University of Virginia
Jonathan Schnyer, Associate Director,
Office of Institutional Assessment and Studies (IAS)
Quantitative Reasoning 2008

PART A: PLAN[1]

1.0 Definition of Terms

1.1: *The central purpose of the University of Virginia is to enrich the mind by stimulating and sustaining a spirit of free inquiry directed to understanding the nature of the universe and the role of humans in it. A specific, articulated goal associated with this purpose is "fostering in students the habits of mind and character required to develop…an ability to test hypotheses and re-interpret human experience.*[2]

These habits of mind and character advance good citizenship in a democratic society, enrich the lives of individuals, and improve communities. The University expects graduating students to effectively use quantitative reasoning to evaluate information and argument, solve problems, and make decisions to these ends.

Quantitative reasoning is defined as "correctly using numbers and symbols, studying measurement, properties, and the relationships of quantities, or formally reasoning within abstract systems of thought to make decisions, judgments, and predictions."

1.2: The following standards have been established for graduating fourth-years:

25% of undergraduates are expected to be highly competent;
75% competent or above;
90% minimally competent or above.

Standards for gain between first-years and fourth-years will be considered after this first administration of the assessment to both first and fourth years. The criterion that will indicate competence for fourth-years is the overall score on UVa's quantitative reasoning test. A score of 11-15 indicates minimal competence, a score of 16-22 indicates competence, and a score of 23-30 indicates high competence.

---

[1] Note: The University's original plan, based on the reporting template established for the 2002-06 competency assessments, was redone to fit the format and outline suggested by Council in December 2007 and then revised in February 2008.

[2] Source: University of Virginia Statement of Purpose and Goals
http://www.web.virginia.edu/iaas/data_catalog/institutional/data_digest/goal.htm

1.3: A graduating fourth-year undergraduate at the University of Virginia will be able to:

(1) Interpret mathematical models such as formulas, graphs, tables, and schematics, and draw inferences from them.
(2) Communicate mathematical information symbolically, visually, numerically, and verbally.
(3) Use arithmetical, algebraic, geometric, and analytic[3] methods to solve problems.
(4) Estimate and check answers to mathematical problems in order to determine reasonableness.
(5) Solve word problems using quantitative techniques and interpret the results.
(6) Apply mathematical/statistical techniques and logical reasoning to produce predictions, identify optima, and make inferences based on a given set of data or quantitative information.
(7) Judge the soundness and accuracy of conclusions derived from quantitative information, recognizing that mathematical and statistical methods have limits and discriminating between association and causation.
(8) Solve multi-step problems.
(9) Apply statistics to evaluate claims and current literature.
(10) Demonstrate an understanding of the fundamental issues of statistical inference, including measurement and sampling.

2.0 Methodology

2.1 UVa's Quantitative Reasoning Assessment will employ a competency-based approach to determining the competence of our fourth-years and longitudinal and cross-sectional approaches to discerning value added.

2.2 The University of Virginia requires 12 hours of math or science as part of its area requirements for all schools except the School of Engineering, which has additional math requirements. Stringent admissions requirements for mathematics mean that first-year students have a strong mathematics foundation upon which the University builds in courses taken to fulfill area requirements and also in major courses. From the University's Course Offering Directory:

> *The faculty established area requirements to ensure that all students have the background and breadth for further learning in a variety of disciplines. In completing these requirements, students explore a wide range of disciplines, points of view, and modes of inquiry. In addition, they investigate unfamiliar areas and thus can make more informed judgments about their major and elective courses…*
>
> *Natural Sciences and Mathematics improve a student's comprehension of the fundamental principles of natural phenomena and of scientific methods as a way of describing and understanding the world.*

[3] After the plan was submitted to SCHEV in June 2007, the Committee (see section 2.3) made this change as the instrument was being developed.

Using longitudinal, cross-sectional, and competency-based approaches allows the University to study the questions of both value-added and competence and will provide a richer data set from which to examine the quantitative reasoning capabilities of our undergraduates.

2.3 The University developed a new instrument to assess the quantitative reasoning abilities of its students. The instrument was developed over several months by the Quantitative Reasoning Assessment Committee, comprised of ten faculty representing various disciplines and schools, and was pilot tested in the fall of 2007 on 50 students. While we were able to obtain information about the discrimination of the individual questions, and make some adjustments, there were not a sufficient number of cases in the pilot to examine the validity of the instrument through factor analysis. While verifying the instrument's reliability will need to take place over time, the questions were carefully developed by experienced faculty instructors and thoroughly considered and improved by the full committee, both before and after the pilot administration in the fall of 2007. Analysis did indicate a very high correlation with the Math SAT score.

2.4 There were two subpopulations from which the samples were drawn: first-time, first-year students entering in the fall of 2007 and fourth-year, fourth academic level students expected to graduate in the spring of 2008. The first-year students will be sampled again in their fourth years (2011) and will be used to assess value-added. The fourth-year student sample will be used to examine competence. The subpopulation sizes for first- and fourth-years are 3,248 and 3326, respectively.

2.6 The University employed a single setting data collection technique. A multiple choice instrument was administered over a period of 3 weeks in February 2008 to groups of approximately 60 students at a time. Students were initially invited by letter to participate, and several email reminders were sent out to reduce nonresponse bias. Those who agreed to participate were assigned to sessions in advance, and email reminders were also sent out to reduce the no-show rate. Of the random sample[4], 456 students signed up to participate (a 33% response rate to the initial invitation) and 82% of those who signed up actually took the test. Students were paid $20 for their participation and provided lunch, and a $100 bonus incentive was added for the top scorers in each major school and discipline.[5]

[4] In addition to these 456 randomly sampled students, another 135 fourth-years, oversampled from several schools and disciplines within the College of Arts and Sciences, were tested as well. The purpose was to have sufficient numbers of fourth-years in each school and each major discipline within the College to report results by school and major discipline within the College. These results are not part of this report but will be included as part of the report to the Provost and deans referred to in 3.2.

[5] The bonus was only offered to ½ of the tested students, as an experiment to see what effect these types of incentives might have on student performance. Results from this experiment were not available as of the writing of this report.

2.7 The measurement strategy employed was a selected-response test. Because of the number of learning outcomes to be assessed, faculty felt that the efficiency of a multiple choice format was required. A cross section of fourth-years was also tested to see a) what their level of competence is and b) how their scores compare to first-years.

2.8 The same measurement strategy will be employed for the post assessment. Gain scores will be computed for individual students and aggregated for the entire cohort. In addition, the mean scores and levels of competence will also be comparable.

2.9 The University will utilize a cross-sectional value-added approach for AY2007-08 and a longitudinal value-added approach in AY2010-11. The University will be able to draw inferences about value added from both the cross sectional comparison between first- and fourth-years in AY2007-08 and the pre-post comparison of a panel of first-year students when they become fourth years in AY2010-11. The University does not expect significant gain scores in the follow up post assessment. We also did not expect significant cross sectional difference scores between first- and fourth-years for students as a whole. Major areas that do not emphasize quantitative reasoning (humanities and fine arts) would not be expected to increase significantly in quantitative reasoning ability. Programs that emphasize specialized quantitative reasoning skills further complicate the assessment of quantitative reasoning from a value-added, institution-wide perspective. Students come to the University with strong quantitative skills, evidenced by their high SAT math scores (sample mean=675). All students have the opportunity to improve upon these skills through math/science courses under area requirements but many choose to focus on courses and programs that do not emphasize quantitative skills. For those who do choose fields that emphasize quantitative reasoning, the University wants to ensure that these students leave the University highly competent in quantitative reasoning. Students who do not choose to focus on improving their quantitative skills leave the University with other valuable and marketable skills learned in their major course of study.

3.0 Process Evaluation

3.1 A detailed report, with analysis and recommendations, will be authored by the Quantitative Reasoning Assessment Committee in the spring and summer of 2008 and submitted to the Provost, all deans of the undergraduate schools, all department chairs and all program and school assessment coordinators at the University (~60). A general report of results will be submitted to the State Council of Higher Education for Virginia in March 2008. Written and electronic versions of these reports will be disseminated directly to pertinent stakeholders. A public website describing the process and the general results has been established.
http://www.web.virginia.edu/iaas/reports/subject/competencies/quantitative.htm

3.2 Competency assessment results have been, and will continue to be, used to improve student learning at the University. While the vast majority of results from competency assessments have been positive, with the University meeting or exceeding its targets, there have been some areas targeted for improvement. Offering "instruction of the highest quality to undergraduates" is a central goal of the University, and core competency assessment results are part of the information stream that will inform efforts to meet that goal. If shortcomings are identified, the Assessment Committee will meet and discuss what the next steps should be at the institutional level. Schools and departments will receive copies of a detailed report to consider in their own discussions. This report will include not only the results of the institutional assessment but also the breakdowns for the schools and major disciplines within the College of Arts and Sciences as well as the committee's analysis of the meaning of the results and any recommendations.

3.3 The University estimates the cost of the development of the instrument and the administration to approximately 500 students to be $18,600. This estimate includes payments to faculty for instrument development, costs associated with inviting students to take the assessment, and administering the test (payments to students, scantron forms). Not included in the $18,600 is IAS staff time. The Associate Director spent approximately 25 hours facilitating committee work, overseeing a staff member's work, and drafting this report; his time equals $1300 in salary and benefits. The staff member spent approximately 165 hours researching instruments and questions, facilitating the drafting of questions, making corrections to and finalizing the pilot and final instruments, communicating with faculty committee members, proctoring and scheduling the test sessions, communicating with students and analyzing the data; her time adds up to $5,000 in salary and benefits. The estimated total cost of this administration, including the development of the instrument, was $24,900.

PART B: STATUS REPORT

4.0 Data Presentation

4.1 Getting students to respond to request for voluntary participation in assessments is always a challenge. We invited about 3X the number of students needed for the assessment and then followed up with email reminders to ensure the highest response rate possible. A number of sessions were scheduled at convenient times and students had their choice of which sessions suited their schedules. Email confirmations were sent out after sessions were assigned and reminders were sent the day before and the day of the sessions to reduce the number of no-shows. Problems of non-response bias are always a concern; therefore analysis was conducted of the demographic representation of the participants, compared to the non-participants.

First-year sample. Five hundred and thirty-five (535) randomly selected first-year students were invited to participate in the quantitative reasoning assessment. Of those, 178 first-year students completed the quantitative reasoning assessment (33 were no-shows). The final sample of 178 first-year students had a higher female to male ratio than that which exists in the population (see Table 1). Accordingly the data were weighted by gender to account for these differences in the sample versus the population. The sample had some differences of representation in terms of race (African-Americans were a bit underrepresented and Asian students were a bit overrepresented) but by and large the racial makeup of the sample looked similar to the population. We did not weight the data by race in addition to gender because the overall makeup of the races looked similar to the population, and we did not feel weighting would impact the results sufficiently to justify the added complication.

Table 1. First-Years by Gender

	Sample n	% of Sample	Population N	% of Population	Weighted n
1st-year female	118	66.3%	1,849	57%	101
1st-year male	60	33.7%	1,399	43%	77
TOTAL	**178**	**100%**	**3,248**	**100%**	**178**

Fourth-year sample. Eight hundred and sixty-two (862) randomly selected fourth-year students were invited to participate in the assessment. Of those, 194 completed the assessment (51 were no-shows). The final sample of 194 fourth-years had a higher female to male ratio than that which exists in the population (see Table 2) so again the data were weighted by gender. Racial representation was similar to first-years and the same decision was made to not weight by race. Representation by school could also be ascertained for fourth-years and was similar to the population.

Table 2. Fourth-Years by Gender

	Sample n	% of Sample	Population N	% of Population	Weighted n
4th-year female	115	59.3%	2001	53.8%	104
4th-year male	79	40.7%	1718	46.2%	90
TOTAL	**194**	**100%**	**3,719**	**100%**	**194**

4.2 It is important to state that UVa's quantitative reasoning test was a challenging one—perhaps too challenging. No students answered more than 27 of the 30 questions correctly. The test asked a wide variety of questions to ascertain student achievement of ten distinct learning outcomes. Developing such an instrument is a major undertaking and while IAS and the Assessment Committee worked diligently to design an instrument that was valid, reliable and informative, the effectiveness of such an instrument will need to be carefully examined and improved over time. Moreover, the learning outcomes that the committee has defined will need to be thoroughly vetted with and disseminated to faculty teaching courses that develop students' quantitative reasoning skills. The University sees this assessment as the continuation of a long term process to address where, how and what we teach our undergraduate students about quantitative reasoning.

Cross-Sectional, Value-added Results. The quantitative reasoning performance of the first-year sample was compared to the performance of the fourth-year sample. There were no mean differences in quantitative reasoning performance between the two groups. However, fourth-year science majors (M=20.27) scored higher than 1st years (M=18.32) $p<.01$ (see Table 3).To understand the implication of these results, the baseline quantitative reasoning skill of all groups needs to be considered. The first-year mean SAT Math score is 678.25, which is 163.25 points higher than the national average. The average mean SAT Math score for fourth-years was not significantly different from first-years. These cross-sectional results indicate that on the whole fourth-year students at UVa maintain high levels of quantitative skill, while science majors gain significantly higher levels of quantitative skill while at the University.

Table 3. Mean Scores on Quantitative Reasoning Test

	Mean Score	Standard Deviation	Minimum	Maximum
1st years	18.60	4.87	2.00	27.00
4th years (all)	18.75	4.97	4.00	27.00
4th-year science	20.50	4.56	6.00	27.00

Competence. The competency standards that were established for graduating fourth years were: "25% of undergraduates are expected to be highly competent; 75% of undergraduates competent or above; and 90% minimally competent or above." Minimal competence was defined as a score of 11-15 on the quantitative reasoning instrument (answering greater than 1/3 of questions correctly), competence was defined as a score of 16-22 on the instrument, and high competence was defined as a score of 23-30 (answering greater than 3/4 of questions correctly). Fourth-year students slightly exceeded the standard for minimal competency with 92.8% of fourth-years achieving minimal competence or above. Fourth-years also slightly exceeded the standard for competence with 77.9% achieving competence or above. Fourth-year students met the standard for high competence with 25.3% achieving high competence or above. However, when the performance of fourth-year science majors was highlighted, the data show that science majors *significantly* exceed the competency standards (see Table 4).

Table 4. Fourth-Year Quantitative Reasoning Competency

	Minimally Competent	**Competent**	**Highly Competent**
4th Years (all)	92.8%	77.9%	25.3%
4th –year Science	97.2%	87.6%	36.9%
Standards for Competency (see Section 1.2)	**90%**	**75%**	**25%**

4.3 After careful consideration of the results, the University may seek additional data from faculty or students about the quantitative reasoning skills of its students. This could take the form of surveys, focus groups, or additional assessments.