SMITH COLLEGE

First-Year Writing Intensive Assessment: 2012-13

April 2013
Jacobson Center for Writing, Teaching and Learning
Office of Institutional Research

Contents

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. According to the Smith College Course Catalogue, 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 compose paragraphs that are appropriately focused and fully developed with evidence, explanation, and analysis;

3) to identify and evaluate, when appropriate, 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 edit work until it is orderly, clear, and free of violations of the conventions of standard written English (grammar, usage, punctuation, diction, syntax).

To achieve these goals, a writing-intensive course requires:

1) students to write a minimum of 15 pages of finished work (not counting drafts and revisions) in at least 3 papers, one of which must be at least 5 pages.

2) students to write in drafts and instructors to provide feedback in at least 2 papers while writing is still in process so that students may incorporate instructors' comments and thus prepare better work.

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 office of admission) 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.

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 aimed to provide an objective evaluation of writing from the perspective of faculty across the college. The 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. A report detailing the results of the 2008-09 writing assessment is available from the Office of Institutional Research.

In January 2013 members of the writing committee worked with the Jacobson Center for Writing, Teaching, and Learning to conduct this writing assessment project a second time. By readdressing the 08-09 assessment, the current project aims to identify how well students write after one writing intensive course, to identify where students' writing is most and least successful, and to ascertain where attention needs to be focused.

In August 2012, the 47 faculty teaching either ENG 118 or FYS courses during Fall 2012 were sent a letter from Julio Alves, co-chair of the Writing Committee and Director of the Jacobson Center and the Writing Program, and Alice Hearst, co-chair of the Writing Committee and Director of First Year Seminars, encouraging them to participate in this assessment, though faculty participation was voluntary. Language about the assessment was provided for instructors to include in their syllabi. Students were asked to prepare a portfolio including what they considered their best thesis-driven paper, a copy of the assignment for the paper, and a reflection explaining why they chose that piece of writing.

First-year students enrolled in the participating WI courses were informed how their papers would be used for the study. In some cases, faculty collected the writing folders from students. Other students were asked to turn in their writing folders at the Registrar's Office or the Jacobson Center, which were then collected, checked for inclusion of all materials, and randomized. A total 420 folders were collected, 83 of which were considered incomplete and removed from the sample,

leaving 337 complete writing samples. Folders were considered incomplete if they did not include the assignment sheet or a reflection in addition to the writing sample.

Finally, all Smith faculty were sent a letter from the Committee on College Writing and Provost Marilyn Schuster inviting them to participate as readers. Volunteers were offered an honorarium 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 B.

A total of 164 student papers were read and scored, representing 30 sections of 22 courses :

- ENG 118: Colloquia in Writing, Section 01 and 03: The Politics of Language; Section 02: To Hell and Back, Trauma and Transformation; Section 04: All the World's a Stage; Section 05: Water, Science and Politics; Section 06: Consumer Culture; Section 07: Language and Power; Section 08: No, Seriously…What's so Funny? Writing about Humor; Section 09: Fakes, Forgeries, and Imposters
- FYS 102: Animal Rights
- FYS 104: God and Evil
- FYS 105: Jerusalem
- FYS 106: Growing Up Asian American
- FYS 108: Curry: Gender, Race, Sexuality and Empire
- FYS 109: Exobiology: Origins of Life and the Search for Life in the Universe
- FYS 114: Turning Points
- FYS 116: Kyoto Through the Ages
- FYS 121: The Evolution and Transformation of the Northampton State Hospital
- FYS 130: Lions: Science and Science Fiction
- FYS 132: Physics for Future Presidents
- FYS 158: Reading the Earth
- FYS 164: Issues in Artificial Intelligence
- FYS 167: Viking Diaspora
- FYS 175: Love Stories
- FYS 179: Rebellious Women
- FYS 191: Sense and Essence in Nature
- FYS 192: America in 1925
- FYS 195: Health and Wellness: Personal Perspectives
- FYS 198: The Global Coffee Trail
- FYS 199: Re-Membering Marie Antoinette

Twelve faculty and seven staff members who also teach gathered in January 2013 to read the papers. While some of the readers were also instructors whose students submitted papers for the assessment, many were not. If WI instructors happened to draw papers written for their classes, they were asked to return those papers and take others. Readers represented the following academic departments: Anthropology (1), Chemistry (1), Education (1), English (3), Exercise and Sports Science (1), Government (2), Italian (1), Physics and Astronomy (1), Religion (1), and Theatre (1). The following offices were represented: Jacobson Center (5), Class Deans Office (1), and Athletics (1).

Student papers 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. The assessment

was adapted from a process developed at Carleton College. Student papers were rated on a three-point scale of "needs work," "pass," or "exemplary." Readers were given 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:

- Statement of thesis
- 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. Readers were also asked 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 twice. Fifty eight of the papers that were read twice received the same overall score from both readers. Eighty seven 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. No papers required a third reader to settle irreconcilable differences in opinions between the first two scorers.

A total of 396 scores (for 164 papers) were recorded. The initial readings that led to a negotiation were recorded but not included in the final analysis—only the negotiated scores went into the final analysis. A weighting variable was used to account for the fact that some papers were read once while others were read twice. When a paper was read only once, that record was weighted as 1.0 for the analysis. For papers that were read twice and resulted in the same overall score, the scores were averaged by applying a weight of 0.5 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 disregarded (weighted as 0) and the negotiated record was used (weighted as 1.0).

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 passing, with mean ratings in the sub-skills ranging from 4.58 to 4.78 points out of a possible 9 (with 9 being strongest). The mean for the overall score is 1.83 (out of 3, with 3 being strongest). Within each sub-skill, the majority of scores fall between high-needs work and high-pass. Low-needs work and mid-exemplary are infrequently awarded scores. Remarkably, only *one* paper was scored as high-exemplary, and only in the sub-skill of Statement of Thesis. These results were very similar to the 2009 assessment, in which 30.7% of papers were scored as *needs work,* 56.8% were given *pass* ratings, and 12.5% achieved an *exemplary* rating.

Table 1: Overall Scores, Detail

Writing Assessment 2012-13
Table 1: Overall Scores

n= 164	Needs Work			Pass			Exemplary		
	Low	Mid	High	Low	Mid	High	Low	Mid	High
Statement of Thesis	1.8%	9.1%	17.1%	19.2%	22.0%	19.5%	7.9%	3.0%	0.3%
Paragraph Development	3.0%	8.2%	19.5%	11.6%	23.5%	21.3%	10.1%	2.7%	---
Complexity of Thought	1.8%	8.6%	18.3%	19.9%	18.3%	19.0%	9.8%	4.3%	---
Command of English	0.9%	7.3%	18.0%	12.8%	26.2%	20.4%	10.7%	3.7%	---
Overall Score		29.3%			58.5%			12.2%	

Writing Assessment 2012-13:
Summary of Scores

■ Needs Work   ■ Pass   ■ Exemplary

Smith College Office of Institutional Research.

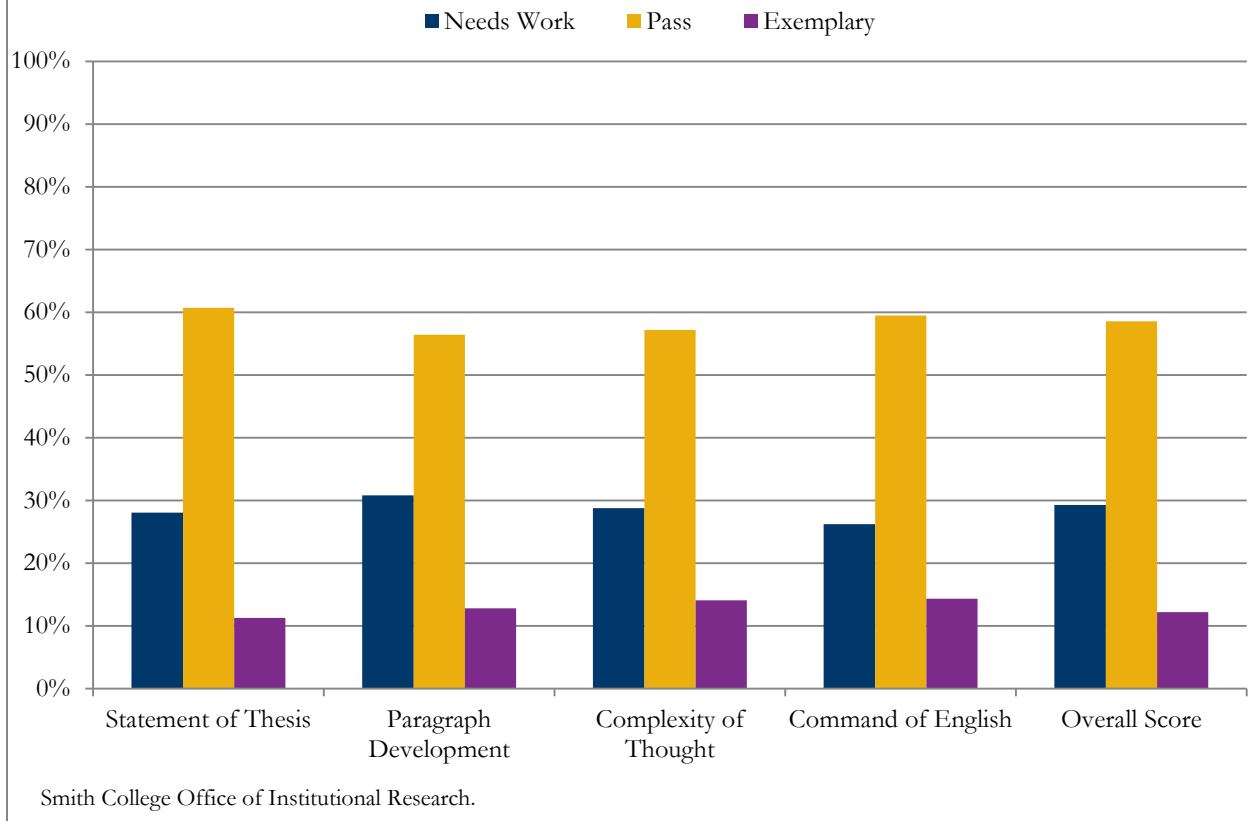**Table 2: Overall Scores, Summarized**

Writing Assessment 2012-13

Table 2: Overall Scores Summarized

n=164	Needs Work	Pass	Exemplary
Statement of Thesis	28.0%	60.7%	11.3%
Paragraph Development	30.8%	56.4%	12.8%
Complexity of Thought	28.7%	57.2%	14.1%
Command of English	26.2%	59.5%	14.3%
Overall Score	29.3%	58.5%	12.2%

Table 3: Mean Scores

Writing Assessment 2012-13

Table 3: Mean Scores

n=164	Mean
Statement of Thesis	4.58
Paragraph Development	4.62
Complexity of Thought	4.62
Command of English	4.78
Overall Score (1-3)	1.83

Tables 1 and 2 indicate that Smith first-years are generally strongest in the areas of Command of English and Complexity of Thought, with over 14% of papers rated as exemplary in these areas. Table 3 shows relatively high mean scores for these skills, as well as Paragraph Development.

Upon closer examination, Paragraph Development seems to be the sub-skill with the greatest variation in score. Although the mean is the same as Complexity of Thought, Paragraph Development actually has the largest percentage of papers needing work. While 30% of papers needed work in that area, almost 13% were exemplary in it. In 2009, Paragraph Development had similar results, with 32% needing work, and 15% being exemplary.

Despite concerns about the increasing proportion of international students at Smith, a group that is adequately represented in this sample, Command of English had the highest mean score, and also the lowest proportion of papers viewed as needing work. Later in this report we delve into the performance of international, and specifically Chinese, students in more detail. While scores on Command of English improved slightly between 2009 and 2013, the gains are not statistically significant.

Statement of Thesis is the sub-skill with the lowest mean score and the smallest percentage of exemplary papers. However, it also has the highest proportion of pass results. Clearly, this is an area in which many Smith students are competent, but few excel. The rubric changed from 2009 to 2013 in the area of thesis and argumentation, so they are not strictly comparable; 2009 results showed a higher proportion of students scoring exemplary in both categories (15.2% in 2009 for thesis, 14.3% in 2009 for supporting arguments, vs. 11.3% in 2013).

Next, we examined the differences among ratings in the two types of writing intensive courses included in the study (ENG 118 and First Year Seminars). The following chart displays the overall score by course type and shows a statistically significant difference. Most notable is the difference in the exemplary category (16.5% exemplary among the FYS papers and only 3.6% among the ENG 118 papers). This is not unexpected, given that ESOL students and those with weak writing skills are encouraged to enroll in English 118.

Writing Assessment 2012-13:
Overall Score* by WI Course Type

■ First Year Seminar ■ English 118

*Difference is statistically significant (p<.05 chi-square)
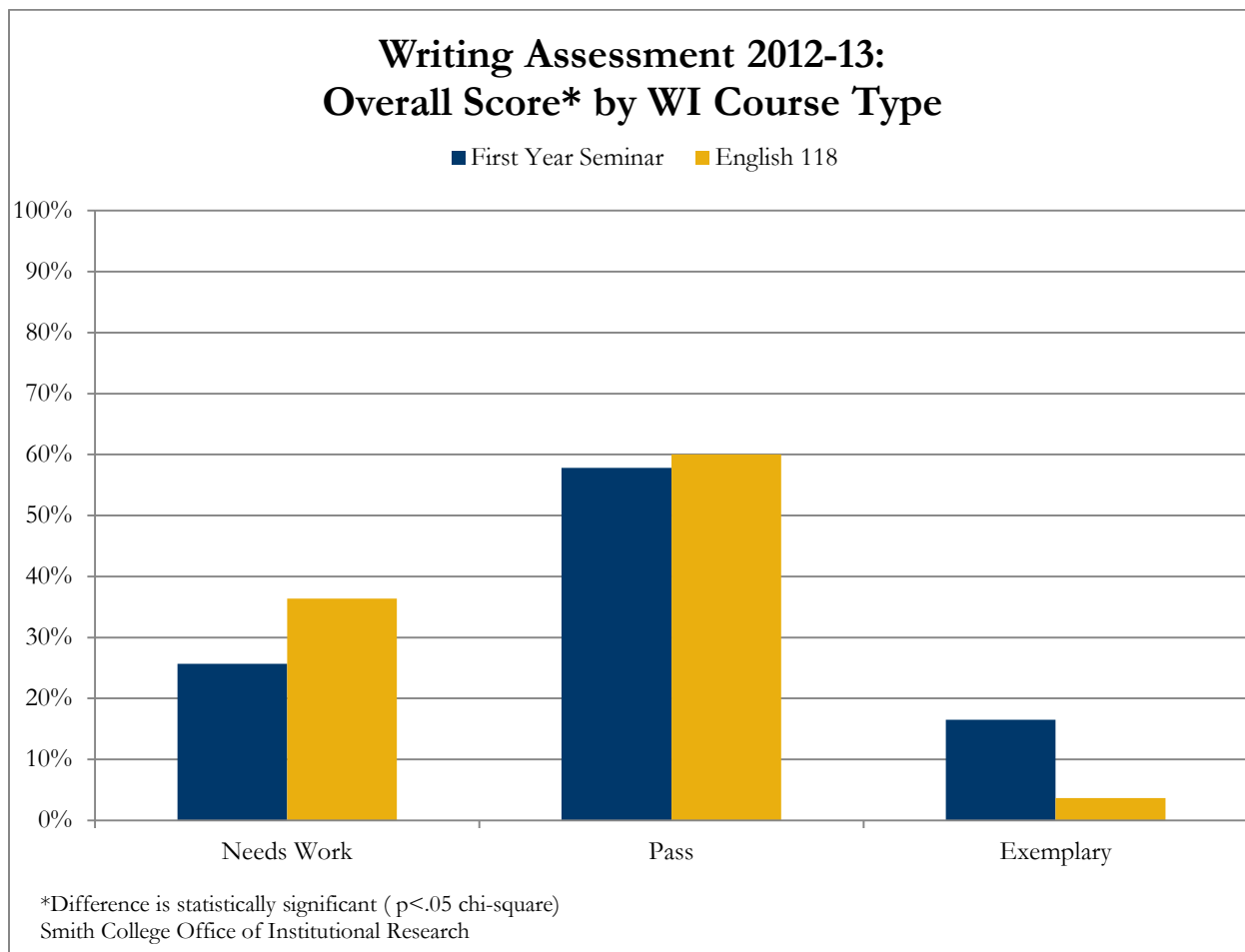Smith College Office of Institutional Research

Table 4, below, provides more detail on the difference by type of WI course. Statistically significant differences between the two types of WI courses appear for the sub-skills Complexity of Thought and Command of English. In both of these skill areas, substantially fewer ENG 118 papers were rated as exemplary, and more were found to need work.

Table 4: Course Type

Writing Assessment 2012-13
Table 4: Course Type

		First Year Seminar (N = 109)	English 118 (N = 55)
Statement of Thesis	Needs Work	27.5%	29.1%
	Pass	58.7%	64.5%
	Exemplary	13.8%	6.4%
Paragraph Development	Needs Work	28.0%	36.4%
	Pass	56.0%	57.3%
	Exemplary	16.1%	6.4%
Complexity of Thought*	Needs Work	26.6%	33.0%
	Pass	54.1%	63.3%
	Exemplary	19.3%	3.7%
Command of English*	Needs Work	23.9%	30.9%
	Pass	56.4%	65.5%
	Exemplary	19.7%	3.6%
Overall Score*	Needs Work	25.7%	36.4%
	Pass	57.8%	60.0%
	Exemplary	16.5%	3.6%

*Difference is statistically significant (p<0.05 chi-square)

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: our analysis only included 25 first-generation and 141 non-first-generation student papers. When the demographic groups we are comparing are this small and separated into three score categories, seemingly large differences in percentages can be attributed to just 1 or 2 individuals. This explains why the apparent score differences between first-generation and non-first-generation students were not statistically significant. Research that targets this group with an over-sample of first-generation student papers would be the next step in further exploring this group.

Overall, 41.7% of papers by first-generation students need work, as compared to just 27.1% of papers by students who are not first-generation. Similar disparities between the two groups are seen for each sub-skill. In 2009, results were also statistically similar, with 34% of first-generation students' papers needing work, and 30% of non-first-generation students being in the lowest category.

Table 5: First Generation Students

Writing Assessment 2012-13
Table 5: First Generation Students^

		Non-First Generation (N = 141)	First Generation (N = 25)
Statement of Thesis	Needs Work	26.8%	35.4%
	Pass	60.7%	60.4%
	Exemplary	12.5%	4.2%
Paragraph Development	Needs Work	29.3%	39.6%
	Pass	56.8%	54.2%
	Exemplary	13.9%	6.3%
Complexity of Thought	Needs Work	26.5%	41.7%
	Pass	58.1%	52.1%
	Exemplary	15.4%	6.3%
Command of English	Needs Work	24.3%	37.5%
	Pass	60.4%	54.2%
	Exemplary	15.4%	8.3%
Overall Score	Needs Work	27.1%	41.7%
	Pass	59.3%	54.2%
	Exemplary	13.6%	4.2%

^ Please use caution when interpreting results due to small cell sizes

Differences are not statistically significant

If we divide the study participants into a high and low group based on their SAT Writing scores, a 670 score is the dividing mark (i.e., half the participants scored above 670 and half scored at or below 670). Table 6 presents the results of the assessment by high and low SAT Writing category.

Generally, we do see the expected trend of those in the higher SAT Writing category receiving better scores on their papers, though no differences were statistically significant. As with the data in Table 5, cell sizes are very small and must be interpreted with care.

Table 6: SAT Writing

Writing Assessment 2012-13
Table 6: High or Low SAT Writing Score

		Low SAT Writing Score (<670) (N = 52)	High SAT Writing Score (670 or higher) (N = 53)
Statement of Thesis	Needs Work	26.0%	22.6%
	Pass	65.4%	60.4%
	Exemplary	8.7%	17.0%
Paragraph Development	Needs Work	36.5%	24.5%
	Pass	52.9%	54.7%
	Exemplary	10.6%	20.8%
Complexity of Thought	Needs Work	31.7%	23.6%
	Pass	58.7%	54.7%
	Exemplary	9.6%	21.7%
Command of English	Needs Work	25.0%	25.5%
	Pass	65.4%	50.0%
	Exemplary	9.6%	24.5%
Overall Score	Needs Work	30.8%	24.5%
	Pass	61.5%	54.7%
	Exemplary	7.7%	20.8%

^ Please use caution when interpreting results due to small cell sizes

Differences are not statistically significant

Next we examined writing assessment scores by the reader rating used in Smith's admission process. Each applicant to Smith is given a numeric rating during the admission process, with a 1 being the best rating an applicant can receive. For this analysis, we divided the participating students into two groups: High Reader Rating (1-4) and Low Reader Rating (5-8).

Table 7 shows that first-years who had better (high) reader ratings tend to score better on the writing assessment, as would be expected. For each sub-skill the group with high reader ratings has a larger percentage of exemplary scores. And for three of the four sub-skills the high reader rating group has a smaller percentage of papers needing work. The exception is Command of English, on which 29.2% of high reader rating papers needed work, as compared to 24.5% of low reader rating papers. These apparent differences are not, however, statistically significant, and the small cell sizes require that we interpret these results cautiously. In 2009, with a larger sample size and therefore greater statistical power, we saw similar results, with statistically significant differences between the two groups on Supporting Arguments, Paragraph Development, Complexity of Thought, and Command of English.

Table 7: High or Low Reader Rating

Writing Assessment 2012-13
Table 7 : High or Low Reader Rating^

		High Reader Rating (N = 60)	Low Reader Rating (N = 104)
Statement of Thesis	Needs Work	24.2%	30.3%
	Pass	60.0%	61.1%
	Exemplary	15.8%	8.7%
Paragraph Development	Needs Work	26.7%	33.2%
	Pass	58.3%	55.3%
	Exemplary	15.0%	11.5%
Complexity of Thought	Needs Work	23.3%	31.9%
	Pass	54.2%	58.9%
	Exemplary	22.5%	9.2%
Command of English	Needs Work	29.2%	24.5%
	Pass	50.0%	64.9%
	Exemplary	20.8%	10.6%
Overall Score	Needs Work	26.7%	30.8%
	Pass	53.3%	61.5%
	Exemplary	20.0%	7.7%

^ Please use caution when interpreting results due to small cell sizes

Differences are not statistically significant

Table 8: Citizenship

Table 8 presents the results of an analysis by citizenship. The very small number of international students in this analysis made it necessary to collapse the writing score categories. This table provides and examination of those papers that need work and those that do not.

Here we do see some statistically significant differences between domestic and international students. Specifically, domestic students are less likely to need work in the area of Command of English and Overall Score. Considering that many international students may not be native English speakers, this is not surprising. As seen in the overall findings, both groups had the highest proportion of needs work ratings in the area of Paragraph Development. International first-years appear to be next weakest in Command of English, while domestic first-years seem to need the most work on Complexity of Thought.

Writing Assessment 2012-13
Table 8: Citizenship^

		International (N = 25)	US Citizen or Permanent Resident (N = 139)
Statement of Thesis	Needs Work	38.0%	26.3%
	Pass or Exemplary	62.0%	73.7%
Paragraph Development	Needs Work	46.0%	28.1%
	Pass or Exemplary	54.0%	71.9%
Complexity of Thought	Needs Work	32.0%	28.2%
	Pass or Exemplary	68.0%	71.8%
Command of English*	Needs Work	44.0%	23.0%
	Pass or Exemplary	56.0%	77.0%
Overall Score*	Needs Work	48.0%	25.9%
	Pass or Exemplary	52.0%	74.1%

^ Please use caution when interpreting results due to small cell sizes
*Differences are statistically significant ($p < 0.05$ chi-square)

Table 9: International Students by Country

Finally, we looked a little more closely at the international student papers. Almost exactly half of the international students in this study were Chinese citizens, so the next table compares that group to all other international first-years in this study. The score categories are once again collapsed to compare papers that need work and those that do not. Even so, the numbers of students in each group is so small that we cannot make statistically-valid inferences about these results. With that caveat, these very preliminary results do not point to particular issues with the work of Chinese students when compared to other international students.

Writing Assessment 2012-13
Table 9: International by Country^

		Other International (N = 12)	Chinese Citizen (N = 13)
Statement of Thesis	Needs Work	50.0%	26.9%
	Pass or Exemplary	50.0%	73.1%
Paragraph Development	Needs Work	58.3%	34.6%
	Pass or Exemplary	41.7%	65.4%
Complexity of Thought	Needs Work	41.7%	23.1%
	Pass or Exemplary	58.3%	76.9%
Command of English	Needs Work	41.7%	46.2%
	Pass or Exemplary	58.3%	53.8%
Overall Score	Needs Work	58.3%	38.5%
	Pass or Exemplary	41.7%	61.5%

^ Please use caution when interpreting results due to small cell sizes

Differences are not statistically significant

Reader Comments:

In addition to rating papers based on a rubric, participating readers were provided a space to write comments on the scoring sheets; the qualitative data below is derived from the reader comments.

Readers rated almost 59% of the papers as pass. Reader comments on passing papers reflected the definitions set forth by the rubric and indicated specific areas in need of work.

> *Generally strong paper- good, but unsurprising thesis. Strong research, but little sense of what her sources are within her text. Paragraphs sometimes sloppy, and organization cloudy.*

> *Overall well written and (certainly) solidly researched. Lacking an initial thesis statement .... The piece is coherent- holds together well. A stronger intro and conclusion would push the quality of work.*

Just under 30% of papers were rated as needs work. In addition to echoing the needs work rubric definition (See Appendix C), many comments specified what made the paper insufficient. Numerous comments were made about poor or missing thesis statements, which echoes the score data seen in Tables 1, 2, and 3.

> *What is her thesis? Good evidence, but not used toward any goal, rambling.*

> *Doesn't prove thesis, superficial. But shows reasonable organization and tries to use evidence.*

> *No real thesis, holes in logic, sloppy paragraph construction*

> *No flow or structure. Some good ideas but they are lost.*

> *The primary reason that this piece needs work is because it makes evident that the writer is a non-native speaker. Yet the effort to clearly articulate the writer's point is evident. She tries to be clear and the tone is sincere. Grammar and sentence structure are key.*

> *Wobbly organization; no sharp thesis and development.*

Readers scored over 12% of the papers as exemplary (see Appendix C for rubric definition of overall exemplary paper). In some cases, reader comments on exemplary papers elaborated on what made these papers stand out:

> *Very nice language- good use of images very thoughtful thesis….*

> *Overall an easy and informative read, well supported by evidence and interesting in its stance and development.*

In other cases, comments on even the exemplary papers indicated room for improvement:

> *I wavered between the high pass and exemplary. She is clear and concise and develops paragraphs well although paper doesn't really flow clearly all the time.*

Some paragraph transition problems and the occasional awkward sentence.

This paper did not firmly support a simple thesis and had minor disorganization within paragraphs.

Finally, because the readers were privy to what the assignment was for each paper, several commented on the paper within that context, or commented on the assignment itself, sometimes with reflections on how assignments were linked to the papers' weaknesses.

Assignment did not ask for an argument. Assignment is not very specific. [Paper is] exemplary for what the assignment asked for.

…It read beautifully despite some grammatically awkward constructions and was interesting and engaging in response to a very open-ended essay prompt.

This is very nicely written within the context of an undemanding assignment.

Very solid efforts to meet the expectations of the assignment.

Seemed to get constrained by the assignment.

Reader Feedback

After two half-day sessions of rating papers and completing negotiations, readers discussed their experience of the assessment process and explored implications for teaching, policy, and future assessments. In addition to providing feedback on the assessment experience and faculty development, the discussion touched on the relationship of writing instruction to the First Year Seminar program and Smith's structures for continued writing instruction.

Assessment Feedback and Faculty Development

Readers reflected on the value of the reading experience, noting particularly the value of reading others' assignments and the dialogue inherent in negotiation. Readers also noted an appreciation for the opportunity to spend time focusing on writing with a group of colleagues. Some of the Jacobson Center staff involved in reading noted the value of seeing work from students who are not taking English 118, noting a difference in the quality of writing between the two groups (which is expected, and seen in the assessment, see Table 4)[1]. These comments are consistent with faculty feedback in prior writing assessment efforts.

In the current assessment, students were asked to write a short reflection explaining why they chose a given paper and what it says about their strengths and weaknesses as a writer. Then readers were asked to rate whether they agreed or disagreed with the students' self-assessment. However, early in the scoring process it became clear that students' self-assessments often focused more on the reasons for choosing a given paper, and few had comments that were easy to agree or disagree with, and readers were eventually advised to disregard that part of the scoring sheet. During the final discussion, this problem was noted.

Even so, another reader reflected on the value of the self-assessments, saying how much students' views varied, both in terms of the extent to which the student's self-assessment was consistent with the reader's assessment, and also in the extent to which students wrote thoughtfully about their own writing. Discussion of revisions, Jacobson Center help, and reflection on the writing process was also seen as notable in student reflections.

Faculty also reflected on the assignments they reviewed. Participants commented on the range of assignments they saw, including brief ones that they found vague or confusing, and others that were excessively long and detailed. Participants agreed that the length of the assignment did not correlate with its quality in participants eyes, with the best assignments being in a variety of formats. Faculty commented throughout the reading session and during the discussion on the many ideas they gleaned from colleagues' assignments. Assignments flagged by faculty as particularly good were identified and will be posted on the FYS Moodle site and made available to ENG 118 instructors by the Jacobson Center.

---

[1] English 118 is a course designed to provide enhanced explicit teaching of writing mechanics compared to other Writing Intensive courses. Students deemed at-risk in writing based on SAT-Writing scores or other risk factors are given a recommendation to take 118 as their Writing Intensive course, leading to an expected difference between the two groups.

Writing Instruction and First Year Seminars

Readers had a lively discussion about the first-year seminar program and its role in writing instruction for entering students. Faculty and instructors discussed the structure of the first-year seminar program. A few participants expressed concern about First Year Seminars being loaded with goals and requirements, including writing, speaking in class, reading at a college level, field trips, etc. Other participants countered with perspectives on the value of First Year Seminars, seeing the course structure as well-suited to writing instruction, with less demand for curricular content than might be found in a more typical introductory disciplinary course. Participants who were familiar with implementing writing across the curriculum noted that these demands on courses in the regular curriculum made that model even more difficult to manage. One participant suggested a "WI2" course that would take place in the second semester of the first year or in the sophomore year, and would operate on a "writing across the curriculum" model, providing writing instruction within a disciplinary context. One participant suggested requiring the writing assessment of all WI instructors, noting its value in providing perspective on student writing.

Participants also engaged the question of whether students should be required to achieve a pass rating on the assessment in order to complete the WI requirement, and how to intervene with those that don't. Further discussion of a second WI course followed, along with suggestions for individualized follow-up through the Jacobson Center or advising.

Summary

First year writing assessment has many benefits as a means of assessing writing, mostly pedagogical. 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 this 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 four sub-scales, the results of the assessment would make a useful starting point for individual conversations with students about their writing.

Based on the 164 papers read for this study, it appears that Smith first-years are generally strongest in the areas of Complexity of Thought and Command of English. The weakest area is Statement of Thesis, with just over 11% of papers found to be exemplary in that area. 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.

The data confirm anecdotal impressions: the majority of our first-year students write well enough (58.5%); some of our students write exceptionally well (12.2%); too many of our students do not write well enough, even after considerable effort (29.3%). We now know that our WI could do more to help our first-year students learn to articulate a thesis and develop paragraphs. 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. 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.

Appendix A: Technical Findings

Analysis of Reliability:

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. These correlations do not take into account any negotiation process that was undertaken subsequently, a process that resolved all disagreements between readers. Table A1 shows that each of the correlations is statistically significant, although none are particularly strong.

Table A1: Score/Re-Score Correlations

Writing Assessment 2012-13
Table A1: Score/Re-Score Correlations

	Correlation Coefficient
Statement of Thesis	.220*
Paragraph Development	.279*
Complexity of Thought	.253*
Command of English	.287*
Overall Score	.204*

* Correlation is significant at the 0.05 level (2-tailed).

Stronger correlations would have been expected if the readers had received extensive training designed to "norm" their ratings to a high degree of reliability. In Smith's current writing assessment process, the 19 readers participated in a single norming exercise prior to reading student papers, and their scoring was not corrected individually to ensure that every reader made consistent judgments. These relatively weak correlations are not surprising given the amount of attention paid to correcting readers who deviated from the gold standard. 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 future writing assessment, additional training could be incorporated to increase inter-rater reliability, however, another highly-normed assessment effort, the NSF's Critical Thinking Assessment Test (CAT) was not well-received by faculty, who felt the scoring process was too mechanized and did not have significant faculty development value. Aware of these concerns, the assessment's organizers decided to privilege the faculty development opportunities and leave the scoring process slightly open-ended, allowing the negotiation process to resolve outliers.

It will be important, as we move forward with further writing assessment projects, to consider to what extent we demand consistency and uniformity in our institutional assessment 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.

Appendix B: Recruitment Materials

Dear Colleagues:

As part of an institutional commitment to examining student quality in writing, Smith will be conducting annual writing assessment workshops with faculty using samples of student writing, selected from a cross-section of the entering class. The results of the study will inform our efforts in further refining the WI requirement, and in designing interventions beyond the WI course for students who need more help with writing after their first semester.

To minimize the administrative burden associated with WI courses, we have attempted to design a process for collecting writing samples that will involve minimal intervention on the part of the instructor. In November, I will send you a folder for each student, into which will be inserted a checklist. That checklist asks students to provide a copy of their best thesis-driven paper, a copy of the assignment for the paper, and a one-page reflection explaining why they chose that piece of writing. Students will submit the folder to you on a date you specify and the Office of Institutional Research will collect those folders from you.

Given that this is a new assessment process and we are requiring that every student participate, we would like to give students a clear understanding of how their papers will be used in the assessment process. Please insert the following statement in your syllabus:

Because writing well is a critical skill to develop, faculty members are always trying to improve the quality of the teaching of writing at Smith. We do this mainly by spending a few days reading and discussing the best work produced by a large number of first-year students in WI courses. Our goal is to identify which writing skills first-year Smith students are mastering and which aspects of writing our faculty need to emphasize in future years. Consequently, toward the end of this course, I will provide you with a folder in which you will include the following:

- *Your best thesis-driven paper (a paper that makes an argument) that is at least 750 words long.*

- *A copy of the assignment for the paper.*

- *A one-page reflection explaining why you chose this particular paper and what it says about your strengths and weaknesses as a writer.*

The completed folder is due [Instructors: Please specify due date]. Your submission will not be recorded on your transcript or in any permanent file at Smith, nor is the folder itself considered in assigning you a grade for this class. Please submit the completed folder even if you are not a first-year student.

The paper need not be one that you have already graded. Many students may choose to submit their final papers, which may not be due until the final day of the term. Depending on what makes sense for your class, the folders should typically be due either on the last day of class or the last day of the term.

Many thanks for your support. Please let me know if you have questions.

All the best,

Alice (FYS) / Julio (118)

Marilyn R. Schuster
Provost and Dean of the Faculty
Smith College
Northampton, Massachusetts 01063
T (413) 585-3000   F(413) 585-3070

RE: Writing Assessment Project

Dear Colleagues:

I am writing to enlist your assistance in the College's writing assessment project carried out by the Writing Committee (a sub-committee of CAP) and the Office of Institutional Research. It involves your commitment for two half days--the afternoon of January 17 and the morning of January 18. The stipend for participating is $200. At the moment, the assessment has three main goals: to help us identify how well students are writing after one writing-intensive course, to try to pinpoint where student writing is most and least successful, and to ascertain where we may need to focus more attention, both in designing new courses and teaching workshops and in sharing pedagogical strategies more generally. Even though it may seem difficult to imagine as you face the end of semester grading tasks, past participants have thoroughly enjoyed the project

The assessment project works as follows: First year students in FYSs and English 118 are being asked to turn in a copy of their best paper from that class this semester, along with a brief reflection indicating why they chose that paper and a copy of the assignment. A random sample of those papers will be culled from the submissions. A group of faculty will then meet to very quickly read and assess those papers, using a rubric and scoring sheet that asks faculty to eventually score the papers as adequate, needs work or exemplary. The project begins with a short "norming" exercise to help faculty determine benchmarks in each category. Some papers get read by two readers; when readers disagree on an assessment, they try to reach consensus.

This exercise was carried out in 2009 by a group of faculty, based upon writing assessments carried out at institutions similar to Smith.   Faculty who participated found it useful and enlightening to look at writing and writing assignments across the curriculum. They also enjoyed discussing with colleagues what the standards are for average, poor and exemplary writing at Smith. The assessment process provides a valuable opportunity to see what is happening across the campus with respect to writing.

I know that it is difficult to contemplate being engaged in a "grading" workshop, particularly at this time of the semester, but assessment is *not* grading as we normally think of it; these papers have already been graded. Though readers read at their own pace, assessment reading is quick reading and the overall assessment is holistic. If you are willing to participate, please contact Alice Hearst, Director of the First Year Seminar program (ahearst@smith.edu) or Julio Alves, Director of Jacobson Center (jalves@smith.edu).

Sincerely,

Marilyn R. Schuster
Provost & Dean of the Faculty

Appendix C: Assessment Rubric, Scoring, & Reading Procedure

Smith College
Writing Intensive Assessment Rubric (2013)

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. **Paragraph development**
Exemplary: Paragraphs are consistently well structured and developed. They are organized around a clearly stated, well-positioned central argument, which is thoroughly and convincingly supported with evidence. The evidence is fully explicated and analyzed. The sentences are coherent (they are clearly linked) and unified (they stay on topic).
Adequate: Paragraphs are well structured but manifest lapses in development. They exhibit all the essential elements—argument, evidence, explication, and analysis—but, often enough, these elements could be developed more fully and/or in a more orderly manner. The sentences may show minor lapses in coherence and unity.
Needs work: The paragraphs are poorly developed. They barely manifest the essential elements, or not at all. The sentences often show major lapses in coherence and unity.

3. **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.

4. **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.

5. **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.

Smith College
Writing Intensive Assessment Scoring Sheet (2013)

Case ID#:_____   Reader ID#:_____

___

Directions for portfolio readers: After reading the portfolio, follow three simple steps: 1. complete the grid below by checking the level of achievement for each category; 2. circle the large-font *Needs Work*, *Pass*, or *Exemplary* (the overall score); and 3. give the student some feedback about her work. Portfolios rated *Needs Work* and *Exemplary* will be read at least twice. Please refer to the rubric for definitions of the categories at each level of achievement.

Step 1.

	Needs Work			Pass			Exemplary		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Statement of thesis or main goal	○	○	○	○	○	○	○	○	○
Paragraph development	○	○	○	○	○	○	○	○	○
Complexity of thought	○	○	○	○	○	○	○	○	○
Command of written English	○	○	○	○	○	○	○	○	○

Step 2.   **Needs Work**   **Pass**   **Exemplary**

Step 3. Comments:_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Step 4. How strongly do you agree with the student's assessment of her own writing?

Strongly Agree   **Agree**   **Disagree**   **Strongly Disagree**