

TO: Larry Schuetz
FROM: The Assessment Advisory Committee
RE: World of Ideas Assessment (WOI) Report
DATE: February 25, 2000

During the spring term, 1999, work began on an assessment of outcomes for WOI. In consultation with Larry Schuetz and instructors for the course, it was decided that evaluating a sample of final papers would be of interest. Three faculty members, Mark Lencho, Emily Hipchen, and Michael Longrie, agreed to work on the project. Last spring, the rubric was developed and piloted. This past fall, a sample of 38 final papers were gathered and scored. The papers came from several sections of the class; instructors were asked to submit 4-5 papers taken at random from each section. In the balance of this report, the process used for scoring the papers and the results are described by Mark Lencho. The following summary information provides an overall picture of the results.

Three criteria—thinking, voice, and literacy—make-up the rubric. Each was scored on a six-point scale—6=outstanding, 5=strong, 4=adequate, 3=limited, 2=seriously flawed, 1=fundamentally deficient. The three raters read each paper and scored each criterion on the six-point scale. The results are summarized as follows:

Averages for thinking--	Rater#1=3.87	
	Rater#2=3.36	Overall average=3.45
	Rater#3=3.13	

Averages for voice--	Rater#1=3.97	
	Rater#2=3.47	Overall average=3.54
	Rater#3=3.18	

Averages for literacy--	Rater#1=3.95	
	Rater#2=3.16	Overall average=3.31
	Rater#3=2.82	

Based on the definitions supplied with the rubric, the scores for all criteria fall between the “limited” and “adequate” descriptors, and there seems to be little difference across criteria. It seems fair to describe the overall performance as average—the center of the continuum.

Coefficient alpha, which estimates the reliability of the scores from the three raters, is .78. For this kind of activity, this would be considered relatively strong agreement among the raters.

As this information is considered by the WOI instructors, the members of the Assessment Advisory Committee hope that there will be an opportunity to discuss the results—especially the appropriateness of the rubric. We believe the data offer a rare glimpse of how our primarily junior-level students are doing. Are they meeting our expectations? At its best, assessment information can focus our attention on specific program elements that might be improved. Other times, the data confirm our collective effectiveness. Ultimately, the goal is to offer our students the best preparation possible; hopefully this report leads us in that direction. In the future, we hope to expand this effort to include an independent evaluation of writing very early in our students’ experience—perhaps in the University Life class—which might help us gauge

improvement. We also plan to continue this assessment in WOI—perhaps on a yearly basis.

MEMO

TO: Steve Friedman

FROM: Mark Lencho

Chair, Ad Hoc Assessment Group for World of Ideas

RE: General Education Evaluation (World of Ideas papers, fall 1999)

DATE: February 3, 2000

cc: Michael Longrie, Emily Hipchen

On January 25, 2000, Professors Hipchen, Longrie and I met to discuss our reading of a set of papers from the fall 1999 semester World of Ideas classes. At the end of the fall 1999 semester, Larry Schuetz had organized the collection of essays from a cross-section of World of Ideas courses, which ended up providing us with 38 papers to read and evaluate. Larry had excised all names, and papers were furnished to us bare, without accompanying assignment.

After numbering the papers and having Sally Lange duplicate them, I had distributed them to Mike and Emily, along with rubrics for their evaluation. The three of us had spent approximately two weeks reading through the papers, scoring each of them on three primary traits-- “thinking,” “voice,” and “literacy”--and providing supporting commentary as necessary. As in the previous reading, we continued to score papers for overall effect as well.

During our hour and a half deliberations, the three of us expressed a range of observations regarding the scoring process, the efficacy of the rubric, the quality of the student writing and thinking, and the institutional context used for generating the writing.

Although the summary of reader scores shows some cross-reader variance, at the same time it is clear that each reader followed a somewhat predictable pattern in regard to his or her own scores. Importantly, best and worst papers were scored consistently across readers. Scoring discrepancies on less distinctive papers resulted not so much in disagreements about the quality of the papers, but in different conceptions of the rubric descriptors and the numbers used to express our evaluation. Does “thinking” include the ability to summarize and repeat, or do such abilities reflect attentiveness and engagement and therefore fall under the description “voice”? Does “literacy” mean solely the ability to write mechanically clean prose or does it also include other skills and knowledge such as awareness of the literary and cultural context in which a writing is set? Are the primary traits meant to be equally weighted, or should we privilege “thinking” over “voice,” “voice,” over “literacy” in reaching our overall evaluation? To what degree are the primary traits discrete markers or are they necessarily integrated? Though we had spent considerable time during our pilot session discussing these matters, it became clear after this last reading that further clarifications were still in order. Finally, we discovered we were still wrestling with questions concerning whether our six point grading scale

was meant to reflect our “A”-“F” grading scale or some independent evaluation metric which could conceivably be quite at variance with grading practices.

For the first time this reading, papers were scored on each of the primary traits listed in the rubric. Having three different scoring dimensions simultaneously at play seemed to encourage multiple readings and some analysis, complementing the holistic nature of the reading that yielded our summary evaluations of each paper. At the same time this feature-by-feature breakdown made us focus our attention on the primary traits to a much greater degree than we had during the pilot reading, when the primary traits merely served as prompts to help direct a single, overall score. If we are to do this exercise in the future, I have every expectation that we would share a much clearer and more uniform conception of each of our scoring categories.

Despite our various divergences, the assessment group reached a number of shared conclusions about the student writing and thinking we saw and about the success of the testing instrument for fostering a situation where students could showcase their efforts to best advantage. On the positive side, it was clear that students, universally, could write expository prose on task that displayed some competence. Additionally, they all showed a reasonable capacity to summarize texts and explain them in terms of what they had been told. On the negative side, as adept as the students were in summarizing information directly provided to them, they generally seemed incapable of constructing a sustained argument marked by critical examination and independent thought. Additionally, students were nearly universally unable to address historically removed texts with a historically informed perspective. Whether it was Huxley, Levi, Plato or Freud, students almost universally assumed these authors were writing in today’s context, addressing their won needs because of the problems that result in their own society. This foreshortened view of history and disconnection with the cultural arc in which our literature is situated led to misunderstandings and weak, self-defeating approaches to the literature. For the majority of students who felt that Huxley is giving his earnest impression of the ideal community in *Brave New World*, the outcome of their exposure to his writing is simply a disdain for his “silly” conceptions. The result could actually be to foster even less appreciation for the classics than what the student had prior to exposure. Perhaps an assignment cycle requiring students to submit papers, get feedback from instructors, and re-submit would help facilitate papers which would show a more constructive engagement with the course readings. However much we felt that a process of composition which builds in a feedback mechanism from informed readers could substantially *improve* the general quality of these papers, we did not feel that the range of different writing assignments created any corresponding *inconsistency* in the quality of the papers.

Finally, all readers of these World of Ideas papers felt like it would be valuable for instructors of World of Ideas to be exposed to our assessment rubric. We would appreciate the chance to explain the method of evaluation and the nature of the results we reached. We would also like to learn more about the nature and variety of approaches used to generate the papers we are evaluating. Ultimately, we would appreciate feedback on the assessment exercise from World of Ideas teachers. After all, it makes

sense that those responsible for generating the student work understand and value the result of its assessment.

Rubric Development for World of Ideas, Spring, 1999

Paper Number :

Rank paper on each of the “primary traits” defined below, using a 6-point scale where points are assigned in the following fashion:

- 6: outstanding
- 5: strong
- 4: adequate
- 3: limited
- 2: seriously flawed
- 1: fundamentally deficient

Thinking:

The writing reflects *critical*, *associative*, and *concessive* thinking, where “critical thinking” means a type of evaluative thinking which is able to uncover implication and context of discourse; where “associative” thinking has to do with the ability to make connections and applications; and where “concessive” thinking has to do with the sensitivity to scope, options, and uncertainty.

Voice:

The writing is successful enough in tone, creativity, energy and structure to make the reader pay attention, to be instructed as well as delighted.

Literacy:

The writer has a grasp of writing conventions. Content is informed by knowledge of scholarship and literary benchmarks.

Notes:

Revised Criteria After Scoring Papers, Fall, 1999

Rank paper on each of the “primary traits” defined below, using a 6-point scale where points are assigned in the following fashion:

6: outstanding	3: limited
5: strong	2: seriously flawed
4: adequate	1: fundamentally deficient

Thinking:

Strong thinkers engage in *critical*, *associative*, and *concessive* thinking, where “critical thinking” means a type of evaluative process which is able to uncover implication and context of discourse; where “associative thinking” has to do with the ability to make connections and applications; and where “concessive thinking” has to do with the sensitivity to scope, options, and uncertainty. They demonstrate that they can think freely, and rise above mere conditioned response. They can uncover, shape, extend, and delimit knowledge. Weak thinkers can at best give back what they have been given. Their thinking is limited to various forms of repeating what they have been told or making affectively based assertions. They give the impression of shaped rather than shaping thinkers.

Voice:

Writers with a strong voice are able to make an impact on the reader and communicate with some resonance. They may come across as especially creative or imaginative. Perhaps they take risks or write with evident energy. They show ownership of what they say, and they seem to be responding to a felt need; it is clear why they are writing and where they are coming from. They get their point across, and, at the very least, they can summarize attentively; they are coherent. Writers with inadequate voice give the impression of being reluctant writers; their efforts are muddled, mute, obscure, or contradictory. Perhaps the paper seems canned or pre-fabricated. They go through the motions but show no commitment. They are unable to compel attention.

Literacy:

The literate writer has a grasp of writing conventions. Content is informed by knowledge of scholarship and literary benchmarks. Literate writers are more than just functionally literate. They can make allusions and convey an awareness of our intellectual traditions. They use language and advance ideas which are culturally informed. They are knowledgeable and they can write appropriate to the level and interests of a serious and scholarly audience. Writers who are inadequately literate may have little understanding of the diction, grammar, punctuation, and usage conventions of formal written English.

They also may appear to have little understanding or knowledge of other literature. Their entire awareness may be circumscribed by popular culture and the here-and-now.

Scores for Each Rater Across Criteria, Fall, 1999

Scores: Reader one/Reader two/Reader three

Paper	Thinking	Voice	Literacy	Overall
1	3-4/2/3	3/2/1	3/2/2	3-4/2/2
2	3/4-3/5	3/4-3/2	3/4/4	3-4/4/4
3	4/2/3	4+/3-2/4	4-/2/1	4/3/3
4	5-/5-4/3	4+/5-4/2	4/5-4/4	4-5/5-4/3
5	5+/5-4/3	4-5/5/5	5/5-4/4	5+/5-4/4
6 (wide spread)*	5/3-2/4	5+/2/3	5/2/2	5/2/3
7	4-/4-3/2	3/4/3	4-/3-2/1	4/4-3/2
8	4-/4/3	3/4-3/4	4/4/4	4/4/4
9	5/3/3	4-5/4-3/3	5/4-3/3	5/4-3/3
10	5+/4-5/6	5/5/4	5/5/5	5+/5/5
11	4-/3-2/2	3-4/3/2	3/3-2/2	4-/3-2/2
12 (wide spread)*	5-6/5-4/2	5-6/5-4/3	5+/4/2	5-6/5-4/2
13	4/3/3	4+/3/2	5/4-3/3	4+/3/3
14	3/4/1	4/4-5/3	4/4/2	4/4/2
15	4-5/3/1	4+/3/4	5/3-2/3	4-5/3/3
16	4-5/5-4/3	5/5-4/3	4-5/4/4	4-5/5-4/3
17 (best)*	5/6-5/5	4-5/6-5/5	5/6/4	5/6-,6-5/5
18	6-/3/4	6/4/5	5+/4-3/3	6-/4-3/4
19	4-5/3/4	4+/3-2/5	4/3/3	4+/3/4
20	2-3/2/3	3/3/1	4/3-2/3	3/3-2/2
21	2/3-2/4	4/3/5	4/3/5	3/3/5
22	3/2/1	3/3/2	3/3/2	3/3-2/2
23	2/2/4	3/2/3	3/1/1	2-3/2/3
24	--/3/3	4+/4-3/4	5/4-3/3	5/4-3/3
25	3+/2??/1	3+/2-1/1	3/2/1	3-4/2/1
26 (worst)*	2/2??/2	3/1/2	2/1/1	2/1/2
27	4-5/4+/3	4+/4/3	3+/2-3/3	4-5/4/3
28	4-5/3/5	5-/4-3/4	4+/3-2/3	4-5/3/4
29	4-5/5+/3	4/5/4	3/2/3	4/5-4/3
30	4-/3/4	4+/3/4	4/3/3	4-/3/4
31	3/5/4	4/5-4/5	4/5-4/5	3+/5-4/5
32	5/6-5/5	5+,5-6/6-5/4	5/6-5/4	5+/6-5/4
33	3-4/3/3	4-/3-2/3	4/3/3	3-4/3/3
34	3/4/3	3/4/3	3/2/2	3/4-3/3
35	5/3-2/3	5/3/3	4+/4-3/3	5/3/3
36	4+/2/2	4+/3-2/2	4-/3-2/2	4-/3-2/2
37	2/4+/3	3/4/3	3-4/3-2/2	2/4/3
38	2/4/3	3/4-3/2	2/3-2/2	2/4-3/2

Scoring summary:

	X Highest Scorer	X Lowest Scorer
One:	22	3
Two:	4	6
Three:	3	16

***The texts of these four papers—#s 26 (worst), 17 (best), 12 (wide spread), and 6 (wide spread)—are included in this report**

Comments for Papers 26, 17, 12, and 6

26. #1: Is this person in college? (L) deficient structures, use of language incoherent in places; (N) unclear opening, no focus; undefined info; undigested info.

#2 Is this a non-native English-speaker??

#3: (T) Plato and Huxley surveyed and found wanting. Concludes that contemporary U.S. society is ideal--a result that shows a glaring lack of *critical* insight. Conclusion is asserted, not argued for. (V) Summative and assertive stances alternate throughout; degree of attention saves paper from the lowest score. (L) Severe language deficit--non-native English speaker. (N) "self-efficient" (self-sufficient); "ideal" (idea); "prospective" (perspective).

17. #1: (V) sometimes difficult but not inane. (N) understands implications/ explains major point/s and differences but lacks *specifics*.

#2: (N) *Very* sophisticated writer. Has good control of Huxley--but as good of Confucius?

#3: (T) Things like "stability," and "community" are shown to be *double edged*, conveying an appreciation of *complexity*. (V) Details that convey writer's absorption in the reading have a residual effect of compelling interest. (L) Confucius and Huxley are compared without reference to historical context or author intent.

12. #1: (V) sophisticated, controlled, effective; (N) clear opening; sees inferences: writer in control of argument, language.

#2: ----

#3: (T) Attempts to take issue and make associations both founder. (V) Embraces the perspective of the newly informed innocent, picking and choosing among agreeable sentiments in Plato and Jesus' Sermon on the Mount. Modesty seems genuine, appropriate for a thinker and for persuasiveness. Writer can engage by summarizing, showing material can register. (L) Coherent, but what else? Plato regularly is unappreciated because there is no recognition of the different historical context in which he writes.

6. #1: (N) explains adequately; shows thinking about material; remains on task; has opinion

#2: ----

#3: (T) Strongly able to see the inherent shortcomings in the effort of building an ideal community and in his own efforts. (V) Impacts though its modesty and recognition of complexity. (L) unable to read Plato and Huxley in context; self-created ideal community seems completely detached from any intelligible intellectual context.