ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education: Effective Program Processes

Section #1: Contact Information of Person Submitting Application

Name	Christopher G. Murphy/ Donna L. Sundre
Title	Cluster Three Coordinator, General Education Program, and Associate Professor of Biology Executive Director, Center for Assessment and Research Studies and Professor of Graduate Psychology /
Institution	James Madison University
Department/Program	Department of Biology and University Studies/ Professor of Graduate Psychology and Center for Assessment and Research Studies/
Street Address	800 South Main Street
City, State, Zip	Harrisonburg, VA 22807
Phone	540.568.2880 (Murphy); 540.568.3483 (Sundre)
Fax	540.568.3333 (Murphy); 540.568.7878 (Sundre)
Email	murphycg@jmu.edu; sundredl@jmu.edu
Signature	

Section #2: Institutional Endorsement

Chief Executive Officer or Chief Academic Officer

Name	Dr. Douglas T. Brown
Title	Provost and Vice President for Academic Affairs
Institution	James Madison University
Phone	540.568.3429
Fax	540.568.7800
Email	browndt@jmu.edu
Signature	

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education:
Effective Program Processes

Section #3: Application Summary

Summarize your award application, identifying the specific general education outcome or outcomes for a single learning domain that provided the focus of your work and achievement. Begin your narrative with a brief description of your institution and the time frame for the work you did, briefly explaining your activities and why you think they improved quality. ***Limit your summary to 150 words.*** Increase the text box size as needed.

James Madison University's Assessment of the Scientific Reasoning Learning Domain in General Education

This application describes James Madison University's (JMU) assessment efforts with respect to scientific reasoning in the General Education program. JMU is a Virginia public comprehensive university of 18,000, with a primary emphasis on undergraduate education. Scientific reasoning is embodied in eight, explicitly stated learning objectives, developed through a sustained, interdisciplinary collaboration that, over 10 years, has led to linkage of these outcomes to the curriculum, scientifically based data collection, and locally developed, direct assessments of learning. Our instruments have generalized and been adopted by others. Assessment data is systematically gathered and reported twice a year, and these results are reviewed and interpreted by a faculty-led steering committee, placed on our websites, reported to the State Council, and provided to students in a new assessment performance feedback system. Assessment has documented student learning gains, led to improvements in the instrument, and guided modifications to the curriculum that have improved student learning.

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education:
Effective Program Processes

Section #4: Award Criteria

Criterion 1: Defining Learning

Describe how your institution developed its operational definition of learning for the outcome(s) discussed in this application. What have you done to understand the importance and value of the outcome(s) to stakeholders both inside and outside your institution? Address all of the following:

- The process used to identify and develop your institution's operational definition of the learning intended by this outcome or outcomes for a single learning domain
- The activities, research, or evidence used to justify this definition
- The individuals involved in the development of this definition
- The process used to communicate this definition to faculty and other interested parties

Limit your response to two pages. Increase the text box size as needed.

Scientific reasoning forms a central part of the General Education program at James Madison University and is directly connected to the university's mission: "We are a community committed to preparing students to be educated and enlightened citizens who lead productive and meaningful lives." To make informed decisions, both in their personal lives and as citizens participating in a democratic society where debates over public policy issues often surround society's most complex and pressing problems, educated and enlightened citizens must wisely incorporate scientific information, drawing that information from reliable sources and applying that knowledge critically. At JMU, scientific reasoning is embodied in Cluster Three of the General Education program. Cluster Three is one of the five clusters, or learning areas, each of which emphasize unique tools, rationales, and methodologies. Cluster Three encompasses two ways of knowing: mathematics and science, which provide the quantitative and scientific reasoning skills that all citizens must employ in understanding and utilizing science in their everyday lives. Because reaching valid conclusions in science depends critically on quantitative reasoning, such reasoning is viewed in Cluster Three as an inseparable component of scientific reasoning.

Learning outcomes in Cluster Three are defined by a set of objectives designed to provide all students with the essential knowledge, skills, and experiences to apply mathematics and science as educated citizens. These objectives were first developed in 1996 by a group of 17 individuals that included faculty from seven mathematics and science departments, a faculty member from assessment, a representative from student services, and a postdoctoral fellow from the general education program. This group examined the learning objectives from other general education programs and after several months of work, produced a group of 17 learning objectives organized under three learning goals. During the development of the objectives, the group sent drafts to the university faculty, followed by multiple, university-wide meetings to gather input from faculty. The objectives were finalized and first implemented in Fall 1997. The short, one-sentence objectives were supplemented with longer descriptions, so faculty would know precisely what was meant by each objective and examples were provided on how the objective might be assessed in classes.

In 2001, a process to revise the objectives was initiated by the faculty committee that oversees Cluster Three. The Cluster Three committee consists of faculty from the eight departments that teach courses in the cluster, a representative from the Center for Assessment and Research Services, and a student representative; the chair of this committee is the coordinator of the cluster (currently Dr. Christopher G. Murphy). The revision was prompted by faculty input indicating that the original objectives were too difficult to implement completely in the number of allotted credit hours and by the results of a survey of faculty teaching courses in the cluster indicating that some objectives were not being covered in any depth in courses.

In this revision, the committee condensed and reorganized the previous 17 objectives into eight objectives, and the three learning goals were removed. The explanations of the objectives were revised, and a section was added to each objective indicating why every student needs to master the objective (i.e., the importance of the objective to the life of an educated citizen). Faculty teaching in the cluster were surveyed about the scope, appropriateness, and level of difficulty of the objectives, as well as the clarity and sufficiency of the new explanations of the objectives. The survey asked faculty to approve or disapprove the revisions, and 100% of those responding approved the changes. The new objectives were implemented in Fall 2003.

The learning objectives for the cluster are posted on the cluster's web page, where the requirements and structure of the cluster may also be found. The objective explanations are supplied to new faculty as they develop courses or sections of existing courses, and communication about objectives also occurs via a Blackboard Organization that enrolls all faculty teaching in the cluster, their department heads, all deans, and the leaders of various units with an interest in the cluster (e.g., directors of advising, coordinators of interdisciplinary programs, student support). The current objectives are:

- Describe the methods of inquiry that lead to mathematical truth and scientific knowledge and be able to distinguish science from pseudoscience.
- Use theories and models as unifying principles that help us understand natural phenomena and make predictions.
- Recognize the interdependence of applied research, basic research, and technology, and how they affect society.
- Illustrate the interdependence between developments in science and social and ethical issues.
- Use graphical, symbolic, and numerical methods to analyze, organize, and interpret natural phenomena.
- Discriminate between association and causation, and identify the types of evidence used to establish causation.
- Formulate hypotheses, identify relevant variables, and design experiments to test hypotheses.
- Evaluate the credibility, use, and misuse of scientific and mathematical information in scientific developments and public-policy issues.

Taken together, these objectives are defined by Cluster Three as embodying scientific reasoning. Individual objectives are assigned to specific groups of courses in the cluster's structure, such that each course is required to cover four or five of these objectives in-depth (defined as explicit treatment of the objective in the course, with the requirement that students apply their knowledge to new situations).

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education:
Effective Program Processes

Award Criteria

Criterion 2: Developing Assessment Methods and Tools
Describe how your institution developed and approved the methods (i.e., direct or indirect measures, stand-alone or assignment-embedded assessments, programs assessed, points or levels at which you do assess assessments, etc.) and tools (such as rubrics) you use to assess the outcome(s) identified in Criterion 1 above. Address all of the following:

- A description of the assessment methods and tools
- The individuals or groups involved in developing the process
- The research upon which the assessment and tools are based
- The process used to assure the necessary commitment and support for assessment methods and tools

Limit your response to two pages. Increase the text box size as needed.

Three hallmarks of JMU assessment practice relate to the infrastructure of support provided by the University, the quality of locally developed instruments, and the scientific rigor of our assessment data collection methods. Each of these will be described in turn as support for our tools and methods.

The infrastructure provided by the University assures sustained attention and resources to all assessment efforts at several levels. Unlike many institutions, we have: 1) a Dean of Undergraduate Studies, Dr. Linda Cabe-Halpern, who oversees all of general education, the Honors Program, and several support services; she reports directly to the Provost; 2) a General Education Council, which meets monthly and is chaired by an elected faculty member and populated by representatives of all five clusters, all contributing general education academic departments, student affairs, students, the library, and the Center for Assessment and Research Studies (CARS); 3) a Cluster Three Steering Committee (described under Criterion 1 above); and 4) CARS, the largest assessment center in the nation. Every cluster has at least one CARS liaison assigned to provide technical and measurement consultation. Dr. Donna Sundre is the Cluster Three CARS liaison and Executive Director of CARS.

Our work in Cluster Three began in 1996 in preparation for our new General Education cluster and has continued for over a decade. During this period, we have witnessed continuous improvement in the definitional clarity of our learning objectives and the crafting of our locally developed instrument that measures scientific reasoning. Our work together has been guided by the scientific foundation provided by researchers with expertise in measurement, science, and mathematics. The worked described above to improve the learning objectives for scientific reasoning has promulgated the writing of more finely tuned assessment items and has stimulated progress in the assessment of scientific reasoning at our institution. We are currently using the ninth version of our locally developed instrument, which is a selected response instrument with 66 items that directly measures student performance. The items focus on scientific reasoning skills and thus do not require discipline-specific factual knowledge; instead, the items make it possible for students who have completed their Cluster Three requirements and mastered the objectives to answer the items correctly regardless of the specific combination of Cluster Three courses they have completed.

JMU funded multiple summer Faculty Institutes where STEM faculty could come together to discuss the learning objectives and to write new selected-response items that were more innovative and interesting and that addressed higher levels of cognition than previous items. Working collaboratively with our STEM faculty, we have learned a great deal about what general education is and how to create appropriate items. We have eliminated items that assess discipline-specific factual information, and replaced those that assess student ability to understand and use science as a way of knowing. This extensive review of the instrument has resulted in a steady increase over the years in the reliability of our instruments (see Table 1 below). Given our success, JMU has been approached by several institutions and has marketed the instrument to others.

We collect data twice a year and prepare reports for each administration. These reports provide thorough psychometric analyses and regularly test faculty-derived hypotheses about student learning and development. We have also conducted qualitative studies to gather information about item quality and student perceptions of our assessment processes. For example, we conducted 'think-aloud' studies with students to determine the strategies used to solve problems. We discovered which items were confusing, boring or trivial and eliminated them. Our students enjoyed providing us with information by which we can improve our practice. We are pleased with the continued improvement of the instruments and our data collection processes.

A final hallmark of JMU assessment practice is our two Assessment Days each year. The first, Fall Assessment Day, takes place in August as an integral part of the required Orientation of entering first-year students. Half of the students are assessed in the morning, and the remaining half, in the afternoon. In fall 2008, we assessed over 4,000 students. Students are assigned to testing locations on the basis of the last 2 digits of their JMU ID; this procedure provides large, random, and representative samples of students for each of our general education assessment instruments (see Table 1 below for sample sizes for Cluster Three assessments). No student completes all assessment tests; they complete only the assessments in their assigned locations. The second, Spring Assessment Day, takes place in February and involves all students with 45-70 cumulative credit hours; this is the midpoint of their undergraduate academic career. Classes are cancelled on this day to avoid time and room conflicts. This spring Assessment Day is also used for graduating senior data collection for assessment in the majors. Because student IDs do not change, we can assure that students will take the same instruments on both of their Assessment Days. We have been using these procedures for almost 25 years, and it is an expected part of JMU student and faculty life. This rigorous design allows many quasi-experimental studies to be conducted and has contributed to the scholarship of assessment through presentations and publications. The entire campus benefits from this assessment support and infrastructure.

Table 1. Sample Sizes and Reliability (α = Cronbach's alpha) for the Scientific Reasoning Tests, Fall 2000 through Spring 2009.

			First-year Students (Fall)		***Sophomores/Juniors (Spring)***	
Acad. Year	**Test Form**	**Number of Items**	***N***	**α**	***N***	**α**
2000-2001	5	50	993	.65	979	.75
2001-2002	5	50	746	.67	801	.77
2002-2003	5	50	1084	.69	1174	.75
2003-2004	6	80	1304	.78	839	.85
2004-2005	7	85	839	.81	770	.87
2005-2006	8	85	1117	.78	526	.86
2007-2008	9	66	1408	.78	1020	.80
2008-2009	9	66	1592	.81	1113	.83

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education:
Effective Program Processes

Award Criteria

Criterion 3: Completing the Assessments
Describe your assessment process, the results of your assessments, and your plans to improve student learning based on your results. Address all of the following:

- The assessment process (both student and assessors' responsibilities)
- The individuals or groups involved in implementing and operating the process
- The process used to collect and verify the data
- The process used to identify and select learning improvements based on the assessments

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education: Effective Program Processes

Limit your response to two pages. Increase the text box size as needed.

As briefly described in Criterion 2, JMU conducts two formal Assessment Days each academic year. We have been conducting these events for almost 25 years, and we have been very satisfied with the rigor of these procedures. It should be noted that the challenges have been many as the campus has grown from 12,000 to over 18,500 students. The planning and conduct of Assessment Days represents a major responsibility of the Center for Assessment and Research Studies, as we now hire and train over 70 proctors to assess over 4,000 entering first year students each fall and over 3,000 sophomores and juniors every spring. The entire JMU community is committed to the assessment process: it appears in our catalog, the Fall Assessment Day is an integral part of the required orientation procedures for entering students, both Assessment Days appear on the campus academic calendar every year, classes are cancelled on the Spring Assessment Day, students are required to participate in Assessment Days, and if they do not attend, a hold is placed on their course registration until they attend a make-up session to complete their assessments. Last spring (2009), we enjoyed an 89% participation rate on Assessment Day, a rate that is higher than that for normal course attendance. We achieve 100% student participation in assessment activities after make-up sessions are completed. Our student focus groups indicate that students know what Assessment Day is and participate with energy.

We have implemented a variety of innovative strategies to engage our faculty, students, staff and proctors in the success of our assessment activities. We regularly gather information from every student during Assessment Days to measure examinee motivation. Using an instrument called the Student Opinion Scale (available for download and free use), we have been able to identify small subsets of individuals who do not put forth good effort. Considerable assessment scholarship has been evidenced with this research via presentations and publications. For example, we know that the factor that explains the most variability in student motivation is not time of day, content matter, paper and pencil vs. computer-based testing, or the size of the room. It is proctors. We concentrated on identifying 'master' proctors, and we modified our selection and our training for proctors. As a result, we have seen significant and substantive changes in examinee motivation, greater confidence and efficacy on the part of proctors in handling their assessment rooms, and much greater satisfaction with their proctoring tasks. Our research long indicated that new entering students presented very few assessment problems; however sophomores and juniors were another matter. Again, we relied on our motivation research and our focus groups to guide the way on identifying problems, trying out solutions, and confirming positive results. Not only have our participation rates improved, motivation scores look quite good (even for sophomores and juniors), but we have initiated new projects that will further improve our work. For example, our focus groups indicated that incentives (e.g., a chance to win a prize) would not motivate students, and participants made a few interesting suggestions. They thought that stickers could be placed on their shirts after assessment completion, similar to those distributed after voting. They also wanted feedback on their performances. We worked with the Information Technology Division to create test score variables in the Student Information System and loaded individual student scores to their records. Students can now access these scores vie our eCampus system. They can click on a link to the CARS website to receive two different kinds of interpretive information about their scores. We are currently conducting a study comparing the motivation and performances of students between those who were told they would receive feedback opportunities with those who were not. Our results indicate significantly higher scores for those who were informed; however, the effect sizes were not large. No motivation score differences were observed. We also know that our students (both first-year and sophomore-juniors) prefer interpretive information that describes their performances in relation to others, rather than in relation to faculty expectations or standards.

We have developed a set of faculty expectations for student performance in a year-long effort involving 38 STEM faculty. We used formal workshops, one-on-one sessions, and independent faculty completion of the standard-setting exercises to maximize participation (we found that the different procedures made no significant difference in the derived standards). The standard-setting procedure required faculty to review the instrument item by item and establish an expectation for the percentage of students who should get each item correct after they had completed their cluster requirements. Percentages for individual items were used to construct expected averages for performance on individual cluster objectives and on the overall instrument. As a result of project participation, our faculty have greatly increased familiarity with our cluster student learning goals and a much greater understanding and appreciation for the quality and breadth of our locally developed assessment instrument. Participation also provided faculty with a stronger contextual understanding of the instrument that will aid them in using and appropriately interpreting Assessment Day reports. We have been able to use these community standards as an additional lens for interpreting student test performances, and our results show a almost tripling of the percentage of students who exceed the faculty-predicted average on the instrument between first-year students and sophomore/juniors.

Identification and selection of learning improvements is done by the cluster committee. This group receives the results of data and statistical analyses conducted by CARS, provides the interpretation of the results, reaches conclusions based on the data, and often requests additional analyses to address questioned raised by the original analyses. Upon receiving the final analyses, the committee provides their final recommendations and then selects the learning improvements to be pursued.

In summary, at JMU assessment involves all key players, and everyone plays an important role. We have engaged in creative and innovative practices to enhance examinee motivation and meaningful participation in our assessment work. Our innovative practices have involved a broad cohort of faculty who now express greater confidence in our assessment findings and the efficacy of our programs. We believe we still have more work to do.

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education:
Effective Program Processes

Award Criteria

Criterion 4: Implementation of Learning Improvements and Evidence of Improved Learning
Describe the steps you took to improve learning and provide evidence of the success of your improvement strategies. Address all of the following:

- A description of your improvement strategies
- The individuals or groups involved in developing these strategies
- A description of how you check your learning improvements to assure they are continuing
- Results of follow-up assessments that provide evidence of sustained improvement

ASSOCIATION FOR GENERAL AND LIBERAL STUDIES

2009 AGLS Awards for Improving General Education: Effective Program Processes

Limit your response to two pages. Increase the text box size as needed.

The first use to which our assessment data were put was the origination of the current General Education program. In the early 80's and 90's JMU developed an outstanding assessment data collection design; however, our ability to detect significant learning gains was largely unsuccessful. We developed local tools with our faculty, but our learning goals and the courses designated to address them were so numerous, we could not demonstrate learning gains. JMU made a very courageous and difficult decision: to sunset the liberal arts core program and to create new, more meaningful learning objectives upon which a program could be built. This led to a core faculty group that worked intensely for over a year to create these new learning goals and objectives (including those for Cluster Three); the result was the new general education program that was implemented in fall 1997. This program required approval of courses on the basis of demonstrated alignment with the new learning goals and objectives for all general education study areas. Courses were allowed to contribute to general education for three years, at which time assessment data would be necessary for continuance. Since that time, we have continued to monitor and refine our learning goals and objectives and to improve our assessment tools. Our student learning outcomes have now been demonstrated regularly over the last 10 years.

The current improvement strategy in Cluster Three is to use the results of our assessment to inform three aspects of the curriculum: 1) mastery of the cluster objectives, 2) program structure, and 3) acceptance of alternative forms of credit for the cluster (i.e., transfer and Advanced Placement credit). Learning improvements are measured as increases between years in the overall score on the assessment instrument, scores for individual objectives, and for individual items, as well as the percentage of students exceeding the faculty predicted average for these three measures.

To improve mastery of cluster objectives, three approaches are used. First, the Cluster Coordinator distributes the annual cluster report to all faculty in the cluster via email and also posts the report to the Blackboard organization for the cluster. The faculty receive as attachments the entire report and a one-page summary that highlights the most important results and provides one or two figures presenting these results. One of these figures shows, for each of the eight cluster objectives, the percentage of students exceeding the faculty-predicted average for the assessment instrument for freshman, sophomores/juniors who have completed the cluster, and sophomores/juniors who have taken some cluster courses but have not completed the cluster. The body of the email highlights those objectives that exhibit the highest and lowest mastery, and faculty are encouraged to reflect on their courses to determine if they can make any changes that could improve mastery of the objectives. This approach has been implemented. Second, the cluster committee examines the alignment between faculty coverage of objectives, student perceptions of objective coverage in classes, and performance on objectives on the assessment instrument. Faculty coverage of objectives and students perception of coverage are determined by surveys. This approach was implemented for the first time this past academic year. Finally, the Cluster Committee conducts an item analysis of the assessment instrument to determine which items students perform poorly on as sophomores/juniors. By identifying commonly chosen incorrect answers, the committee identifies common misconceptions, and those misconceptions are conveyed to the faculty, who are encouraged to reflect on their course to see if they can modify their courses to determine if these misconceptions can be dispelled. This approach is being implemented in the upcoming academic year.

Assessment results are used to assess the efficacy of program structure in Cluster Three. At the beginning of 2006-2007, a new structure for the cluster was implemented. The previous structure consisted of seven integrated groups or “packages” of courses, each containing three to four courses. Students chose a package and were required to complete that package. Different packages had different themes (e.g., environmental issues) or served different subpopulations of students (e.g., science majors, pre-service teachers). The package structure was originally developed to produce a set of integrated, interdisciplinary courses in which connections between disciplines were made explicit, with the hope that mastery of cluster objectives would be higher than in might occur in a more open structure (e.g., a “menu” approach). The impetus to reconceptualize the structure of the cluster arose from the logistical constraints of maintaining the packages in the face of rapid growth of the student body and faculty. In deliberations among the members of the cluster committee, the question arose as to whether tight integration did indeed improve learning outcomes compared to looser structures. A comparison of the assessment results among packages showed no relationship between the degree of integration and student performance. Although the original goal had been for each package to be tightly integrated, in reality, there existed variation among packages in the degree of integration. For some packages, the package was thoroughly integrated from the beginning, and integration and communication both within and between courses was maintained by regular meeting of faculty teaching the package courses. Other packages were began with much less integration and experienced little or subsequent communication among faculty. A pre-/post-test analysis show no significant difference among packages in the amount improvement in assessment scores. Furthermore, more integrated packages did not produce greater improvements than less integrated packages. For example, students in the tightly integrated package serving predominantly the integrated science and technology majors showed the same mastery of objectives and improvement in mastery as did students in the least integrated package, which served primarily math and natural science majors. Given these results, the committee decided that changing the package structure to the current, more flexible structure would solve the growing logistical problems without compromising student learning. This decision has been supported by performance on the assessment instrument: the average score for sophomores/juniors did not decrease under the new structure, as would be predicted if delinking the courses within packages resulted in poorer mastery of objectives. In fact the opposite has been observed: slight increase in student performance on the assessment instrument over the three years since the restructuring.

Assessment results have also been used to inform policy on the acceptance of advance placement (AP) and transfer credit. Assessment results consistently show that performance of students on the assessment instrument is positively correlated with the amount of AP credit earned in Cluster Three courses (credit is awarded if students receive sufficiently high scores on the AP exam). In contrast, there is generally either no relationship, or a negative relationship between the amount of transfer credit earned and performance on the assessment interest. These results provide support for two Cluster Three policies: 1) accepting AP credit for any cluster course for which the student has received the minimum score for the analogous AP exam, and 2) limiting to one the number of courses that can be transferred to meet cluster requirements for native (non-transfer) students.