

**UNDERGRADUATE WRITING  
AT THE UNIVERSITY OF VIRGINIA**

A Core Competency Assessment Report

Submitted by the Writing Assessment Committee

Coordinated by the Office of Institutional Assessment and Studies  
Report written by IAS staff (Lois Myers and Sarah Schultz Robinson)

March 15, 2010

Writing Assessment Committee:

Liz Arkush, Anthropology

Carl Trindle, Chemistry

Scott Deveaux, Music

Jon D'Errico, English

Greg Colomb, English

Claire Chantell, English

Ryan Cordell, English

Josipa Roksa, Sociology

Dorothe Bach, Teaching Resource Center

Deandra Little, Teaching Resource Center

## TABLE OF CONTENTS

Executive Summary	2
Assessment Process and Methodology	3
Summary of Results	5
Committee Findings	9
Recommendations	11
Appendices	13
A- Rubric for CLAS, Architecture and Nursing Assessment	15
B- Rubric for FWR Assessment	18
C- Inter-rater Reliability	20
D- Representativeness Analysis	21
E- Mean Scores with and without FWR-exempted	23
F- Mean Scores by Discipline	24
G- Mean Scores pre- and post-FWR	25
H- Mean Scores FWR-participants vs. Exempted	27

## EXECUTIVE SUMMARY

Five separate assessments constitute the overall assessment of undergraduate student writing that was conducted in 2008-2009 by the Writing Assessment Committee. This report describes results primarily from the assessment of writing by fourth-year students from the College (with some students from the Schools of Architecture and Nursing), although it also incorporates findings from the assessment of first-year students participating in the First Writing Requirement (FWR). The assessments were designed to address three sets of questions:

1. How well do fourth-year students write? Where is their writing strongest? Weakest?
2. How well do fourth-year students write as compared to first-years?
3. How much do students who took FWR courses improve their writing, and how do they compare to students who were exempted from FWR?

Except for the pre-post FWR course assessment, the comparisons were all cross-sectional (first-year students *vs.* fourth-year students).

Results suggest that fourth-year students are not as competent in writing as expected, although their writing shows a statistically significant improvement over first-year students' writing. Students who took FWR improved their writing, and participation in the FWR course helped to close the gap between students who took FWR and students who were exempted from FWR.

The committee recommends 1) requests for more information about writing instruction post-FWR, 2) a re-examination of the learning outcomes, 3) a true pre-post (rather than cross-sectional) study with a larger sample, 4) an investigation of current support for writing instruction, and 5) completion of the online, self-guided writing support program for students.

## I. ASSESSMENT PROCESS AND METHODOLOGY

In spring 2008, the Office of Institutional Assessment and Studies (IAS) initiated the assessment of undergraduate student writing competency. In collaboration with faculty in the English Department and the Assistant Director of the Writing Center, IAS developed an assessment plan that would both examine the overall writing competency of undergraduates and allow for individual schools and the Writing Program to assess discipline-specific writing. The University mandates that all undergraduates meet a two-part writing requirement before graduation: students must 1) either place out of or complete a First Writing Requirement (FWR) course, and 2) take a Second Writing Requirement (SWR) writing-intensive course. Three types of FWR courses are offered through the Writing Program (introductory; ESL; and accelerated). SWR courses are offered within the disciplines.

All schools were offered the opportunity to conduct discipline-specific assessments. Faculty from the School of Commerce, the School of Engineering and Applied Sciences, the Bachelor of Interdisciplinary Studies in the School of Continuing and Professional Studies (BIS/SCPS), and the Writing Program in the College of Arts and Sciences opted to work with IAS to plan separate assessments. The remaining schools—Architecture, Nursing, and the College of Arts and Sciences—were represented in a combined assessment. The five assessments included:

1. Commerce- an assessment of student performance pre- and post-writing course using a scoring rubric reflective of the Integrated Core Experience writing curriculum;
2. Engineering- a cross-sectional comparison of first-year and fourth-year student writing in Science, Technology, and Society (STS) courses using a scoring rubric reflective of the STS curriculum;
3. BIS/SCPS- a cross-sectional comparison of student writing in introductory and capstone courses, employing a scoring rubric reflective of the BIS curriculum;
4. College/Architecture/Nursing- a cross-sectional comparison of first-year and fourth-year student writing and a comparison of fourth-year students' actual performance to expected performance, applying a scoring rubric based on learning outcomes affirmed by the Writing Assessment Committee;
5. Writing Program/First Writing Requirement (FWR)- an assessment of student performance pre- and post-FWR writing courses employing a scoring rubric reflective of the FWR curriculum.

This report primarily describes the process, results, and findings from the assessment of writing by students in the College/Schools of Architecture/Nursing; it also incorporates information from the Writing Program/FWR assessment. The report closes with a set of recommendations.

The Writing Assessment Committee convened in August 2008 to plan an assessment and delineate criteria by which to assess writing by students from the College and the Schools of Architecture and Nursing. To evaluate overall competency and improvement in competency (“value-added”), the Committee decided to solicit and assess first- and fourth-year student papers submitted to fulfill course requirements. The Committee crafted a set of seven learning outcomes, created a scoring rubric based on those outcomes (Appendix A), and set targets for achievement. Student papers were to be scored on a scale of 1-4 (not competent, minimally competent, competent, highly competent) for each of the seven learning outcomes:

1. Introduction
2. Parts of Argument I-Claims and Sub-claims
3. Parts of Argument II-Evidence

4. Counterarguments
5. Cohesion and Coherence
6. Audience and Tone
7. Conclusions

The committee established the following standards for competence:

- 40 percent of students would be highly competent (avg. total score=3.5 or above)
- 85 percent of students would be competent or better (2.5 or above)
- 100 percent of students would be minimally competent or better (1.5 or above)

Simultaneously, IAS and Writing Program faculty planned a “pre-post” assessment of student learning in the FWR courses. For fifty students taking FWR, papers from the beginning of the course and the end of the course would be assessed and the results compared. Writing Program faculty designated four overall learning outcomes and created a scoring rubric to facilitate assessment (Appendix B). Some learning outcomes were scored on a scale of 1-4 while others were rated as “present” or “absent”.

In August 2008, IAS created a stratified<sup>1</sup> random sample of FWR courses from which first-year student papers were to be collected. At a meeting convened by IAS, course instructors were given an overview of the two assessments—their rationale and methodologies—and asked to collect and share all papers that had been written to fulfill the first writing assignment of the course. In September, the instructors submitted 179 student papers to be evaluated for the College/Architecture/Nursing assessment. Fifty of these papers were randomly selected to serve as the sample of “pre-FWR” papers. At the end of the FWR course, final papers from these same 50 students were collected to serve as “post-FWR” papers.

IAS also created a stratified random sample of fourth-year students, taking into account gender, ethnicity, GPA, Verbal SAT score, FWR-exempt status, and school or discipline within the College. In December 2008, an initial letter from Milton Adams, Vice Provost for Academic Programs, invited students to submit an example of their best writing for the assessment. The invitation included the seven learning outcomes and asked students to submit 10-15-page papers that best fit the learning outcomes and that had been written in their fourth year. Students were promised confidentiality. Over the next four months, four email reminders were sent. Each student who contributed a paper was entered into a lottery to win one of seven \$99 gift certificates to Barnes and Noble bookstore. Overall, 424 students submitted papers, and from those, 84 percent were accepted for assessment. The remaining 16 percent did not fit the criteria (e.g., were “hard” science papers, case studies, experimental research papers, etc.). Less than five percent of papers were from students in Architecture or Nursing.

All papers were de-identified to protect student confidentiality.

During seven work sessions in spring term 2009, 22 instructors assessed student work using the scoring rubric. Each session began with a norming exercise to assure that evaluators were in accord in the application of the rubric. Papers were randomly assigned to evaluators. Evaluators rated each paper according to the seven learning outcomes contained in the scoring rubric, giving each outcome a rating from 1 (not competent) to 4 (highly competent). Two evaluators scored each student paper,

---

<sup>1</sup> Stratified for two factors: level of FWR course (3 levels offered: Introduction; Introduction-ESL; Accelerated) and for instructors’ years of experience.

and the two evaluators' scores for each learning outcome and for the overall rubric were averaged to create scores for each paper.

Inter-rater reliability was moderate for both assessments. For the College/Architecture/Nursing assessment, reliability ranged from 0.41 to 0.61, averaging 0.54.<sup>2</sup> For the Writing Program/FWR assessment, inter-rater reliability ranged from .01-.62, averaging .31 on a scale of zero to one.<sup>3</sup> (Appendix C.).

Reflecting on the assessment process, the Committee observed that:

- In addition to the 424 fourth-year students who submitted papers, 36 expressed interest in participating but had not written a 10-15 page paper in their fourth year.
- The rubric measured what the Committee intended it to (validity) and was applied in a consistent manner (reliability).
- Some of the papers did not fit or satisfy all of the learning outcomes. This was particularly true of the learning outcome, “counterarguments”: nearly three-quarters of papers included little or no consideration of counterarguments.
- First-year papers and fourth-year papers may not have been fully comparable. The difference in length (avg. 3-4 pages for first-years *vs.* avg. 10-15 pages for fourth-years) probably reflects substantive differences in assignment and expectations.

## II. SUMMARY OF RESULTS

### *Study sample*

Compared to the population of fourth-year students, the fourth-year sample contained a disproportionate number of students with higher GPAs (mean GPA for sample=3.33 *vs.* 3.18 for the population) and higher verbal SAT scores (673 for the sample *vs.* 651 for the population). (Appendix D). Since these variables correlated positively with writing performance, their over-representation may have affected the assessment results. Also over-represented in the fourth-year sample were women (71% in sample *vs.* 62% in population), non-minority students (70% *vs.* 65%), and humanities/fine arts majors (43% *vs.* 36%). Gender, minority status, and discipline, however, were *not* significantly related to writing performance.

The first-year student sample consisted of all students enrolled in a stratified random sample of FWR courses and thus is assumed to be representative of students who participated in FWR.

### *Student performance*

Fourth-year students fell below the committee's standards for writing competence: 8 percent were highly competent, 61 percent were competent or above, and 99 percent were minimally competent or above (Table 1).

---

<sup>2</sup> Intra-class Coefficient (ICC): scale of 0 to 1 with zero indicating little or no agreement and one indicating perfect agreement.

<sup>3</sup> Unlike the other rubrics, the FWR rubric contained numerous dichotomous items; the presence of dichotomous variables tends to lower reliability scores, hence the lower reliability for FWR.

Table 1. Fourth-Year Writing Competence (CLAS, Architecture, Nursing)

	Minimally Competent or above	Competent or above	Highly Competent
All Fourth-Years	99%	61%	8%
<b>Standards for Competency</b>	<b>100%</b>	<b>85%</b>	<b>40%</b>
N	351	216	28

Students' scores did not reach competency (an average score of 3) for any single learning outcome (Table 2).

Table 2. Mean Score for Each Learning Outcome—Fourth Year

	<b>4<sup>th</sup>-Years Mean Score</b>
Introduction	2.52
Parts of Argument I- Claims and Sub-claims	2.68
Parts of Argument II- Evidence	2.82
Counterarguments	1.97
Cohesion and Coherence	2.93
Audience and Tone	2.85
Conclusions	2.75
<b>Final Score</b>	<b>2.64</b>
N	355

Students did perform better on some learning outcomes than on others: not surprisingly, students scored the best on Cohesion and Coherence, followed by Audience and Tone and Parts of Argument-Evidence. Students scored the lowest on Counterarguments. Even among higher performing students<sup>4</sup>, Counterarguments yielded the lowest scores. It is possible that a sizeable percentage of students did not address counterarguments in their papers at all: 39 percent of students received a rating of 1, which was defined as “Either does not acknowledge any counterargument or responds to them in an excessively combative or hostile manner” (Table 3). One third received a rating of 2, which suggests that these students did at least acknowledge counterarguments (“Acknowledges counterarguments, but does not fully explain them. Responds to counterarguments in a cursory manner, or combatively rather than dialogically.”). Even when Counterarguments was removed from the Final Score, the average performance of fourth-year students still fell below “competent” (2.64 with Counterargument vs. 2.76 without Counterargument).

<sup>4</sup> For analysis purposes, fourth-year students were divided into three relative groups based on final scores—“low” performers (n=122), “middle level” performers (n=113), and “high” performers (n=120).

Table 3. Percentage of Ratings by Score for Each Learning Outcome – Fourth-Year

	1.00	2.00	3.00	4.00
Introduction	8.0%	46.0%	32.0%	14.0%
Parts of Argument I – Claims & Sub-claims	6.0%	35.0%	43.0%	16.0%
Parts of Argument II - Evidence	3.0%	35.0%	41.0%	21.0%
Counterarguments	39.0%	33.0%	21.0%	7.0%
Cohesion and Coherence	1.0%	24.0%	55.0%	20.0%
Audience and Tone	4.0%	30.0%	45.0%	21.0%
Conclusions	5.0%	35.0%	41.0%	19.0%

Comparing high performing (top third) students with low performing (bottom third) students, lower-performing students received on average consistently poorer scores across all learning outcomes. The differences were especially large for Introduction, Evidence-Claims/Sub-claims, Audience/Tone, and Conclusions (Table 4).

Table 4. Mean Scores of Low and High Performers

	Low	High	High - Low Difference	Effect-Size
Introduction*	1.91	3.20	1.29	0.59
Parts of Argument I – Claims & Sub-claims*	2.02	3.33	1.31	0.64
Parts of Argument II - Evidence*	2.22	3.38	1.16	0.53
Counterarguments*	1.43	2.63	1.2	0.39
Cohesion & Coherence*	2.45	3.41	0.96	0.52
Audience & Tone*	2.19	3.45	1.26	0.62
Conclusions*	2.11	3.35	1.24	0.60
<b>Final Score*</b>	<b>2.05</b>	<b>3.25</b>		
N	122	120		

\* Mean differences in performance are significant at  $p < .001$ .

Comparing first-year and fourth-year students' writing, **fourth-year students ( $M=2.64$ ) significantly outperformed first-year students ( $M=2.02$ ) on all learning outcomes and overall (Table 5).**

Table 5. Mean Differences between First- and Fourth-Years

	<b>1<sup>st</sup>-Years Mean Score</b>	<b>4<sup>th</sup>-Years Mean Score</b>
Introduction*	2.03	2.52
Parts of Argument I- Claims and Sub- claims*	1.98	2.68
Parts of Argument II- Evidence*	2.04	2.82
Counterarguments*	1.73	1.97
Cohesion and Coherence*	2.25	2.93
Audience and Tone*	2.04	2.85
Conclusions*	2.05	2.75
<b>Final Score*</b>	<b>2.02</b>	<b>2.64</b>
N	179	355

\*Mean difference is significant at  $p < .001$

This result may overestimate the difference since the fourth-year sample contained both students who placed out of and who had participated in FWR whereas the first-year comparison group contained only FWR-participants. To correct for this potential bias, the analysis was repeated with fourth-year FWR-place-outs deleted. This analysis confirmed the original finding: fourth-year students ( $M=2.54$ ) still significantly outperformed first-years ( $M=2.02$ ) on all learning outcomes and overall. (Appendix E).

Among fourth-year students in the College, there were no significant differences in writing performance among humanities/fine arts majors, science/math majors, and social science majors after controlling for GPA and verbal SAT score (Appendix F).

Students who participated in FWR showed significant improvements in writing performance for 11 of the 16 rubric items across all four learning outcomes. On five of the rubric items that relate to Motivation, Style, and Core Elements of Argument, students' performance did not improve significantly (Appendix G).

A final analysis compared the performance of fourth-year students who had completed the FWR and those who were exempted from taking FWR. Students who were exempted had significantly higher Verbal SAT scores and GPAs than did students who participated in FWR (These two sets of scores, in part, determine eligibility for exemption.). After controlling for Verbal SAT score and GPA, the overall mean score of exempted students was 2.72 and that for FWR participants was 2.60. This difference was not statistically significant, although the difference did approach significance ( $p=.07$ ). Students who were exempted significantly outperformed FWR participants on two learning outcomes: Parts of Argument—Evidence ( $p=.02$ ), and Audience/Tone ( $p=.01$ ). (Appendix H).

### III. Committee Findings

This assessment was designed to address three key questions:

1. How well do fourth-year students write?
2. After fulfilling the two-part writing requirement and completing four years of study at UVa, how much has student writing improved? (value added)
3. How well do students who took FWR courses write as compared to students who were exempted from FWR?

The assessment was sufficiently robust to provide useful information regarding undergraduate students' writing competency and improvement in competency. The representativeness of the fourth-year sample, however, does suggest caution: the sample contained a larger proportion of students with higher GPAs and verbal SAT scores than occur in the University population as a whole. The differences were statistically significant but substantively small. While small, this overrepresentation would tend to yield artificially higher scores for fourth-year students.

#### 1. How well do fourth-year students write?

The assessment results suggest that the undergraduate program fosters, on average, competent writers, but not highly competent writers. The overall score for fourth-year students' writing averaged 2.64 (on a scale of 1-4); this was toward the low end of the range designating competency (2.50-3.49). Fewer than 1 in 10 fourth-year students were judged highly competent (3.5-4.0); only six in 10 were judged competent or better (2.50-4.0), leaving about 4 in 10 judged less than competent (incompetent to minimally competent: 1.0-2.49). The Committee had expected 40 percent of fourth-year students to be rated as highly competent and 85 percent as competent or better.

How do we explain these results?

- First, it was not a matter of one or two disciplines pulling down the average scores. On the whole, fourth-year students across the discipline groups in the College of Arts and Sciences (humanities/fine arts, social sciences, science/math) wrote equally well.
- Fourth-year students' papers may not have provided the strongest evidence of student writing as defined by the learning outcomes. While students were asked to submit papers that reflected both their best writing and the learning outcomes, the papers probably represented a wide variety of assignments, some of which may not have corresponded well to the learning outcomes. This variability would tend to lower average final scores. However, even for specific learning outcomes that we would expect to be widely-supported—such as cohesive/coherent writing—student work did not meet expectations (only 20% highly competent; 75% competent or better).
- Students may not have submitted their best writing, but rather their favorite paper, most recent paper, most easily located paper, or the first paper that came to mind. This possibility would also have contributed to lower average scores.
- One or more of the learning outcomes proffered by the Committee may not be widely shared and supported by instructors. One learning outcome in particular—counterargument—stood out for its low performance (mean score of 1.97 is less than competent; only 28 percent of students rated competent or better). Instructors may not ask students to address counterargument in their papers. Even if we remove counterargument from the analysis, however, student performance did not rise to expectations.

- Fourth-year students believe that they are good to excellent writers. In 2008<sup>5</sup>, fourth-year students were asked to rate their own writing ability on a scale of 1 (poor) to 4 (excellent); their self-ratings averaged 3.5 (Table 6). Fifty-five percent considered themselves to be excellent writers, with 97 percent rating their proficiency as good or better. Moreover, they believe that their writing has improved considerably: they assessed their writing as incoming first-year students as good (avg. score=3.0). This self-assessment should be interpreted with caution as students may over-estimate their skills<sup>6</sup>. This disconnect between students' self-assessment and the results of this competency assessment, however, may suggest that instructors in the disciplines are holding students to a different standard than ENWR instructors do, that students are receiving the message that they are good to excellent writers, and/or that students have not internalized ENWR writing standards.

Table 6. Fourth-year Students' Self-Assessment re: Writing

Ability to write clearly and effectively:	Poor	Fair	Good	Excellent	Total responses	Mean (range 1-4)
When started at UVa	1.6%	19.5%	54.6%	24.2%	706	3.0
As fourth-years	0.1%	3.4%	41.7%	54.8%	703	3.5

- The Committee's stated expectations were not grounded in evidence. Given that this was the first administration of the assessment, the standards may need to be re-examined now that data can inform the discussion of expected standards.

Even taking these considerations into account, the results for fourth-year students are concerning. More than 61 percent of students should graduate with writing skills judged "competent" or better. More than 8 percent should graduate as "highly competent writers."

## 2. After fulfilling the two-part writing requirement and four years of study at UVa, how much have students improved their writing?

Fourth-year students showed a statistically significant, but substantively small to moderate improvement in competency over first-year students, regardless of whether they were exempted from FWR or successfully completed FWR and regardless of discipline. Both first-year and fourth-year students were weakest in the area of counterarguments, and the difference between first- and fourth-year performance was smallest for this learning outcome.

For three reasons, the improvement in competency may overestimate actual improvement:

- Some of the improvement may be attributable to the small but statistically significant overrepresentation of those with higher GPAs and verbal SAT scores in the sample of fourth-year students.
- The first-year students were all FWR participants even though the comparison fourth-year group contained both FWR participants and FWR-exempt students. When first-year

<sup>5</sup> Undergraduate First and Fourth Year Student Survey, 2008:

[http://www.web.virginia.edu/iaas/surveys/past/2008-09/under09\\_1st4thyear.htm](http://www.web.virginia.edu/iaas/surveys/past/2008-09/under09_1st4thyear.htm)

<sup>6</sup> Richard Aurm and Josipa Roksa. 2010. "Learning During Unsettled Times: College Student Academic Performance and Recent Experiences in the U.S." Research Committee on Social Stratification and Mobility (RC28), International Sociological Association, Haifa, Israel (May)

students are compared to FWR-participant fourth-year students, the improvement is still statistically significant but is slightly smaller.

- The first-year papers, being shorter (3-4 pages), may have been written in response to a less demanding assignment than the fourth-year papers (10-15 pages).

### 3. How well do students who took FWR courses write as compared to students who were exempted from FWR?

As reflected in comparison of scores pre- and post-FWR, students who participated in FWR improved their writing for a majority of the criteria. But did the FWR experience raise students' performance to that of FWR-exempted students? Students with high verbal SAT (or comparable test) scores are exempted from taking FWR; all others are required to take and pass FWR. As such, comparing FWR-participants' and FWR-exempt students' performance is a little like comparing apples and oranges, and indeed, initial analysis found a statistically significant difference in performance. This SAT/GPA factor can be controlled for, however, through statistical analysis so as to arrive at a truer comparison. When SAT/GPA was controlled for, the difference between the two groups was no longer statistically significant, although it *approached* significance. This finding suggests that participation in FWR does help to close the gap between FWR-participants and FWR-exempt students.

## IV. Recommendations

The success of the writing requirement rests on two assumptions: 1) the FWR is effective in ensuring that first-students students who matriculate with lesser skills learn to write as well as their peers who placed out of FWR, and 2) the SWR is effective in improving students' competency beyond FWR levels. While the results of this assessment affirm both assumptions, it's a weak affirmation.

Taken together, the results of this assessment portray a "mixed bag." The FWR does appear to reduce the gap between FWR-participants and FWR-exempt students. Beyond that, on average, students appear to make only modest progress for the next three years. The results are in the right direction—fourth-year students perform better than first-years—but it appears that, by their fourth year, over a third of students do not achieve competency. Fewer than 10 percent write well enough to be considered highly competent according to this assessment. In contrast, nearly all fourth-year students surveyed report that their writing is good to excellent. Are students internalizing different standards than those applied in this assessment, not receiving feedback and thus applying their own standards, or just inclined to high self-regard?

These results are a wake-up call. They suggest that either students receive little instruction/ feedback regarding their writing after FWR, or the standards and instruction they receive differ from those supported by ENWR, or both.

In response to these results, the Committee raised three key questions, followed by seven recommendations:

1. What writing standards are applied throughout the academic curricula, and how do students measure up?
2. How much writing instruction is provided beyond FWR, and how can it be enhanced?
3. How can we make writing instruction a priority?

**Recommendation 1:** Solicit information from faculty members who teach SWR courses and those who supervise students writing capstone projects or distinguished major theses: to determine:

- Do instructors consider it their responsibility to teach students to write?
- What instruction and direct feedback do they provide to students?
- Why is instruction provided?
  - To improve students' ability to write within the discipline?
  - To help students hone their writing skills?
- Do they use the same criteria as were applied in the writing competency assessment? If not, how do criteria differ?
- What do assignments require? Exposition? Argumentation?
- Is writing quality a factor in grading?
- What is their perception of students' writing competency?
- What training have they received in writing instruction? What training or support would they find useful?

**Recommendation 2:** Solicit information from students regarding SWR experience:

- How much writing instruction did they receive?
- How much feedback about their writing did they receive?
- How much did the instructor value (weigh) writing when giving grades?
- Are they better writers after SWR than they were before? If yes, how?

**Recommendation 3:** Identify, document, and learn from academic programs (e.g., Psychology) that have adopted learning outcomes regarding writing and implemented curricula and pedagogy to support them. Over half of programs in the College cite writing as a learning outcome in assessment plans, although the learning outcomes vary widely (e.g., “write persuasively” vs. “articulate an argument” vs. “write publishable paper”). Do the programs promulgate writing standards and foster writing instruction across courses? If yes, how?

**Recommendation 4:** Convene a panel to:

- Re-examine, confirm, or, if necessary, revise the learning outcomes. Are these the writing skills that undergraduates should possess upon graduation? If not, what should be expected?
- Address these questions: Who assumes responsibility for teaching students to write? What are their goals? How do they address their goals? The University is home to multiple answers to these questions, ranging from ENWR to the discipline-specific writing instruction provided by Psychology.

**Recommendation 5:** Using the confirmed or revised learning outcomes, verify these assessment results by conducting a longitudinal (pre-post) competency assessment, beginning with students matriculating in Fall 2011 and following up with assessments of the same students' work in 2014-

2015. Increase the sample to 300 students and include the ability to control for those who were exempted from FWR and for other demographic factors such as GPA and verbal SAT scores.

**Recommendation 6:** Research available and potential vehicles for faculty professional development in the area of writing instruction. The TRC offers a limited number of workshops on writing instruction. Are they well attended? Can we create a program, modeled on the Teaching Fellows program, to include pedagogical and curricular support for student writing? Investigate programs at other institutions (e.g., George Mason University and University of Illinois) that provide significant support for teaching writing throughout the curriculum.

**Recommendation 7:** Complete development of the online self-guided tutoring program. This program, intended to supplement classroom instruction and Writing Center services, presents principles of argumentation, with illustrations and opportunities for practice and feedback. Once the program is available, it will provide curricular content in a highly flexible format so that students can learn content as needed and at their own pace. Make the program available both to students for personal use and to instructors as a tool for reinforcing writing instruction.

## APPENDICES

- A- Writing Assessment Rubric for College of Arts and Sciences and Schools of Architecture and Nursing
- B- Writing Assessment Rubric for First Writing Requirement (FWR)
- C- Analysis of Inter-rater Reliability
- D- Representativeness of the Fourth-year Sample
- E- Mean Writing Assessment Scores with and without FWR Exempted Students
- F- Writing Assessment Performance by Discipline
- G- Mean Writing Assessment Scores pre- and post-FWR
- H- Writing Performance by FWR Participants vs. FWR Exempted Students

## **APPENDIX A**

### **COLLEGE OF ARTS AND SCIENCES AND SCHOOLS OF ARCHITECTURE AND NURSING WRITING ASSESSMENT RUBRIC**

#### **Introduction**

4 - Introduction clearly explains and proposes to solve a problematic attitude, idea, or practice in the writer's field or subject of study. The proposed solution (the thesis or claim) is innovative, compelling, and convincing, motivating reader interest in both the paper and the larger field or subject.

3 - Introduction explains and proposes to solve a problematic attitude, idea, or practice in the writer's field or subject of study. The proposed solution (the thesis or claim) motivates reader interest in the paper.

2 - Introduction gestures toward a problematic attitude, idea, or practice in the writer's field or subject of study, but the problem is implicit rather than explicitly stated. Proposed solution (the thesis or claim) is unclear, unfocused, or too simplistic to satisfactorily address the problem.

1 - Introduction demonstrates confusion or misunderstanding about the writer's field or subject of study. Introduction provides no context for readers, is vague, and proposes no solutions (no thesis or claim).

#### **Parts of Argument I – Claims and Sub-claims**

4 - Argument is well balanced, with specific, insightful, debatable claims and sub-claims. Ideas progress in a logical sequence, work to support a clear structure, and claims are placed at appropriate intervals.

3 - Argument is balanced, with specific claims. Ideas progress in a logical sequence and claims are placed at appropriate intervals.

2 - Writers argument is unwieldy; some claims are implicit, unclear, or nearly indisputable. Makes some generalizations without support.

1 - Argument is unbalanced or impossible to identify. Claims are either entirely unclear or indisputable. Frequently makes illogical generalizations without support.

#### **Parts of Argument II – Evidence**

4 - Claims are almost always supported by precise, authoritative, and varied evidence.

3 - Claims are usually supported by authoritative and varied evidence.

2 - Evidence is sometimes insufficient, unreliable (unauthorized sources, anecdote, etc.), or only loosely connected to claims.

1 - Claims are usually unsupported by evidence.

#### **Counterarguments**

- 4 - Acknowledges and sufficiently explains counterarguments, responding thoroughly and convincingly through dialogue rather than verbal combat.
- 3 - Acknowledges and explains counterarguments, responding through dialogue rather than verbal combat.
- 2 - Acknowledges counterarguments, but does not fully explain them. Responds to counterarguments in a cursory manner, or combatively rather than dialogically.
- 1 - Either does not acknowledge any counterarguments or responds to them in an excessively combative or hostile manner.

### **Cohesion and Coherence**

- 4 - Writing is cohesive and coherent: information flow within sentences and paragraphs is logical and consistent. Sentences are strong, expressive, and varied in construction. Prose is stylistic and compelling. Grammar, diction, and spelling are nearly perfect.
- 3 - Writing is cohesive and coherent: information flow within sentences and paragraphs is logical and consistent. Sentences are clear, but may be formulaic or tedious. Document contains some common errors in grammar, diction, and spelling.
- 2 - Writing is fragmented: information flow between sentences and paragraphs is inconsistent. Sentences demonstrate little or no variety in style. Syntax may be irregular, and the document may contain persistent errors in grammar, diction, and spelling that hamper meaning.
- 1 - Writing is incoherent and fragmented. Problems with syntax create barriers to reader understanding, and the document contains pervasive errors in grammar, diction, and spelling.

### **Audience and Tone**

- 4 - Exhibits a thorough understanding of the goals, readers, situation, purpose, and structure of their argument, and writes in a style appropriate to each. Authorial tone is consistent, mature, and engaging. The language appropriately academic.
- 3 - Exhibits a solid understanding of the goals, readers, situation, purpose, and structure of their argument. Authorial tone is consistent, though perhaps unrefined or static. Language is appropriately academic.
- 2 - Exhibits an inconsistent understanding of the goals, readers, situation, purpose, and structure of their argument. Authorial tone is uneven or immature. Language is occasionally nonacademic.
- 1 - Exhibits little or no understanding of the goals, readers, situation, purpose, and structure of their argument. Argument is severely hampered by pervasive stylistic problems. Authorial tone is unbalanced and immature. Language is frequently nonacademic and distracts from the argument.

### **Conclusions**

- 4 - Conclusions are logical, convincing, clearly expressed, and consistent with those proposed in the introduction.
- 3 - Conclusions are logical and consistent with those proposed in the introduction.

2 - Conclusions are logical, but either unconvincing, not clearly expressed, or inconsistent with those promised in the introduction.

1 - Conclusions are not drawn, indecipherable, illogical, or inconsistent with the information presented in the rest of the document.

## APPENDIX B

### FIRST WRITING REQUIREMENT (FWR) WRITING ASSESSMENT RUBRIC

- 1 [Motivation] Is there a canonical problem statement? (Status Quo, Destabilizing Moment, Costs & Consequences)  
1 Yes (If this answer is "Yes," please automatically mark "Yes" for rubric item 2.)  
0 No (If this answer is "No," please read and respond to rubric item 2.)
- 2 [Motivation] If not, does the introduction state a question that motivates the reader to read on?  
1 Yes  
0 No
- 3 [Motivation] Point (claim) first or point last?  
1 First  
0 Last
- 4 [Core Elements of Argument] Does the paper have a clear claim that responds to the question?  
1 Yes  
0 No
- 5 [Core Elements of Argument] Are the reasons easily identifiable?  
1 Yes  
0 No
- 6 [Core Elements of Argument] Do the reasons support the claim?  
1 Yes  
0 No
- 7 [Core Elements of Argument] Is the evidence sufficient?  
4 Entirely sufficient  
3 Mostly sufficient  
2 Somewhat sufficient  
1 Insufficient
- 8 [Core Elements of Argument] Is the evidence appropriate?  
4 Entirely appropriate  
3 Mostly appropriate  
2 Somewhat appropriate  
1 Inappropriate
- 9 [Core Elements of Argument] Is the evidence reliable?

- 4 Entirely reliable
- 3 Mostly reliable
- 2 Somewhat reliable
- 1 Unreliable

10 [Core Elements of Argument] Is the evidence properly cited?

- 4 All evidence properly cited
- 3 Most evidence properly cited
- 2 Some evidence properly cited
- 1 No evidence properly cited

11 [Interacting with Readers] Does the text acknowledge other points of view in Common Ground?

- 1 Yes
- 0 No

12 [Interacting with Readers] Does the text acknowledge other points of view in Acknowledgement and Response?

- 1 Yes
- 0 No

13 [Interacting with Readers] Does the text approach readers with a dialogic or combative stance?

- 4 Dialogic
- 3 Mostly dialogic
- 2 Mostly combative
- 1 Combative

14 [Style] Does the paper use appropriately academic language?

- 1 Yes
- 0 No

15 [Style] Are the sentences clear and direct? (Characters and Actions)

- 4 Sentences always clear and direct
- 3 Most sentences clear and direct
- 2 Sentences frequently unclear or indirect
- 1 Sentences pervasively unclear or indirect

16 [Style] Does the paper flow? (Old-to-New, Topic Strings)

- 4 Excellent information flow
- 3 Good information flow
- 2 Fair information flow
- 1 Poor information flow

## APPENDIX C

### INTER-RATER RELIABILITY

Inter-rater reliability is a measure of how much consensus there is in the ratings made by different evaluators. The intra-class coefficient (ICC) was used to measure the reliability of the raters in this assessment because the ICC takes into account the differences in ratings for individual segments along with the overall correlation between raters. The ICC ranges from zero to one, with zero indicating little or no agreement and one indicating perfect agreement between raters. Overall, the inter-rater reliability was moderate, at 0.54 for both the first-year papers and the fourth-year papers, indicating that raters did not often mark each individual paper with the *exact* same score for each learning outcome. The reliability fluctuated based on the learning outcomes, with some learning outcomes being rated more consistently than others (for a complete listing of reliabilities, see Appendix B). There was also a difference in rating style among the raters; some raters were more critical than others. Another reason for lower reliability was the high number of raters to papers (approximately 12 raters for 179 first-year papers and 14 raters for 355 fourth-year papers). The advantage to having more raters is that it increases faculty buy-in for, and experience with, assessment.

The raw agreement, however, presents a more easily interpreted view of the raters' agreement. On average, approximately 50 percent of the ratings were exact matches between rater 1 and rater 2, 44 percent of ratings differed by only one point, and only 6 percent of ratings differed by more than one point. In order to correct for the ratings that differed by a point or more, all final scores on each learning outcome are the average score of both raters. Thus rater differences were reduced. Nevertheless, future assessments should include a more extended norming session, and perhaps additional "mini" norming sessions midway through the process.

Table C1. Inter-Rater Reliability

	First-Year	Fourth-Year
Introduction	0.54	0.57
Parts of Argument I – Claims & Sub-claims	0.61	0.56
Parts of Argument II - Evidence	0.49	0.56
Counterarguments	0.49	0.61
Cohesion & Coherence	0.41	0.41
Audience & Tone	0.62	0.57
Conclusions	0.61	0.49

**APPENDIX D****REPRESENTATIVENESS ANALYSIS****Table D1. Gender**

	Sample	Population
Female	71.0%	62.0%
Male	29.0%	38.0%
Total N	355	2657

**Table D2. Ethnicity**

	Sample	Population
White	70.1%	64.7%
African-American	8.2%	11.6%
Asian	9.6%	10.5%
Hispanic	4.5%	5.1%
Other	7.6%	8.3%
Total N	355	2657

**Table D3. School and Discipline within CLAS**

	Sample	Population
College of Arts & Sciences Humanities and Fine Arts	42.8%	35.5%
College of Arts & Sciences Science and Math	25.6%	28.8%
College of Arts & Sciences Social Science	27.0%	25.6%
College of Arts & Sciences Undeclared	0.0%	1.5%
Architecture	2.5%	3.3%
Nursing	2.0%	4.2%
Education	0.0%	1.3%
Total N	355	2657

**Table D4. GPA**

	Sample	Population
Low (0.00-3.02)	22.8%	33.2%
Middle (3.03-3.44)	31.3%	33.3%
High (3.45-4.00)	45.9%	33.5%
<b>Mean</b>	<b>3.33</b>	<b>3.18</b>
Total N	355	2657

**Table D5. Verbal SAT Score**

	Sample	Population
Low (350-610)	20.0%	31.4%
Middle (620-680)	32.8%	32.7%
High (690-800)	47.2%	35.9%
<b>Mean</b>	<b>672.5</b>	<b>650.9</b>
Total N*	335	2225

\*7% of students do not have Verbal SAT Scores

**Table D6. Exempt Status**

	Sample	Population
Exempted	31.8%	29.1%
ENWR	68.2%	70.9%
Total N	355	2657

**APPENDIX E**  
**MEAN WRITING ASSESSMENT SCORES WITH AND WITHOUT EXEMPTED STUDENTS**

**Table E1. Mean Scores on Learning Outcomes –  
First-Years and All Fourth-Years (including Exempted)**

	First- Year	Fourth- Year	Effect- Size
Introduction*	2.03	2.52	0.11
Parts of Argument I – Claims & Sub-claims*	1.98	2.68	0.21
Parts of Argument II - Evidence*	2.04	2.82	0.26
Counterarguments*	1.73	1.97	0.02
Cohesion & Coherence*	2.25	2.93	0.26
Audience & Tone*	2.04	2.85	0.26
Conclusions*	2.05	2.75	0.20
<b>Final Score*</b>	<b>2.02</b>	<b>2.64</b>	<b>0.25</b>
N	179	355	

\* Mean differences in performance are significant at  $p < .001$ .

**Table E2. Mean Scores on Learning Outcomes –  
First-Years and Fourth-Years (excluding Exempted)**

	First- Year	Fourth- Year	Effect- Size
Introduction**	2.03	2.42	0.09
Parts of Argument I – Claims & Sub-claims**	1.98	2.57	0.18
Parts of Argument II - Evidence**	2.04	2.7	0.23
Counterarguments*	1.73	1.92	0.02
Cohesion & Coherence**	2.25	2.83	0.22
Audience & Tone**	2.04	2.72	0.22
Conclusions**	2.05	2.64	0.17
<b>Final Score**</b>	<b>2.02</b>	<b>2.54</b>	<b>0.22</b>
N	179	242	

\*Mean differences in performance are significant at  $p < .01$ .

\*\* Mean differences in performance are significant at  $p < .001$ .

**APPENDIX F**  
**WRITING ASSESSMENT PERFORMANCE BY DISCIPLINE**

**Table F1. Fourth-Year Writing Competency by Discipline within the College of Arts and Sciences**

	Minimally Competent or above	Competent or above	Highly Competent	N
Humanities and Fine Arts	98%	64%	8%	152
Science and Math	100%	51%	9%	91
Social Science	99%	67%	7%	96

**Table F2. Mean Scores on Learning Outcomes by Discipline within the College**

	Humanities/Fine Arts	Science and Math	Social Sciences
Introduction	2.57	2.43	2.53
Parts of Argument I – Claims & Sub-claims	2.72	2.56	2.7
Parts of Argument II - Evidence	2.86	2.74	2.79
Counterarguments*	1.88	1.93	2.14
Cohesion & Coherence**	3.04	2.8	2.91
Audience & Tone <sup>†</sup>	2.95	2.71	2.83
Conclusions	2.79	2.64	2.79
<b>Overall Score</b>	<b>2.69</b>	<b>2.55</b>	<b>2.67</b>

\* Mean differences in performance between Social sciences and Humanities/Fine Arts significant at  $p < .05$ .

\*\* Mean differences in performance between Humanities/Fine Arts and Science/Math significant at  $p < .01$ .

† Mean differences in performance between Humanities/Fine Arts and Science/Math significant at  $p < .05$ .

**Table F3. Mean Scores on Learning Outcomes by Discipline within the College after Controlling for Verbal SAT Score and GPA**

	Humanities/ Fine Arts	Science and Math	Social Sciences
Introduction	2.51	2.48	2.59
Parts of Argument I – Claims & Sub-claims	2.63	2.61	2.74
Parts of Argument II - Evidence	2.79	2.78	2.83
Counterarguments*	1.81	1.97	2.19
Cohesion & Coherence	2.98	2.84	2.94
Audience & Tone	2.88	2.75	2.86
Conclusions	2.72	2.69	2.83
<b>Final Score</b>	<b>2.62</b>	<b>2.59</b>	<b>2.71</b>
N	152	91	96

\* Mean difference in performance between Social Sciences and Humanities/Fine Arts significant at  $p < .05$ .

**APPENDIX G**  
**MEAN WRITING ASSESSMENT SCORES PRE- AND POST-FWR**

**Table G1. Beginning-of-Course (Pre) vs. End-of-Course (Post)**

		<b>Pre</b>	<b>Post</b>	
[Motivation] Is there a canonical problem statement? (Status Quo, Destabilizing Moment, Costs & Consequences)	YES	8%	38%	
	YES/NO	38%	30%	
	NO	54%	32%	
	Mean	0.27	0.53	**
[Motivation] If not, does the introduction state a question that motivates readers to read on?	YES	54%	78%	
	YES/NO	34%	22%	
	NO	12%	0%	
	Mean	0.71	0.89	**
[Motivation] Point (claim) first or point last?	FIRST	82%	82%	
	FIRST/LAST	16%	16%	
	LAST	2%	2%	
[Core Elements of Argument] Does the paper have a clear claim that responds to the question?	YES	40%	60%	
	YES/NO	42%	32%	
	NO	18%	8%	
	Mean	0.61	0.76	*
[Core Elements of Argument] Are the reasons easily identifiable?	YES	40%	46%	
	YES/NO	34%	40%	
	NO	26%	14%	
	Mean	0.57	0.66	
[Core Elements of Argument] Do the reasons support the claim?	YES	50%	46%	
	YES/NO	36%	38%	
	NO	14%	16%	
	Mean	0.68	0.65	
[Core Elements of Argument] Is the evidence sufficient?		2.21	2.68	**
[Core Elements of Argument] Is the evidence appropriate?		2.42	2.69	*
[Core Elements of Argument] Is the evidence reliable?		2.38	2.79	**
[Core Elements of Argument] Is the evidence properly cited?		2.30	2.76	**

[Interacting with Readers] Does the text acknowledge other points of view in Common Ground?	YES	34%	52%	
	YES/NO	36%	34%	
	NO	30%	14%	
	Mean	0.52	0.69	*
[Interacting with Readers] Does the text acknowledge other points of view in Acknowledgement and Response?	YES	34%	52%	
	YES/NO	36%	34%	
	NO	30%	14%	
	Mean	0.52	0.69	*
[Interacting with Readers] Does the text approach readers with a dialogic or combative stance?		2.45	2.74	*
[Style] Does the paper use appropriately academic language?	YES	70%	92%	
	YES/NO	30%	8%	
	NO	0%	0%	
	Mean	0.85	0.96	**
[Style] Are the sentences clear and direct? (Characters and Actions)		2.96	3.02	
[Style] Does the paper flow? (Old-to-New, Topic Strings)		2.58	2.77	

\* $p < .05$ , \*\* $p < .01$

## APPENDIX H

## WRITING PERFORMANCE BY FWR PARTICIPANTS VS FWR EXEMPTED STUDENTS

**Table H1. Mean Performance of FWR Participants vs. FWR-Exempted Students**

	Participants	Exempted Students
SAT Verbal	645	734
GPA	3.22	3.55
Writing Assessment Final Score	2.54	2.86
N	242	113

**Table H2. Mean Performance after Controlling for SAT and GPA (Estimates)†**

	Participants	Exempted Students
Introduction	2.50	2.57
Parts of Argument I – Claims & Sub-claims	2.63	2.74
Parts of Argument II - Evidence*	2.75	2.96
Counterarguments	1.96	1.98
Cohesion & Coherence	2.89	2.98
Audience & Tone*	2.77	2.99
Conclusions	2.69	2.84
Final Score	2.60	2.72
N	242	113

\*Mean differences in performance are significant at  $p < .05$ .

†Verbal SAT was held constant at 672 and GPA at 3.33