



## First-Year Writing Intensive Assessment: 2008-2009

December, 2009

Jacobson Center for Writing, Teaching and Learning  
Office of Institutional Research and Educational Assessment



## Contents

Background.....	3
The data .....	9
Overall Scores:.....	9
Scores by Student Demographics:.....	12
Reader Comments:.....	18
Reader Feedback .....	24
Conclusion .....	26
Appendix A: Technical Findings .....	A1
Appendix B: Recruitment Materials.....	B1
Appendix C: Assessment Rubric, Scoring, & Reading Procedure.....	C1

## Background

Smith College, a highly-selective liberal arts college for women with an open curriculum, does not require students to fulfill distributional requirements outside the major. Smith only requires that students take at least one course designated as Writing Intensive (WI) in their first year. To obtain a writing intensive designation, courses must meet the following criteria:

*Writing intensive courses will devote a significant amount of class time to teaching students to write with precision, clarity, economy, and some degree of elegance.*

*That is to say,*

- 1) to articulate a thesis or central argument, or to create a description or report, with an orderly sequence of ideas, apt transitions, and a purpose clear to the intended audience;*
- 2) to support an argument and to enrich an explanation with evidence;*
- 3) when appropriate, to identify and to evaluate suitable primary and secondary sources for scholarly work, demonstrating awareness of library catalogues and databases and of the values and limitations of internet resources;*
- 4) to incorporate the work of others (by quotation, summary, or paraphrase) concisely, effectively, and with attention to the models of citation of the various disciplines and with respect for academic integrity;*
- 5) to compose paragraphs that are unified and coherent;*
- 6) to edit work until it is orderly, clear, and free of violations of the conventions of standard written English (grammar, usage, punctuation, diction, syntax).*

Writing intensive courses include English 118, an introductory writing composition course; interdisciplinary first year seminars; and courses in the general curriculum that meet the writing intensive guidelines. Students with weak writing skills (identified by the dean of enrollment) are encouraged to enroll in English 118; all students, however, including those judged to have weak writing skills, are free to select the WI course of their choice.

In 2006, Provost Susan Bourque convened a group of faculty and administrators to discuss systematic writing assessment at Smith. In previous years, the Smith Committee on College Writing had broken ground in the area of writing assessment with a series of exploratory studies, which, though limited, raised critical questions:

1. How can we assess more effectively how well Smith students write by the end of their writing-intensive (WI) course? (Grades are one indicator, but they often encompass much more than the students' writing ability.)
2. How can we determine over the years whether the quality of first-year student writing is declining, improving, or holding steady?
3. How can we better advise students about their writing needs after the first semester and in the sophomore year?
4. What data do we need to help us plan course offerings in writing in the spring semester and beyond?
5. In what aspects of writing do our students excel and struggle, and what implications do such strengths and weaknesses hold for our pedagogy?

In the fall of 2006, the writing of first-year students upon completion of their WI requirement was assessed. Over the course of the 2006-2007 academic year, the quality of the writing of sophomores produced in the course of their first two years at Smith was assessed. Reports detailing the result of these two assessments are available from the Office of Institutional Research & Educational Assessment.

For the 08-09 WI assessment, members of the initial writing assessment committee worked with the current Smith College writing committee and the provost to develop a new assessment project that leveraged the strengths of the prior two projects, and addressed questions the prior assessments raised. Specifically, the 08-09 writing assessment project aims to provide an objective evaluation of writing from the perspective of faculty across the college. The current project was designed to have adequate statistical power to draw conclusions and to have the student's writing independently evaluated by a faculty member other than her writing instructor. Both of these concerns were limitations in the prior writing assessments.

In August 2008, the 36 faculty teaching Fall 08 WI courses were sent a letter from Provost Bourque encouraging them to participate in this assessment. Faculty participation was voluntary and, in the end, 30 of the 36 WI instructors submitted student papers for review. Participating faculty collected and handed in portfolios containing each student's writing for the entire semester and indicated which piece was, in their opinion, the student's "best paper." Only the best paper was evaluated for the WI assessment.

Within the WI courses, student participation was also voluntary. First-year students in participating WI courses were sent letters explaining how their papers would be used for the study, indicating that opting out and participating anonymously were options available to them. Finally, all Smith faculty were sent a letter from the Committee on College Writing inviting them to participate as readers. Volunteers were offered an honorarium of \$600 for their time. Although there was initial concern that more faculty than were needed signed up to participate, in the end everyone who wanted to was able to be a reader. Copies of each recruitment letter can be found in Appendix A.

A total of 305 students in 32 sections of 22 courses participated (some participating faculty taught more than one course):

- ARH 101, Approaches to Visual Representation
- BIO 110, Introductory Colloquia: Life Sciences for the 21st Century
- CLT 202, Western Classics in Translation, from Homer to Dante.
- ENG 118, Colloquia in Writing, Section 01: Consumer Culture; Sections 02 and 03: Mixing Memory and Desire: Language and the Construction of Experience; Section 04: The Politics of Language; Section 07: To Hell and Back.
- ENG 120, Colloquia in Literature, Section 03: Modern Irish Writing, Section 04: Writing American Lives
- ENG 199, Methods of Literary Study
- ENG 202, Western Classics in Translation, from Homer to Dante
- FYS 113, Meanings and Values in the World of Work
- FYS 127, Adaptation
- FYS 128, Ghosts
- FYS 130A, Lions: Science and Science Fiction
- FYS 141, Reading, Writing and Placemaking: Landscape Studies
- FYS 142, Reacting to the Past
- FYS 147, Science and Politics of Food, Water and Energy
- FYS 153, Excavating Women

- FYS 157, Literature and Science: Models of Time and Space
- FYS 160, The End of the World as We Know It: The Post-Apocalyptic Novel
- FYS 162, Ambition and Adultery: Individualism in the 19<sup>th</sup>-Century Novel
- FYS 164, Issues in Artificial Intelligence
- FYS 165, Childhood in the Literatures of Africa and the African Diaspora
- GOV 100, Introduction to Political Thinking I.
- LOG 100, Valid and Invalid reasoning: What Follows from What?

Twenty-four faculty members gathered in January 2009 to read the papers. While some of the faculty reading were also participating WI instructors, many were not. If WI instructors happened to draw papers written for their classes, they were asked to return those papers and take others. The following departments were represented: Afro-American Studies, Anthropology, Art, Astronomy, Biology, Comparative Literature, East Asian Languages and Literatures, East Asian Studies, Education and Child Study, English, Geology, German Studies, Government, Mathematics & Statistics, Music, and Sociology.

Although each course that was represented is designated as writing intensive, the number and length of papers assigned varied widely. In some cases, 2 or 3 lengthier papers were assigned, while in other cases more than 10 shorter papers were handed in. Of all the papers handed in (not just those that were read for the assessment) the average page length is 3.97 pages<sup>1</sup>. When average paper length is examined by professor, we see that the average by course ranges from 1.64 to 6.07 pages.

Student papers marked as “best” were reviewed by faculty members with a rubric developed by the Jacobson Center and based on the goals for WI courses developed by Smith’s Writing Committee, using a reading process based on one used at Carleton College. Student papers were rated on a three-point scale of “needs work,” “pass,” or “exemplary.” We gave readers a scoring sheet for each paper, which asked them to assess the students’ current abilities—as either pass, needs work, or exemplary—in the following areas:

---

<sup>1</sup> It should be noted that three of the courses represented were excluded from this analysis because those professors only handed in each student’s single best paper. One additional course was also excluded from this analysis because all papers assigned were short “response papers” instead of formal papers.

- Statement of thesis
- Development through supporting arguments
- Assembly & presentation of evidence
- Paragraph development
- Complexity of thought
- Command of written English

The Smith rubric described each criterion at each level of ability. Within each level, readers chose among high, medium, and low proficiency. We also asked readers to assess the paper overall as pass, needs work, or exemplary, and to write a final comment for the student. (Thus, the scale for each of the specific areas of writing was the expanded 1-9, with 1 being “low needs work” and 9 being “high exemplary.” The scale for the overall paper score was simpler: Readers were asked to assess the paper as needs work, pass, or exemplary.) The reading began with a norming exercise using three papers from an earlier WI study. All participating faculty read, scored, and discussed the three papers in order to gain some consensus on the scoring process and ensure some degree of inter-rater reliability.

Following the model developed by Carleton College, papers that were given an overall score of pass were initially read only once. Papers falling into either the needs work or exemplary categories were then read a second time. As time permitted, the passing papers were then read a second time, to allow for an analysis of inter-rater reliability. In the end, all but 19 papers were read at least twice. One hundred and thirty two of the papers that were read twice received the same overall score from both readers. One hundred and fifty one of the papers that were read twice resulted in disagreement over the final score and went into negotiation. In these cases, readers met at the end of the reading process to discuss their differences, settle on a rating, and complete a new scoring sheet. Finally, three papers required a third reader to settle irreconcilable differences in opinions between the first two scorers.

A total of 761 scores were recorded. The readings that went into negotiation were recorded but not included in the final analysis—only the negotiated scores went into the final analysis. A weighting variable was then used to account for the fact that some papers were read once while others were read twice or even three times. When a paper was read only once,

that record was weighted as 1.0. For papers that were read twice and resulted in the same overall score, a weight of 0.5 was applied to each of the two records. Papers that were read twice, resulted in a disagreement, and were then negotiated resulted in three records in the database. In these cases, the two initial readings were weighted as 0 and the negotiated record was weighted as 1.0. Finally, when the two readers of a paper could not come to an agreement and a third reader was used, the ratings of the third reader were weighted as 1.0 while the initial readings were weighted as 0.

Any interpretations of this study's findings should consider its limitations. It is widely understood that college writing is a developmental process, and that one semester of writing instruction is only the start of student writing development. The skills and techniques learned and practiced in a student's WI course will be implemented, honed, and refined in other courses throughout her college career; a summative assessment at the conclusion of a first-semester writing course may reasonably be considered premature. Our instructions to faculty participants clearly acknowledged this issue and recommended that they set their standards at a level appropriate for first-semester first-year students.

## The data

### *Overall Scores:*

On average, readers rated first-years' writing as low- to mid-pass, with mean ratings in the sub-skills ranging from 4.52 to 4.68 points out of a possible 9 (with 9 being strongest). The mean for the overall score is 1.82 (out of 3, with 3 being strongest). Within each sub-skill, the majority of scores fall between high-needs work and mid-pass. Low-needs work and high-exemplary are the least commonly awarded scores.

**Table 1: Overall Scores, Detail**

	Needs Work			Pass			Exemplary		
	Low	Mid	High	Low	Mid	High	Low	Mid	High
Statement of Thesis	2.3%	9.2%	15.0%	21.3%	21.3%	15.7%	7.0%	6.5%	1.7%
Supporting Arguments	1.8%	9.8%	16.2%	22.2%	21.8%	13.8%	8.2%	5.0%	1.2%
Assembly of Evidence	2.2%	6.3%	21.0%	18.0%	22.5%	14.7%	8.3%	5.7%	1.3%
Paragraph Development	2.5%	8.8%	20.4%	18.2%	22.5%	13.0%	8.5%	4.8%	1.2%
Complexity of Thought	2.2%	8.1%	19.6%	23.2%	20.8%	11.6%	9.2%	4.2%	1.2%
Command of English	1.5%	10.8%	17.0%	21.3%	22.0%	13.3%	7.0%	5.2%	1.8%
Overall Score		30.7%			56.8%			12.5%	

**Table 2: Overall Scores, Summarized**

	Needs work	Pass	Exemplary
Statement of Thesis	26.5%	58.3%	15.2%
Supporting Arguments	27.8%	57.8%	14.3%
Assembly of Evidence	29.5%	55.2%	15.3%
Paragraph Development	31.7%	53.8%	14.5%
Complexity of Thought	29.9%	55.5%	14.6%
Command of English	29.3%	56.7%	14.0%
Overall Score	30.7%	56.8%	12.5%

**Table 3: Mean Scores**

	<b>Mean</b>
Statement of Thesis	4.68
Supporting Arguments	4.59
Assembly of Evidence	4.66
Paragraph Development	4.54
Complexity of Thought	4.52
Command of English	4.56
Overall Score (1-3)	1.82

When we examined the differences among ratings in the three types of courses included in the study (ENG 118, First Year Seminars, and other Writing Intensive courses) a statistically significant difference ( $p < .05$  chi-square) among course types emerged in just one category: command of written English. Within this sub-skill, papers written for English 118 courses had a significantly higher percentage of needs work scores (50%) than papers written in First Year Seminars (26%) or other WI courses (25%). This is not surprising, given that ESOL students and those with weak writing skills are encouraged to enroll in English 118.

**Table 4: Course Type**

		<b>First-Year Seminar</b>	<b>ENG 118</b>	<b>Other WI</b>	<b>Total</b>
Statement of Thesis	Needs work	25.4%	34.8%	24.6%	26.6%
	Pass	61.9%	52.2%	55.1%	58.5%
	Exemplary	12.7%	13.0%	20.4%	14.9%
Supporting Arguments	Needs work	26.8%	35.2%	26.3%	27.9%
	Pass	60.0%	54.9%	55.7%	58.0%
	Exemplary	13.2%	9.9%	18.0%	14.0%
Assembly of Evidence	Needs work	27.1%	40.7%	28.7%	29.6%
	Pass	60.3%	46.2%	50.3%	55.4%
	Exemplary	12.6%	13.2%	21.0%	15.1%
Paragraph Development	Needs work	27.5%	44.9%	33.7%	31.8%
	Pass	58.5%	44.9%	49.4%	53.9%
	Exemplary	14.0%	10.1%	16.9%	14.2%
Complexity of Thought	Needs work	31.2%	31.9%	26.3%	30.0%
	Pass	55.7%	54.9%	56.3%	55.7%
	Exemplary	13.1%	13.2%	17.5%	14.3%
Command of English*	Needs work	26.0%	50.0%	25.0%	29.4%
	Pass	61.4%	38.0%	56.7%	56.5%
	Exemplary	12.6%	12.0%	18.3%	14.0%
Overall Score	Needs work	29.3%	41.3%	28.1%	30.8%
	Pass	60.9%	47.8%	53.9%	57.0%
	Exemplary	9.8%	10.9%	18.0%	12.2%

\* Difference is statistically significant at the .05 level

### *Scores by Student Demographics:*

First-generation students appear to be slightly, but not statistically significantly, weaker writers, though we must be cautious in our interpretation of small numbers: 46 first-generation and 264 not first-generation. The score differences between first-generation and not first-generation students were not statistically significant; further research will be necessary to confirm these findings statistically.

Overall, 34.1% of papers by first-generation students need work, as compared to 30.1% of papers by students who are not first-generation. Similar disparities between the two groups are seen for statement of thesis, supporting arguments, assembly of evidence, and complexity of thought. Interestingly, the sub-skills of paragraph development and command of written English show *smaller* percentages of first-generation papers needing work. Yet each of these apparent differences is both small and not statistically significant. This contradicts the findings of previous WI studies, in which large and sometimes significant differences between the two groups of students were found.

**Table 5: First Generation Students**

		First-Generation		Non First-Generation	
		Count	%	Count	%
Statement of Thesis	Needs work	14	30.0%	66	25.9%
	Pass	23	51.1%	152	59.6%
	Exemplary	9	18.9%	37	14.5%
Supporting Arguments	Needs work	14	30.3%	70	27.4%
	Pass	26	57.3%	148	57.9%
	Exemplary	6	12.4%	38	14.7%
Assembly of Evidence	Needs work	14	30.3%	75	29.4%
	Pass	27	59.6%	139	54.4%
	Exemplary	5	10.1%	42	16.2%
Paragraph Development	Needs work	12	26.4%	84	32.6%
	Pass	27	62.1%	134	52.3%
	Exemplary	5	11.5%	39	15.0%
Complexity of Thought	Needs work	14	32.2%	75	29.5%
	Pass	26	59.8%	140	54.8%
	Exemplary	4	8.0%	40	15.7%
Command of English	Needs work	13	29.2%	75	29.4%
	Pass	27	60.7%	143	56.0%
	Exemplary	5	10.1%	38	14.7%
Overall Score	Needs work	16	34.1%	78	30.1%
	Pass	26	57.1%	147	56.8%
	Exemplary	4	8.8%	34	13.1%

The paper ratings correlate with SAT Writing scores and admissions reader ratings. Dividing SAT Writing scores into quartiles (<590, 590-650, 660-690, 700-800) reveals that the percentage of needs work portfolios is highest among low-scoring students and lowest among high-scoring students. Nearly half of the papers by students in the lowest quartile need

work, as do more than one-third of papers in the second quartile, as compared to only 29% and 18% in the top two quartiles, respectively. The differences between quartiles are statistically significant at the .05 level (chi-square) for all sub-skills. Generally, the percentage of papers needing work drops most sharply between the third and fourth quartile: for supporting arguments, paragraph development, complexity of thought, and command of written English. Taking supporting arguments as an example, we see that between 31.7% and 37.5% of papers written by students in the first three quartiles need work, but this drops dramatically to just 13.2% in the highest quartile.

Table 6: SAT Writing Quartile

		SAT WR <590		SAT WR 590-650		SAT WR 660-690		SAT WR 700-800	
		Count	%	Count	%	Count	%	Count	%
Statement of Thesis*	Needs work	22	41.3%	21	29.4%	10	19.6%	14	18.1%
	Pass	28	52.9%	42	58.0%	32	66.0%	44	56.8%
	Exemplary	3	5.8%	9	12.6%	7	14.4%	20	25.2%
Supporting Arguments*	Needs work	20	37.5%	25	34.7%	16	31.6%	10	13.2%
	Pass	31	58.7%	40	55.6%	28	56.1%	49	64.5%
	Exemplary	2	3.8%	7	9.7%	6	12.2%	17	22.4%
Assembly of Evidence*	Needs work	25	47.1%	24	33.1%	13	26.5%	14	18.2%
	Pass	25	48.1%	39	54.9%	28	56.1%	46	59.7%
	Exemplary	3	4.8%	9	12.0%	9	17.3%	17	22.1%
Paragraph Development*	Needs work	22	42.2%	28	38.5%	17	34.7%	15	19.4%
	Pass	26	51.0%	40	55.2%	25	50.0%	44	56.8%
	Exemplary	4	6.9%	5	6.3%	8	15.3%	19	23.9%
Complexity of Thought*	Needs work	20	37.5%	28	39.2%	16	32.3%	11	13.7%
	Pass	32	60.6%	35	48.3%	25	51.0%	51	66.0%
	Exemplary	1	1.9%	9	12.6%	8	16.7%	16	20.3%
Command of English*	Needs work	27	51.0%	28	38.2%	14	28.1%	10	12.9%
	Pass	25	47.1%	38	52.8%	28	58.3%	48	61.9%
	Exemplary	1	1.9%	7	9.0%	7	13.5%	20	25.2%
Overall Score*	Needs work	27	48.6%	25	34.2%	14	28.6%	14	18.2%
	Pass	26	47.7%	44	60.3%	28	57.1%	47	61.0%
	Exemplary	2	3.7%	4	5.5%	7	14.3%	16	20.8%

\* Difference is statistically significant at the .05 level

If we divide the study participants into a high and low group based on their SAT Writing scores, a 650 score is the dividing mark (i.e., half the participants scored above 650 and half scored at or below 650). Differences between these two groups are statistically significant for the overall rating and all sub-skills.

Paper ratings also correlate with high and low admissions reader ratings (high being 1-4 and low being 5-8) with 24.6% of papers by highly rated students and 35.7% of papers by lower rated students needing work. However, this apparent difference in the overall score of the paper is not statistically significant. Significant differences by reader rating were detected for the following sub-skills: supporting arguments, paragraph development, complexity of thought, and command of written English. For example, more than 20% of the papers written by highly rated students were scored as exemplary in the areas of paragraph development, complexity of thought, and command of written English, as compared to fewer than 10% of papers written by lower rated students.

**Table 7: High or Low Reader Rating**

		High Reader Rating		Low Reader Rating	
		Count	%	Count	%
Statement of Thesis	Needs work	27	22.6%	52	30.8%
	Pass	70	58.2%	96	57.5%
	Exemplary	23	19.2%	20	11.7%
Supporting Arguments*	Needs work	26	21.6%	55	32.7%
	Pass	71	60.2%	96	56.8%
	Exemplary	22	18.2%	18	10.4%
Assembly of Evidence	Needs work	31	25.8%	55	32.7%
	Pass	64	53.8%	94	56.0%
	Exemplary	24	20.3%	19	11.3%
Paragraph Development*	Needs work	36	30.4%	58	34.4%
	Pass	59	49.4%	94	56.0%
	Exemplary	24	20.3%	16	9.6%
Complexity of Thought*	Needs work	31	26.1%	55	32.9%
	Pass	64	53.8%	96	58.0%
	Exemplary	24	20.2%	15	9.1%
Command of English*	Needs work	34	27.9%	54	32.1%
	Pass	62	51.3%	98	58.6%
	Exemplary	25	20.8%	16	9.3%
Overall Score	Needs work	30	24.6%	61	35.7%
	Pass	73	59.8%	93	54.9%
	Exemplary	19	15.6%	16	9.4%

\* Difference is statistically significant at the .05 level

### ***Reader Comments:***

In addition to rating papers based on a rubric, faculty readers were provided a space in which to write comments on the scoring sheets; the qualitative data is derived from the reader comments. Readers rated 56.8% of the papers as pass.

Reader comments reflect the rubric definition of a passing paper (see Appendix C).

*“This is an adequate paper. The writer has command of the structure and mechanics of writing but it does not in argument reach beyond adequate.”*

*“A good essay that accomplishes what it sets out to achieve - carefully designed, textually supported and ultimately persuasive. Central claim is good -- complexity of thought pleases rather than impresses.”*

More than 30% of papers were rated as needs work. In addition to echoing the needs work rubric definition (See Appendix B), some comments elaborated on what made the paper insufficient. In some cases, authors struggled in all areas defined by the rubric.

*“The argument is not broken down into paragraphs. Claims need to be better articulated, more organized, and better supported. Tone of language needs to be more concise. References should be examined more carefully.”*

*“The student seems unprepared for college writing. Her central purpose, trajectory, and arguments are unclear and not well organized. I do not understand the general purpose of the book after reading her analysis. She seems to note a vague awareness of the author’s purpose (that life is composed of instances), but a lack of overall organization prevents this point from being integrated well. At times, quotes from the memoir are strung together without any contextualizing or analysis. Though she seems to integrate outside sources, as the assignment dictates, these sources are not well integrated and are not used as evidence to support any sort of claim.”*

Other needs work papers were generally difficult to understand.

*“Although the author’s ideas about the topic may well be clear, her communication is not. Many sentences are difficult to understand. Poor punctuation and choppy phrasing make comprehension a formidable task. This is compounded by occasional repetition and superfluous interjection. Furthermore, the assigned question (about the portrayal of women) is only answered tangentially and incompletely. Development of argument is not clear.”*

Some needs work papers fell into that category not because the author did poorly in all areas defined in the rubric, but because of specific shortcomings.

*“The position of the writer is clear, and she gathers some good evidence to support her position. Yet, problems with the writing obfuscate her meaning in many passages. Words are misused and there are many typos. As for*

*the argument, the paper begins in strong support of the supremacy of [one position], but ends with a defense of [an incompatible position].”*

*“Great start, clearest thesis statement I have read in two days, then a very poorly constructed, not well thought out, argument. All three reasons seem very similar to me and half of the paragraph rambles without supporting the argument coherently.”*

Some subscales were rated in many respects as both exemplary and weak. Statement of thesis was most often rated as an extreme: either needs work or exemplary. Readers rated 26.5% of the papers as needing work in this area and 15.2% as exemplary. Polarities are more apparent among other skills, with readers rating much more writing as needs work than exemplary. For example, readers rated 31.7% of portfolios as needing work in paragraph development and only 14.5% as exemplary.

### ***Exemplary Skills:***

Readers scored 12.5% of the papers as exemplary (see Appendix B for rubric definition of overall exemplary paper). Few papers received exemplary ratings in every category of the rubric. When this did occur, readers more often than not elaborated on what made these papers stand out.

*“Not only did the student fulfill the expectations of the assignment, she also crafted a profoundly heart-felt meditation on a deeply personal memento. The integration of her personal experiences with the analysis of the required texts was seamless. Extraordinary.”*

*“This is an excellent paper for a first year student. She carefully develops her main argument and supports it using scholarly texts and clear examples. She does a good job teasing out the contradictions and complexities of this [topic]. The student writes with a confidence that I would expect of a sophomore or junior level student.”*

*“This is the strongest paper I’ve read so far, both in substance and expression. The student is clear without sacrificing subtlety, and she collects and explains examples intelligently and persuasively. There are times when the style could profit from a bit more discipline, but in general, she writes fluently and with complexity and reflects in her writing excellence of her observations and of her thinking in general.”*

Not all exemplary papers received exemplary marks in every category of the rubric. In these cases, skillful work in one or more areas made up for shortcomings in others.

*“Multiple theses present in first paragraph, which is the main one? Need attention to detail e.g. ‘he’ vs. ‘his’. Body paragraphs would be strengthened by [the] integration of evidence with thesis. Overall, after a shaky start I became captivated.”*

*“This is a lovely paper, close to the text of the book, sensitive to the ways in which father and son see each other and to the way we see them. The only thing I’d change would be the final paragraph, which tends to summarize the argument you’ve made -- not necessary in a short paper. Use the final paragraph to add one final thought or perspective -- or investigate one last passage -- in relation to your completed argument.”*

*“Some (few) grammatical errors. Body paragraph lacks a terminal integration with the thesis. These are minor grievances with an otherwise exemplary essay.”*

An analysis of the subscales shows that students’ strongest skills are assembly of evidence and statement of thesis, with more than 15% of papers being rated as exemplary in these areas. Comments about exemplary uses of evidence echo the rubric definition, which focuses on the thoughtful use of plentiful and convincing evidence to support an argument.

*“This well-articulated argument is supported by carefully selected and persuasive evidence.”*

Reader comments on statement of thesis endorse the rubric definition, which calls for a substantive, articulately stated, and thought-provoking central claim.

*“Your thesis is intriguing and well-supported.”*

*“Every paragraph is relevant to her topic, and every sentence contributes in some clear way to her thesis.”*

Just under 15% of papers received exemplary ratings in the areas of complexity of thought and paragraph development. Reader comments indicate that complex thinking can make for an outstanding paper and that some writers had excellent paragraph development.

*“Complexity of thought takes this paper into exemplary category.”*

*“She writes fluently and with complexity, and reflects in her writing excellence of her observations and of her thinking in general.”*

*“Paragraphs well-developed.”*

*“Strong sense of the structures of academic writing: paragraphing, development, use of evidence, linguistic flow.”*

*“Every paragraph is relevant to her topic, and every sentence contributes in some clear way to her thesis.”*

Comments related to students’ use of supporting arguments (14.3% exemplary) reiterate the rubric definition of the subscale.

*“She carefully develops her main argument and supports it using scholarly texts and clear examples.”*

*“Paper leads the reader through a convincing argument.”*

On the command of written English subscale, 14.0% of the papers received an exemplary rating. Some comments about exemplary work on this subscale stretched beyond the rubric definition.

*“Written with admirable command of English -- witty as well as correct, with a sure sense of style.”*

*“A fine paper, written with complete command of the subject and of English.”*

*“Demonstrates good command of English.”*

### ***Weakest Skills:***

A group of skills emerged as weakest among WI students, with slightly less than one-third of papers being rated as “needs work” in each area. These areas are: paragraph development, complexity of thought, assembly of evidence, and command of written English. Additionally, more than one-quarter of papers received a needs work rating in the areas of supporting arguments and statement of thesis. This is similar to the results seen in the prior two writing assessments, in which paragraph development and supporting arguments were both found to be areas of weakness.

Students’ weakest skill is paragraph development (31.7% needs work). In their narrative feedback to students, readers often elaborated on the rubric definition of a paper that needs work in this category, with comments ranging from lack of a focus and/or central arguments to lack of cohesion.

*“Paragraphs often include unrelated ideas, and the student switches her focus without transition.”*

*“Paragraphs also lack central arguments. . . Many of the sentences make no sense in the context of the paragraphs in which they are embedded.”*

*“Paragraphs are not well defined nor developed.”*

*“Paragraphs ramble. I am not clear what connection each paragraph has to the essay.”*

*“Your paragraphs wander from topic to topic with little direction.”*

*“Your paragraphs are not unified or coherent, meaning that the links between them are unclear, and they do not always have main points.”*

Complexity of thought (29.9% needs work) follows paragraph development as one of the students' weakest skills. Papers needing work in this area received reader comments that expanded on the rubric definition.

*"The thesis itself lacks the kind of development that would show more complex thought. . . The language, too, lacks complexity and differentiation."*

*"Points are not concise, and the arguments are unsophisticated."*

*"Both the vocabulary and the thinking remain on a very basic level."*

*"Demonstrates no analytical thought process. . ."*

Assembly of evidence follows paragraph development as another of the students' weakest skills (29.5% needs work).

Beyond the rubric definition, readers explained that evidence lacked integration, contradicted the main argument or thesis, were poorly organized, or were used inappropriately/ineptly.

*"The paper is hampered by too many long block quotations. . . too often, these long quotes replace rather than support the student's own readings or insights."*

*"While the paper presents evidence to support the main point, it does not do so effectively -- indeed the relevance of the evidence is not clear."*

*"The evidence is carefully presented to show [an argument], which contradicts both the introductory and concluding claim."*

*"The textual evidence that you present does not always illustrate the point you are trying to make."*

*"Evidence is mixed up. Do you have any objective evidence rather than just author's opinion?"*

*"The evidence from the text that you present does not always relate to the point you're trying to make."*

*"Paragraphs also end in evidence (quotes) and do not provide a connection between the material in the paragraph with the thesis."*

On the command of written English subscale, 29.3% of the papers received a needs work rating. Faculty readers commented mostly on troubles with grammar and writing that was awkward.

*"Awkward use of English makes reading this paper difficult."*

*"The student seems to struggle with some basic rules of English grammar. . . Phrases are awkward at times and take several reads to understand. . . her ideas are trapped in muddled writing and misused words."*

*"Grammatical errors and awkward use of English detract from the paper."*

Of the total number of papers, 27.8% received a needs work rating on the supporting arguments subscale. Comments that expanded on the rubric definition were related to inconsistencies and/or contradictions in the argument, lack of direction, and repetition.

*“The speech could benefit from a more clearly articulated argument and examples.”*

*“One has the feeling that the writer has made up her mind (not a particularly original or debatable conclusion), but that she really has no idea of how to write an argument in support.”*

*“There is no sense of direction in the argument.”*

*“Your argument is hampered by a simplistic structure--after a brief introduction, you compress the plots of each story into 1-2 long paragraphs, followed by one paragraph for the analyses of each work--cutting short the development of your argument.”*

*“The argument that supports [the thesis] is a repetition of the same point with a slight variation in wording.”*

More than one-quarter (26.5%) of papers received a needs work rating on statement of thesis. Reader comments reveal that there is sometimes a complete lack of a thesis statement, making the paper difficult to follow.

*“This paper had an unwieldy, ill-defined topic that resulted in generalization, disorder, and difficulties in integrating research into the analysis.”*

*“The thesis statement is so weak, it is never clear to the reader where this is going or what argument is being made.”*

*“This paper has no central focus--it feels a bit like a list of generally related topics, but with no central thesis it doesn't add up to a convincing argument for anything.”*

*“The paper needs a thesis or a main point for the reader. Because it is not present all of the thoughtful marshalling of the details in the paper are rather disjointed. Where is the writer taking us? The conclusion seems to contain a thesis but it is muddled in a paragraph that is still explicating the poem and not leading us to a bigger picture.”*

*“Because there is no thesis, paragraphs aren't developed, which further leads to... lack of analysis.”*

*“There's no thesis here: this paper essentially says “I'll discuss this topic,” rather than taking a position on the subject. As a result there's a rambling, unfocused quality to the paper.”*

## Reader Feedback

After rating the papers, readers discussed their impressions of the portfolios and the assessment process. The conversation focused mostly on improvements that could be made to Writing Intensive courses and the assessment process, particularly if Smith were to institutionalize the assessment process as a requirement for all students.

### Requirements of a WI Course:

- Readers understand the tradition of giving faculty the freedom to teach their own courses, yet still feel that a writing intensive course should be assigning papers of more than 1-2 pages in length.
- Readers thought that WI class assignments ought to be crafted to simultaneously allow students to demonstrate both the formulation of an argument and knowledge of the mechanics of writing.

### Requirements of Assessment Project:

- Readers would like to see more consistency across the papers being assessed, both in length and genre.
- Some concern was expressed over assessing a student on the basis of just one paper, especially if this were to become a Smith-wide requirement.
- There is a desire for the assessment rubric to address issues of proofreading and proper use of citations.
- Readers felt it was critical to have a copy of the assignment in hand when assessing papers.

Readers commented that they learned much from the assessment process. For example, in reading a great variety of assignments, they developed ideas about how to improve their own. They commented on how strong assignments produce strong writing and bad assignments produce bad writing. In reading students' writing in a variety of courses, they also readjusted their standards, which they may have set too high or too low. Finally, faculty enjoyed and benefited from talking to their colleagues in different disciplines about writing and grading. In other words, they enjoyed and benefited

from being part of an interdisciplinary culture of writing, which is writing across the curriculum at its best. One person said, “I wish I could do all my grading this way.” Grading is often a tedious, solitary activity. Done in an interactive group with colleagues, it can become an enjoyable learning experience.

## Conclusion

First year writing assessment has many benefits as a means of assessing writing, mostly pedagogical. Students benefit from the assessment by receiving feedback about their writing skills as they enter their second year. Faculty benefit from exposure to a wide range of assignments and student writing, which leads to individual and collaborative reflection about their own expectations and practices, and they benefit from engaging with colleagues across the disciplines in a pedagogical endeavor, colleagues whom they may not see much and with whom they may not engage much. Ultimately, the reading experience will lead to greater collegiality and better, more consistent teaching of writing.

The results of the Fall 08 writing assessment should prove useful in beginning to answer the questions set forth by the Committee on College Writing (see page 3). The assessment of each paper by one or more faculty members provides an additional measure of the quality of first year student writing. Together, the course grade and the scores a particular student received on her paper provide a more complete picture of that student's writing ability. And since each paper was scored on six sub-scales, the results of the assessment would make a useful starting point for individual conversations with students about their writing.

Based on the 305 papers read for this study, it appears that Smith first-years are generally strongest in the area of statement of thesis, with an overall mean of 4.67 and 15.2% of papers receiving an exemplary score. The three weakest areas are complexity of thought, command of English, and paragraph development. This general sense of first year writing skills may help in planning future writing intensive course options and allow for a particular focus on the weaker areas.

To examine trends in the quality of first year student writing over time, an assessment such as this would need to be repeated several times.

The enthusiasm with which Smith faculty volunteered to submit writing to the assessment project and volunteered to read papers for two and half days speaks highly of their commitment to the teaching of writing. Moreover, this commitment cuts deeply across disciplines and divisions.

The data confirm anecdotal impressions: the majority of our first-year students write well enough (57%); some of our students write exceptionally well (13%); too many of our students do not write well enough, even after considerable effort (31%). We now know that our WI courses do a very good job teaching students to articulate a thesis but not as good a job (perhaps even a poor job) teaching them to write strong paragraphs and grammatically correct, well punctuated sentences, both of which we can and should teach if students are to succeed in their first year, transition smoothly to higher level courses, and be successful in their lives after Smith. We may not be able to do much about the third weak skill—complexity of thought—but have faith that it will emerge over time with practice.

The bulk of our “needs work” students are in the high end of needs work, just as a considerable number of sister students are in the low “pass” category. These students need courses that devote a substantial amount of class time to teaching writing and instructors willing to spend a substantial amount of time outside of class individually meeting with students and commenting on drafts.

The data show that we are successfully placing many of the weakest writers in ENG 118, where much class time is devoted to the teaching of writing and instructors spend much individual time with students, but they also show that a single course is not enough, because these students appear to lag behind their peers after one semester of instruction. No one expects students to learn to write in one semester, or perhaps even two, but it appears very appropriate to develop a two-semester sequence for the weakest writers.

We have reached a point, however, where ENG 118 cannot accommodate all students who need intensive work in writing. We need to decide whether to expand the offerings of courses like ENG 118 (and the pilot ENG 119, of which we offer only one section), rethink the mandates for all WI courses to require instructors of these courses to spend more time teaching writing, or both.

Now that we have good data, we need to make the hard choices.

# Appendix A: Technical Findings

### Analysis of Reliability:

In order to assess inter-rater reliability, two distinct analyses were conducted. First, a comparison of mean scores by reader allows us to gauge overall variation among the raters. Second, correlations of the multiple scorings of individual papers allow us to examine the consistency of the rating process.

We looked at each rater's average rating (overall score). This revealed that the lowest mean rating was 1.40 and the highest was 2.25 (on a scale of 1-3). Further, we learn that the difference between the highest mean (2.25) and the two lowest (1.40 and 1.54) is statistically significant, but that apparent differences among the mean overall scores of all other raters are non-significant. This indicates that, in general, the raters were similar in their assessment of the student papers. Next, we can do the same thing for each of the sub-scores on the rubric. Table A1, below, provides some basic information on the variation by reader in mean scores (on a scale of 1-9) for each of the six areas.

**Table A1: Mean Scores by Sub-Category**

	<b>Lowest Rater Mean</b>	<b>Highest Rater Mean</b>	<b>Overall Mean</b>	<b>Std. Deviation</b>
Statement of Thesis*	3.00	6.50	4.67	1.92
Supporting Arguments*	2.59	6.54	4.63	1.93
Assembly of Evidence*	2.83	6.50	4.65	1.93
Paragraph Development*	2.15	6.23	4.59	1.93
Complexity of Thought*	3.05	5.83	4.55	1.90
Command of English*	2.89	5.95	4.57	1.85

\*Difference between highest and lowest mean rating is statistically significant and the .05 level.

Here we see that the statement of thesis mean score, by reader, ranged from 3.00 to 6.50, with an overall mean of 4.67. For each sub-category, the difference between the lowest and highest rater means is statistically significant, but this is largely due to the presence of one rater with consistently low mean scores and two with consistently high mean scores. In fact, if we remove those three outliers from the analysis, we get much more consistent mean ratings and far fewer statistically significant differences (see Table A2).

**Table A2: Mean Scores by Sub-Category, Excluding Outliers**

	<b>Lowest Rater Mean</b>	<b>Highest Rater Mean</b>	<b>Overall Mean</b>	<b>Std. Deviation</b>
Statement of Thesis	3.95	5.77	4.70	1.95
Supporting Arguments*	3.22	5.41	4.63	1.96
Assembly of Evidence	3.71	5.35	4.62	1.99
Paragraph Development	3.47	5.45	4.62	1.93
Complexity of Thought	3.50	5.55	4.56	1.93
Command of English*	3.53	5.95	4.60	1.89

\*Difference between highest and lowest mean rating is statistically significant and the .05 level.

Taking each paper that was read twice, we can examine the correlation of the first scoring and the second scoring on each of the sub-scales and the overall score. Table A3 shows that each of the correlations is statistically significant, although not particularly strong.

**Table A3: Score/Re-Score Correlations**

	<b>Correlation Coefficient</b>
Statement of Thesis	0.340*
Supporting Arguments	0.290*
Assembly of Evidence	0.268*
Paragraph Development	0.272*
Complexity of Thought	0.228*
Command of English	0.381*
Overall Score	0.281*

\* Correlation is significant at the 0.01 level (2-tailed).

Stronger correlations would have been expected if the readers had received explicit training on using the rubric in a reliable and consistent manner. Due to time and resource limitations, the 24 readers were not provided extensive training, but participated in a single norming exercise prior to reading student papers. Keeping that in mind, and the fact that the

24 readers came from disciplines across the college and were often reading papers from disciplines other than their own, the above coefficients are not surprising. Still, if the goal is to come to a unitary conclusion about each student's writing in relation to a College-wide standard, there is room for improvement. In any future writing assessment, additional training could be incorporated to increase inter-rater reliability.

It will be important, as we move forward with further writing assessment projects, to consider to what extent we value consistency and uniformity in our ratings of student writing. While it is of course necessary to have highly reliable assessment measures for any process that serves as a criterion for the awarding of a credential or that serves a gatekeeping function, preventing students from progressing through college, to date, Smith has not used its writing assessment processes in these ways. It is notable that Carleton College, the institution on which Smith has based many of its procedures, does not norm its assessment process to a high degree of reliability; instead it relies on the negotiation process to correct for outliers. It is worthwhile to consider whether Smith should adopt a more formal norming procedure, or whether Carleton's process, which yields a level of reliability described above, is more appropriate for Smith.

#### **Score Correlations:**

Further analysis of the extent to which student attributes and writing performance co-vary for all students shows strong internal correlations among the different measures. SAT writing scores and reader ratings correlate with each other and the overall quality of a paper, with significance at the 0.01 level. Correlations among the different rubric criteria also correlate highly at the 0.01 significance level. In other words, if a student's writing is exemplary, is adequate, or needs work in one domain, it is very likely that it will do so in all other domains (see Table A4).

**Table A4: Correlations**

		Statement of Thesis	Supporting Arguments	Assembly of Evidence	Paragraph Development	Complexity of Thought	Command of English	Overall
Statement of Thesis	Pearson Correlation	1.000	.804**	.737**	.769**	.706**	.653**	.767**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000
	N	300	298	298	297	296	298	295
Supporting Arguments	Pearson Correlation	.804**	1.000	.862**	.859**	.791**	.705**	.819**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000
	N	298	300	299	298	297	298	295
Assembly of Evidence	Pearson Correlation	.737**	.862**	1.000	.839**	.794**	.703**	.800**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000
	N	298	299	300	298	297	298	295
Paragraph Development	Pearson Correlation	.769**	.859**	.839**	1.000	.779**	.741**	.776**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000
	N	297	298	298	300	296	297	295
Complexity of Thought	Pearson Correlation	.706**	.791**	.794**	.779**	1.000	.704**	.722**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000
	N	296	297	297	296	298	298	293
Command of English	Pearson Correlation	.653**	.705**	.703**	.741**	.704**	1.000	.715**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000
	N	298	298	298	297	298	300	295
Overall	Pearson Correlation	.767**	.819**	.800**	.776**	.722**	.715**	1.000
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	
	N	295	295	295	295	293	295	305

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Cronbach's alpha shows strong internal consistency of the scale. The alpha of .951 indicates that each of the six components of the rubric measures the same thing (presumably, writing skill). However, with such a high alpha, we may actually consider the possibility that some of the items are redundant and the rubric could be simplified. Alphas for pairs of subscales range from .790 to .926 (see Table A5).

**Table A5: Cronbach's Alpha**

		<b>Cronbach's Alpha</b>
Statement of Thesis	Command of English	0.790
Assembly of Evidence	Command of English	0.826
Complexity of Thought	Command of English	0.826
Statement of Thesis	Complexity of Thought	0.827
Supporting Arguments	Command of English	0.827
Statement of Thesis	Assembly of Evidence	0.848
Paragraph Development	Command of English	0.852
Statement of Thesis	Paragraph Development	0.869
Paragraph Development	Complexity of Thought	0.876
Supporting Arguments	Complexity of Thought	0.883
Assembly of Evidence	Complexity of Thought	0.885
Statement of Thesis	Supporting Arguments	0.891
Assembly of Evidence	Paragraph Development	0.913
Supporting Arguments	Paragraph Development	0.924
Supporting Arguments	Assembly of Evidence	0.926

Here we see that statement of thesis and command of English is the pair with the smallest alpha, although it is still very strong at .790. Additionally, the following pairs have alphas of more than .90: assembly of evidence & paragraph development, supporting arguments & paragraph development, and supporting arguments & assembly of evidence.

# Appendix B: Recruitment Materials

Dear WI Instructors,

I write to ask you to participate in an important pilot project for fall semester. The project comes as a recommendation from the Writing Committee, who consider it a vital part of our efforts to develop our students' writing abilities.

The project will comprehensively assess how well our first-years' writing abilities meet faculty expectations after one term. The results will enable us to decide whether we want to make first-year writing assessment a permanent feature of a Smith education and further serve to identify particularly weak writers early in their college careers.

We have structured the project to involve minimal intrusion into your course. We request that you ask your students to keep portfolios of the papers they write for your course and that you collect the portfolios at the end of the semester. Finally, we ask that you select what you think is each student's best paper. We are not asking you to assess the writing.

We hope for full participation to enable us to get a detailed picture of our students' writing abilities. Please e-mail Hayley Spizz to indicate your willingness to participate ([hspizz@email.smith.edu](mailto:hspizz@email.smith.edu)). The Writing Committee is planning a light reception for all participating faculty on Friday, September 19, 4:30-5:30, to discuss the project in more detail, to answer any questions you might have, and to give you a sense how this project fits within our larger vision for the future of student writing at Smith.

The Writing Committee and I look forward to our work together.

Sincerely,

Susan C. Bourque  
Provost

Dear WI Instructor,

Last week Provost Susan Bourque wrote to you asking you to participate in an important pilot project for fall semester. We hope you have e-mailed Hayley Spizz to express your willingness to participate. If you have not, we hope you will now ([hspizz@email.smith.edu](mailto:hspizz@email.smith.edu)). We consider this project a vital part of our efforts to develop our students' writing abilities.

We have conducted a number of writing-assessment studies over the past few years and are now at the point of deciding whether we want to recommend that first-year writing assessment become part of the Smith curriculum. To do that with confidence, however, we will need a comprehensive understanding of our entering students' abilities after one term. We are particularly interested in identifying weak writers early in their college careers and providing them with more instruction in the spring semester and the sophomore year.

We have structured the project to involve minimal intrusion into your course. We ask the following:

1. We ask you to ask your students to keep clean copies of all the papers they write in your course in a folder. We also ask that you monitor, in whatever way you deem appropriate, that students are adding papers to the folders as they write them. We will provide you with the folders.
2. We ask that you collect the folders on the last day of classes.
3. We ask that you select what you deem to be each student's best paper, identify it with a check mark, and attach the assignment for that paper, if the student has not already done so. You need not do this for students in your WI course who are beyond the first year.
4. Finally, we ask that you mail or deliver the folders, labeled with student names and with the best papers identified, to the office of Institutional Research & Educational Assessment in Room 305 of College Hall. We ask that you do this by January 8, 2009, the date grades are due. The papers will be used in a faculty assessment exercise during the month of January.

We encourage you to use the attached writing rubric, now familiar to many of you, in selecting each student's best paper. We also invite you to use the rubric in teaching your students to write well. It may help you explain what constitutes good writing and what the specific writing goals of WI courses are.

We invite all participating faculty to a light reception on Friday, September 19, 4:30-5:30 p.m., in Seelye 207, when we will discuss the project in more detail, answer any questions you might have, and give you a sense how this project fits within our larger vision for the future of student writing at Smith.

We look forward to working with you on this project and to seeing you on September 19.

We thank you in advance.

All the best,

Doreen Weinberger  
Associate Professor, Physics, & Co-chair, Writing Committee

Julio Alves  
Director, Jacobson Center, & Co-chair, Writing Committee

Attachment: WI Writing Assessment Rubric

December 9, 2008

Dear Participating WI Instructor,

Thank you for your involvement in the WI writing assessment. The Provost's office, writing committee, and I are excited about the opportunity to collect data about student performance on writing from a cross section of the entering class. The results of this study will inform our efforts in further developing the WI requirement, and in designing interventions beyond the WI course for students who need more help with writing after their first semester.

In an effort to minimize the administrative burden associated with WI courses, we have attempted to design a process that will involve minimal intervention on the part of the instructor. As you know, we have asked instructors to collect a portfolio of clean, unmarked papers, and to turn those portfolios in to the IR office. You may do this by dropping them off in the registrar's office or the IR office in College Hall, by dropping them off at the Jacobson Center, or by sending them via campus mail to the IR office. If you would like to arrange for an alternate collection method, please contact Ingrid Boedker at 3021.

Given that this is a new assessment process, we would like to give students a clear understanding of how their papers will be used in the assessment process. Would you please distribute the following statement in class, read it, or post it to your moodle course? This statement will give students information on how they may opt out or make their participation anonymous. If a student decides to opt out, there should be no consequences associated with that choice. Please just ask her not to turn in a portfolio. If she would like to participate with her identity masked, she can follow the instructions in the letter below.

Thank you again for your participation in this project. Please do not hesitate to call me at x3022 with any questions.

Sincerely,

Cate Rowen  
Director of Institutional Research & Educational Assessment

Dear WI Instructor,

Thank you again for agreeing to participate in the 2008-09 WI Writing Assessment. This letter will provide you with information about how to complete your participation.

As you know, we've asked participating WI instructors to work with students to keep portfolios of student work throughout the semester, and to turn those portfolios in at the end of the semester.

- We ask that for each **first-year** student in your class, you review the portfolio and **mark the one paper that you deem to be the student's best writing**. You may mark that paper with a check, the notation "best," or any other obvious notation on the paper.
- Please **include assignments for all the papers you submit**. The most convenient way to do this is probably to provide a packet of assignments for your course, and to double-check that every "best" paper clearly indicates which assignment it is for. If this is not clear on the paper, please indicate it along with your mark of "best," either by writing the name or number of the assignment on the paper.
- As you know, students may opt not to participate, or they may choose to turn in papers with their names removed. Some faculty members have only collected "best" papers from students; we will accept these papers as well.
- When you have assembled the students' portfolios and marked the best papers, all you need to do is **get them to the Institutional Research & Educational Assessment office**. There is a variety of ways to do this.
  - If you have paper copies,
    - You may drop them off in any of the following locations:
      - The Jacobson Center
      - The Registrar's Office
      - The IR Office (305 College Hall)
      - The Clark Science Center Office (Burton 115);
    - You may send them via campus mail. Please address the envelope(s) to Ingrid Boedker, IR, 305 College Hall.
  - If you have electronic copies,
    - You may burn them to a CD or store them on a flash drive (to be returned to you) and send them via campus mail.
    - You may transmit them via email to [iboedker@email.smith.edu](mailto:iboedker@email.smith.edu).
  - You may arrange for other methods of transmittal by contacting Ingrid at extension 3021 or at the above email address.

We hope that this process is not burdensome for you, and if there is anything we can do to make it easier, we are eager to assist. Please contact me or Ingrid to discuss any questions you may have.

Thank you again for your involvement in this project. We are excited about the writing assessment process and the opportunity it will give us to further our understanding of Smith's writing program, and of the needs of first-year Smith students.

All the best,

Cate Rowen  
Director of Institutional Research & Educational Assessment  
[crowen@email.smith.edu](mailto:crowen@email.smith.edu)  
3022

Dear WI Student,

Your instructor has agreed to participate in an assessment project on student writing. The purpose of this assessment is to help the College gain a better understanding of writing skills and needs among first-year students who have just completed a WI course. The results of the study will be used to provide feedback to WI instructors about the areas of strength and weakness among WI course participants, and to design courses and support mechanisms that can support students who need additional writing instruction after the WI course.

To do this, your instructor will ask you to provide a clean, unmarked copy of each of the assignments you completed for his or her WI course this semester. Your instructor will submit this portfolio of assignments to Smith's Office of Institutional Research and Educational Assessment, where it will be coded with a case number. This case number will allow the IREA office to link the scores your portfolio receives with other academic and demographic information about you. No individually-identifiable information, however, will be reported at any time.

One paper from your portfolio will be read by one or more faculty members at a three-day workshop during January. Readers will review your writing and will score it using a rubric based on the learning outcomes established for WI courses. The results of that scoring will be stored securely in a research database that will be used to prepare a report about writing skills among first-year Smith students. This report will present all data in group form: no student names will ever be used. Additionally, the scores you receive on this paper will never be linked with your academic record: they are for research purposes only. After the study is over, your portfolio will be stored in the Office of Institutional Research & Educational Assessment and may be used for additional research on other topics.

**Your participation in this project is entirely voluntary.** You may refuse to take part in the study at any time without affecting your relationship with your instructor or Smith College. Your decision will not result in any loss of benefits to which you are otherwise entitled. If you do not wish to participate, you may inform your instructor that you would like to opt out, and he or she will not submit a portfolio for you. If you wish to participate, but to have your identity masked from portfolio readers, you may remove your name from the papers you submit in your portfolio. Please insert a separate piece of paper in your portfolio indicating your identity (name or 99 number) so that we can provide a code that will allow us to access your demographic information. We will remove the identifying paper from your portfolio before it is reviewed at the workshop.

If you have any questions about this assessment, or about your participation in it, please contact me or Ingrid Boedker in the Office of Institutional Research & Educational Assessment (x3021) or Julio Alves in the Jacobson Center. Thank you in advance for your participation.

Best,

Cate Rowen  
Director of Institutional Research & Educational Assessment

From: The Committee on College Writing  
To: First-Year Students in Writing-Intensive Courses

We would like you to become part of a project aimed at bettering writing instruction at Smith, for your class and for future first-year classes. The Committee on College Writing has asked this fall's teachers of writing-intensive courses to collect unmarked copies of all the graded essays you write for their courses, to indicate the strongest essay, and to pass the essays on to us. In January we will read through the essays to determine what students have learned over the semester, and how ready they are for written assignments in later college courses. Once we have a sense of where students are at this point in their careers, we'll be able to plan more accurately for more advanced courses in writing, and we'll be able to fine-tune our first-year writing-intensive courses for future generations of students.

The results of the study will not, of course, be recorded on your transcript, or in any permanent file at Smith. But we believe that your participation in the project will have a side-benefit for you as a writer. Your advisors will be able to advise you more accurately about your strengths and weaknesses as a writer, and recommend courses that will further strengthen your skills.

We ask you to participate in this study. If your instructor asks you to print the essays out yourselves, we would be happy to pay for the extra expense. Just notify Julio Alves, Director of the Jacobson Center for Writing, Teaching and Learning, and he'll recompense you for the copying.

Thanks. This study will be valuable to future generations of Smith students, who will benefit from what we learn.

Yours,

Bill Oram, Chair  
College Committee on Writing

**To: Smith College Faculty Members**  
**From: The Committee on College Writing**

We'd like to invite you to participate in a group that will meet this January to evaluate how well our first-year students write after one writing-intensive course.

We (roughly twenty-five faculty members) will meet in for three mornings and an afternoon in January (the mornings of January 21, 22 and 23 and the afternoon of the 23<sup>rd</sup>). The morning will be spent reading the best paper from each of the first-year students who have taken writing-intensive courses this fall, and evaluating it according to a set of carefully defined criteria. During the final afternoon we'll talk about what we've found out about the quality of first-year writing, and we'll consider what we should do with our conclusions.

Participants will receive an honorarium of six hundred dollars, but there is a good deal more to be gained. Those who did a similar study eighteen months ago thought that they learned a good deal about what to look for in writing, and how to go about remedying the problems that students often have. And we promise that the lunches will be good.

We hope that you'll be willing to join the group. To put yourself on the list, send an email to Hayley Spizz at [hspizz@email.smith.edu](mailto:hspizz@email.smith.edu).

Thanks very much.

*The Committee on College Writing*  
Bill Oram, English, Chair  
Julio Alves, Jacobson Center, Co-Chair  
Karen Pfeiffer, Economics  
Sharon Seelig, English  
Doreen Weinberger, Physics  
Susan Etheredge, Education & Child Study

**To: Smith College Faculty Members Participating in January Reading**  
**From: The Committee on College Writing**

Thank you for your willingness to participate in the group that will meet this January to evaluate how well our first-year students write after one writing-intensive course.

As you know, we (roughly twenty-five faculty members) will meet in for three mornings (9-12:30) and an afternoon (1:15-2:45) in January (the mornings of January 21, 22 and 23 and the afternoon of the 23<sup>rd</sup>). The morning will be spent reading the best paper from each of the first-year students who have taken writing-intensive courses this fall, and evaluating it according to a set of carefully defined criteria. During the final afternoon we'll talk about what we've found out about the quality of first-year writing, and we'll consider what we should do with our conclusions.

You'll receive an honorarium of \$600 for your participation.

Even though we meet only until 12:30 the first two days, we'll provide lunch.

If you have any questions, please send them to Hayley Spizz at [hspizz@email.smith.edu](mailto:hspizz@email.smith.edu), who will collect questions and pass them on to us. We'll get back to you as soon as we can.

Thanks very much.

*The Committee on College Writing*  
Bill Oram, English, Co-Chair  
Julio Alves, Jacobson Center, Co-Chair  
Karen Pfeiffer, Economics  
Sharon Seelig, English  
Doreen Weinberger, Physics  
Susan Etheredge, Education & Child Study

## Appendix C: Assessment Rubric, Scoring, & Reading Procedure

## Writing Assessment Rubric

### 1. Statement of the thesis (or main goal)

*Exemplary:* The paper makes a bold central claim that is substantive, articulately stated, and thought-provoking.

*Adequate:* The paper makes a fine central claim, but it could be more substantive, articulately stated, and/or thought-provoking.

*Needs work:* The paper makes no central claim, or one that is irrelevant, unclear, or self-evident.

### 2. Development of the thesis through supporting arguments

*Exemplary:* The thesis is broken down into a series of logical, effectively sequenced supporting arguments.

*Adequate:* The thesis is broken down into a series of supporting arguments, but the arrangement manifests minor lapses in logic and sequencing.

*Needs work:* The thesis is not broken down into a series of supporting arguments, or the arrangement of the arguments manifests major flaws in logic and sequencing.

### 3. Assembly and presentation of evidence

*Exemplary:* Arguments are supported with plentiful, carefully selected, fully convincing evidence. The evidence is carefully and seamlessly arranged, leaving the reader confident and eager to move forward.

*Adequate:* Arguments are supported with sufficient evidence, but its presentation manifests minor lapses in selection and arrangement. The reader is satisfied and ready to move forward.

*Needs work:* The arguments are unsupported or weakly supported with evidence, and the presentation of the evidence shows major lapses in selection and arrangement. The reader is unconvinced, disoriented, frustrated, and reluctant to move forward.

### 4. Paragraph development

*Exemplary:* Paragraphs are structured around a clearly stated, well-positioned central argument. The sentences are coherent (clearly linked to each other) and unified (focused on the central argument). Paragraphs fully present and analyze supporting evidence.

*Adequate:* Paragraphs are structured around a central argument, but its statement could be clearer or better positioned. The sentences may show minor lapses in coherence and unity. The presentation and analysis of the evidence are present, but one or the other may be underdeveloped.

*Needs work:* The paragraphs do not state a central argument or state it unclearly. The sentences show major lapses in coherence and unity. The presentation and analysis of the evidence are minimal, or one or both may be missing. 2

## **5. Overall complexity of thought**

*Exemplary:* The complexity of thought in the paper impresses. The discussion is intricate and original, stretching well beyond what is immediately apparent.

*Adequate:* The complexity of thought in the paper pleases, but does not impress. The discussion is competent and mildly original, but it does not stretch, or stretches very little, beyond what is immediately apparent.

*Needs work:* The paper shows little or no complexity of thought. The discussion is oversimplified, unoriginal, and self-evident, ultimately boring.

## **6. Overall command of written English**

*Exemplary:* The paper is free of errors in grammar, punctuation, and mechanics. The writer's voice is engaging and the prose fluent.

*Adequate:* The paper manifests minor, occasional errors in grammar, punctuation, and style. The writer's voice is appropriate and the prose competent.

*Needs work:* The paper manifests major, persistent errors in grammar, punctuation, and style. The writer's voice is inappropriate or inconsistent, and the prose substandard.

## **7. Overall**

*Exemplary:* This paper exceeds expectations for a writer at this point in her college career.

*Adequate:* This paper meets expectations for a writer at this point in her college career.

*Needs work:* This paper does not meet expectations for a writer at this point in her college career.



JANUARY READING SESSIONS  
JANUARY 21 – 23, 2009  
MARY MAPLES DUNN ROOM  
PIERCE HALL, SECOND FLOOR

**SCHEDULE**

Wednesday, January 21

Noon – 1:00 p.m. Lunch and Introduction  
1:00 – 1:30 p.m. Norming Exercise  
1:30 – 4:00 p.m. Paper Reading (*refreshments available at 2:30 p.m.*)

Thursday, January 22

9:00 – 9:15 a.m. Refreshments and Introduction  
9:00 a.m.–12:30 p.m. Paper Reading (*refreshments available at 10:30 a.m.*)  
12:30 – 1:30 p.m. Lunch

Friday, January 23

9:00 – 9:15 a.m. Refreshments and Introduction  
9:15 a.m. – 12:30 p.m. Paper Reading (*refreshments available at 10:30 a.m.*)  
12:30 – 1:30 p.m. Lunch  
1:30 – 2:30 p.m. Paper Reading  
2:30 – 2:45 p.m. Conclusions

Readers

Christine Andrews, Art  
Ryan Brandau, Music  
Jack Czajkowski, Education and Child Study  
Sara Eddy, English Language & Literature  
Katherine Halvorsen, Mathematics & Statistics  
Virginia Hayssen, Biological Sciences  
Alice Hearst, Government  
Jonathan Hirsh, Music  
Robert Hosmer, English Language & Literature  
Barbara Kellum, Art  
Jina Kim, East Asian Studies  
Jocelyne Kolb, German Studies  
Daphne Lamothe, Afro-American Studies  
Denise Lello, Biological Sciences  
Larry Meinert, Geology  
Caroline Melly, Anthropology  
James Middlebrook, Art  
Katwiwa Mule, Comparative Literature  
William Oram, English Language & Literature  
Cornelia Pearsall, English Language & Literature  
Meg Thacher, Astronomy  
Tina Wildhagen, Sociology  
Sujane Wu, East Asian Languages and Literatures  
Suzanne Zhang-Gottschang, Anthropology

Other Participants

Julio Alves, Director, Jacobson Center for Writing, Teaching and Learning  
Cate Rowen, Director, Institutional Research and Educational Assessment  
Hayley Spizz, Office of the Provost/Dean of the Faculty