completions

Number of Package A Courses Completed	N	Total	Quantitative Reasoning
One	8	58.0	29.0
Two	82	59.0	29.8
Three	83	58.7	30.2
Four	83	60.3	31.0

There were no significant differences between scores based on number of package A courses completed.

Package B – Neither of the analysis of variance tests resulted in significant overall effects (Total test score: $F_{2,83} = .63$, $p = 0.54$; Quantitative reasoning: $F_{2,83} = .53$, $p = 0.59$). Mean performances for Total and Quantitative Reasoning scores are presented in Table 16 by the number of courses completed. There was a slight increase in test score performances with increased course completion, but no significant differences were observed.

Table 16. Test Performances across Package B Course Completions

Number of Package B Courses Completed	N	Total	Quantitative Reasoning
One	8	58.4	30.8
Two	19	60.2	30.9
Three	59	63.2	33.0

There were no significant differences between scores based on number of package B courses completed.

Package C – While test scores increased consistently with increasing courses completed in Package C for both the Total Test score and the Quantitative Reasoning subscore, no significant differences were detected (Total test score: $F_{2,49} = 1.92$, $p = 0.16$; Quantitative Reasoning: $F_{2,49} = 1.31$, $p = 0.28$).

Table 17. Test Performances across Package C Course Completions

Number of Package C Courses Completed	N	Total	Quantitative Reasoning
One	5	41.4	21.6
Two	21	52.0	27.1
Three	26	55.6	28.3

There were no significant differences between scores based on number of package C courses completed.

Package D – No significant differences were observed for Package D student course completions for the Total Test score or the Quantitative Reasoning subscore (Table 18), (Total scores: $F_{2,56} = 1.02$, $p = 0.37$; Quantitative Reasoning: $F_{2,56} = 2.39$, $p = 0.10$).

Table 18. Test Performances across Package D Course Completions

Number of Package D Courses Completed	N	Total	Quantitative Reasoning
One	8	61.7	32.5
Two	28	57.4	27.5
Three	23	61.3	31.9

There were no significant differences between scores based on number of package D courses completed.

Package E – Test scores tended to increase with increasing courses completed in Package E for both the Total Test score and the Quantitative Reasoning subscore (Table 19). These increases were not statistically significant for Total Scores ($F_{2,23} = 2.65$, $p = 0.09$), but they were for Quantitative Reasoning: $F_{2,23} = 4.00$, $p = 0.03$). Students completing three Package E courses scored significantly higher on the Quantitative Reasoning subtest than did students that had completed only one course.

Table 19. Test Performances across Package E Course Completions

Number of Package E Courses Completed	N	Total	Quantitative Reasoning
One	6	51.2	24.3 ^a
Two	2	55.0	27.0
Three	18	61.9	32.2 ^b

Superscript letters indicate groups with significant differences ($p < 0.05$) based on Tukey multiple comparisons procedure.

Package F – While scores tended to increase with the addition of completed courses in the package, no significant differences were observed for either the Total Test score ($F_{2,16} = .44$, $p = .65$) or for the Quantitative Reasoning subscore ($F_{2,16} = .95$, $p = .41$). These results are presented in Table 19. Small samples sizes again may have contributed to these null findings. We leave it to the faculty to interpret whether the obtained results bear educational or practical significance.

Table 19. Test Performances across Package F Course Completions

Number of Package F Courses Completed	N	Total	Quantitative Reasoning
One	1	58.0	28.0
Two	10	67.3	33.6
Three	8	67.1	36.5

There were no significant differences between scores based on number of package F courses completed.

Interpretation

Very few statistically significant results were obtained in this set of analyses. We had a reduced sample size for the spring 2002 assessment in comparison with the previous year. In many cases, it is very difficult to identify with precision those students that have commenced their package work, because several courses are shared across several packages. Many students were left out of these analyses. This may have contributed to the inability to discern significant effects that may be present. Further, in most packages, the number of sophomore/juniors with only one course completed was

relatively low. Nonetheless, the relatively consistent finding across packages that test scores increase with increasing numbers of package courses completed supports the hypothesis that packages enhance mastery of objectives. We will need to replicate this set of analyses with future Assessment Day samples. Hopefully, we will be able to collect data from a larger sample.

IV. General Conclusions

Overall, existing data support the hypothesis that Cluster Three courses enhance mastery of Cluster objectives. Sophomores and juniors taking NW-5 scored higher than freshman. Among sophomores and juniors, scores on NW-5 increased with increasing numbers of Cluster Three related courses, and the magnitudes of the increases were slightly greater when students with transfer or AP credit for Cluster Three courses were excluded. For the vast majority of courses, there was a positive correlation between course grades and scores on NW-5, and the median correlation coefficient was significantly greater than zero. Within most packages, performance on the NW-5 increased with increasing numbers of courses. A multiple regression analysis revealed that increased relevant coursework at JMU or via AP credits significantly and positively predict NW-5 Total and Quantitative Reasoning test scores. Further, the accumulation of college credit hours is a significant and negative predictor of test scores, while transfer credits in science and mathematics do not predict NW-5 test performances. These are intriguing findings. The finding that cumulative credit hours is a negative and significant predictor of NW-5 test scores is a direct refutation of the rival hypothesis that increases in scores might be attributable solely to academic maturation. The finding that transfer credits do not predict test scores is also important and worthy of additional study. Although alternative explanations might be generated for each of these results (see discussions above), it is more difficult to generate an alternative explanation for all of these results, and therefore, the conclusion that Cluster Three courses enhance student learning with respect to Cluster objectives seems reasonable.

Despite the overall conclusion that Cluster Three courses appear effective in meeting the Cluster objectives, several issues remain. First, the Cluster faculty have not determined a standard for judging whether the observed average test scores, for either the overall test score or individual objectives, represent an overall poor, moderate, or strong performance. Similarly, it is unclear whether the observed increases in scores with increasing courses completed represent small, medium, or large increases. It is recommended that Cluster faculty establish standards for unsatisfactory, satisfactory, and excellent performance.

Second, there are some problems in inferring cause and effect from nonexperimental studies such as this assessment. Because there can be no control group in which students are assessed for their mastery of Cluster Three objectives without taking Cluster Three courses, we cannot be absolutely certain that increases in assessment scores between freshman and sophomore/juniors or with increasing numbers of Cluster Three courses completed are due to Cluster Three courses.

However, our ability to add new analyses has strengthened our contention that Cluster Three courses significantly contribute to mastery of Cluster objectives.

Third, coverage of objectives on NW-5 is very uneven. Nearly one-third of the objectives (5 of 17) are not assessed on NW-5, and the remaining objectives represent from 2% to 10% of the exam. The Cluster Three Committee has entertained a reconfiguration of the goals and objectives that will reduce their number. When the objectives to be assessed become stable, more equal and complete coverage of objectives with new test items will be planned. We will work toward better coverage of the Cluster objectives.

Finally, the current assessment strategy, while having the potential to inform Cluster Three faculty about the degree to which students are achieving Cluster objectives, does little to inform faculty on how they might improve their courses to enhance mastery of objectives. To be most informative to Cluster faculty, assessment should indicate to faculty within packages or courses the degree to which their students are meeting each objective. For example, it would be very helpful for faculty to know how well students in their package or course scored on the assessment of a particular objective or how well grades in their course correlate with scores for a particular objective. Currently, the low reliability of the assessment of individual objectives (see Table 2) precludes examination of performance on specific objectives. Thus, faculty within any course or package cannot use the current assessment to inform them about which objectives need the most work in their course or package. Cluster faculty should consider alternative assessment strategies that more fully inform them about specific objectives.

V. Recommendations

1. Develop standards for unsatisfactory, satisfactory, and exceptional performance on NW-5.
2. Assess all objectives and equalize coverage of objectives on NW-5.
3. Develop assessment strategies that allow faculty to determine which objectives need the most improvement in a course or package.

This concludes the spring 2002 report. We leave any additional interpretation to the faculty that teach in Cluster Three. We also welcome recommendations for additional analyses and means by which the results can inform better teaching and learning at JMU.