

Summary of Responses

From the general discussion:

- Students are unable to synthesize knowledge; they can regurgitate but they can't do anything with it.
- Problem with this scoring process is that departments, schools, and university do not repeat the process.
- Outcomes and pedagogy are not the same thing. Need to focus on both.

Red – Questions

Blue – My reaction to/reflection about the comments

GE Scoring for IB and IC June 20-22, 2003 Scorer Feedback

Now that you've participated in scoring essays from IB and IC courses, we'd like to know what you think about the process. We'd also like to gather suggestions you have—for faculty teaching IB/IC and for future scoring sessions. Please help us by answering the questions below.

1. What is your overall evaluation of the process we used for scoring? Please circle the number below that best represents your overall view of the process.

1	2	3	4	5
It did not go well		It was OK but needs improvement		It went smoothly as is
3 – 7				
4 – 6				
5 – 3				
Average: 3.4				

What went well?

- Well organized; task clear; rubrics mostly clear.
- Enough scorers; scorers took responsibilities seriously.
- Scoring and consultation with other faculty; learning about criteria for other disciplines is important if there is to be an integrative approach. Also opportunity to see assignments from other disciplines.
- Getting it done together as a group.

What could be improved?

- Scoring rubrics need some fine-tuning; may not fit all disciplines equally well.
- Assignments need to be tailored to the rubrics. Instruction (writing prompts) to students could be improved; would make them easier to score and scoring more accurate.
- More time for practice scoring and discussion, improving inter-rater reliability; also include bias training so scorers aren't influenced by student's opinion or choice of topic. If we were making decisions about individual students these would be more important issues. Since we are not, I think that they're probably not necessary. In hindsight, since they finished so quickly, I probably would have had the scorers spend more time doing this, and possibly double score papers as well. However, I really anticipated the process would take MUCH longer than it did so I tried to move the agenda after a couple of hours. I also think a number of those scoring really did not know much about the structure of GE or the move to outcome-based assessments so this added to their confusion. I really did explain it but for some it went in one ear and out the other.
- Have pre-established anchor papers.

This would be desirable. I tried to identify some but decided I couldn't make the judgment for all areas. We need an "expert" group to do this. We could use papers from this round for writing and critical thinking, choosing some that scored at each level to see if we could get agreement. For other integrative areas, however, we'd need content-specific papers. Having anchor papers could also inform faculty about expectations and they could be used as a focus for discussion at the department and/or school level. All this would only work if faculty really bought into the process and wanted to improve student performance.

- Needed a larger pool of IC papers. Yep!
- Felt there was a rush to read lots of papers in a day rather than use all the time allotted. Up until mid-afternoon, I still thought we would use another day. However, when faculty realized that they could finish it up, almost all wanted to try and get it done. I went along with their wishes. I'm not convinced anything would have been gained by trying to slow down the process at that point. At the end, some volunteered to read papers from a different area (writing instead for critical thinking, for example). Looking at the scores, they didn't seem different from those given by the "experts" who had participated in the faculty scoring discussions for the rubric.

2. Was this scoring process a productive one? Please circle the number below that best represents your overall view of the scoring process we used.

1	2	3	4	5
It was pretty much a waste of time and money		It was somewhat productive but needs improvement		It was a productive process likely to yield useful information
1 – 2				
3 – 4				
4 – 7				
5 – 4				
Average: 3.6				

The “ones” focused on the fact that there’s a disconnect between the process and the classroom; may not have had an impact on the classroom and the lack of congruity between the assignments and the rubrics.

How could the process be made more productive?

- Integrate into the classes themselves. There is such a disconnection that I’m not sure what if anything will be accomplished.
- The assignment did not necessarily correlate well with the rubrics’ goals. I am loath to suggest that teachers be made to give proscribed assignments, [but] if you are hell-bent on doing this I think you need to get better matches between assignments and rubrics (or rather, outcomes [and] goals.)
- Improve the prompts. Demand participation. Require chairs to review GE syllabi for compliance. Remove/restrict section numbers for those departments that don’t comply.
- Make sure all courses are represented; standardization of use of rubrics; more uniform participation of courses.
- Rubric discussions.
Not clear whether this is among scorers or faculty more broadly. This individual thinks the rubrics need some revision.
- Add an “explanation of score” column to rubrics so scorers could explain their reasoning. Who would read it? How would it affect the process or the results?
- 1-4 scales are good for forcing a “tight” response, but did not have enough flexibility for students who were rated as not developing but not competent. [Suggested having a 2.5.]
- Scorers were told they could use .5’s if they wished and I recorded them that way. Some wanted to use finer gradations (2.3 for example) but I rounded up or down to .5. I don’t think the process is sufficiently precise to warrant this type of precision!
- More IC classes need to submit papers. Review the IC scoring rubric to include language skills.

- Break down the rubrics to specifics. Some rubrics don't reflect certain subject areas. For example: critical thinking and the papers on science subjects. Since most of the writing requirement has certain percentage for language mechanics, include language assessment in the rubrics.
Wanted a column added to IB and IC rubrics on language mechanics. Seems to me that the faculty can use this as grading criteria but that for our purposes, it should just be in writing.
- If most of the assignments are 5-10 page papers, the rubric should feature fewer (more global) requirements.
- Since I didn't understand what the ultimate goal was, it was very hard for me to tell whether the goal had been accomplished.
- Where does this info go? How does it impact actual classroom pedagogy?
- Rules for obvious plagiarism; rules for papers that don't "fit" the paradigm.
- Some raters move ore quickly than others. Perhaps setting up tables with piles of papers to grade in the center and grabbing as you go would have been better.
- For improving the number of papers received: Ask department chairs to select participating faculty on a rotating basis, and have the deans' secretaries call the faculty to ensure enough samples are received.

Is it worth doing?

- Depending on the purpose. If it's an academic exercise in futility then continue on this path. If you really want to assist students and faculty then it needs to be forced into the classroom.
It's not clear to me what this means so I included it in its entirety.
- No.
- Ideally, in my opinion, the study should be done by faculty who teach courses in IB and IC. Accepted faculty in the [scoring] workshop should be ones who have taught courses in the areas so they can use the feedback.
- Yes. The small sample tells us what is going wrong; we don't need ten more years of study to reaffirm what we already know.
- Yes. If you get full participation.
- Maybe. But that depends on how the info will be used. What will be done with it? It was useful for me as a teacher to see how different writing assignments worked—and didn't work.
- Absolutely! I am eager to see the results in order to guide improvement to IB courses in my department.
- Absolutely!
- Yes. Finding a rubric to best represent the intentions of the university is a slow, painful process. However, I believe that it is important.
 - Yes.
 - Yes. I agree that it was very worthwhile to read papers based on a variety of assignments and academic disciplines.
 - This question goes hand in hand with the question of whether university-wide GE assessment is worth doing. The answer to the second question is "yes." But whether this process was precise enough or relevant enough to contribute to and produce reliable data about actual student skills and knowledge is unclear to me. Clearly multiple data sets are needed to reveal a trend. But the trends are only as accurate as the data.
 - I think the information was useful and interesting. However I have my doubts that the faculty will make curricular changes appropriate to meeting the IB goals.
 - Yes.
 - Probably. Allow independent overview. Should be some feedback to instructors collectively and individually.
 - Yes! Important for providing exemplars and samples to teaching staff and students!
 - I believe that examining these documents serves two very important purposes: (1) the process allows the university-at-large the opportunity to evaluate the state of student writing and critical thinking. (2) provides data to properly "prompt" faculty to use rubrics and enforce higher writing standards across the board.

3. Think about the characteristics of the assignments that were used for the essays you read.

What did you notice about the assignments that worked well?

- Clear instructions.
- They were directly related to the rubrics.
- Clear about expecting analysis, arguing a position.
- Clear and well-define. Assignment had a well-defined goal and information on style, sources, and expectations.
- The good ones were more general, encouraged multiple hypothesis, evaluation; not just rote fact-reporting.
- Very clear instructions are given to the students; the faculty indicated availability to coach the students during the process.
- Having sufficient information about what the student's assigned instructions were.
- The assignments that worked well had the following in common: papers corresponded to the assignment; the assignment was clear and specific.
- Clear instructions about the requirements.
- From a critical thinking perspective, assignments that emphasize an understanding of more than one perspective on an issue, and then call for students to use their critical thinking skills to defend and critique diverse perspectives, are better than reports.
- They were lengthy, very clear, required multiple activities in order to accomplish (listen, write), (observe, write).
- Good papers seemed to occur at random. I saw no obvious correlation with the syllabus.
- They were situated, meaning they had a fairly well articulated rhetorical framework. They called for student decision-making without a well-defined context. The prompts were fairly explicit about purpose.
- Comprehensive instruction; awareness of desired outcomes; broadly based subjects.
- Professors with highly specific grading rubrics on the paper were an asset for raters and students.
- Professor that provided prompts seemed to receive a greater adherence to the overall objective of the assignment; clear and concise directions yield the best papers.

What did you notice about the assignments that did not work well?

- Students had too many choices and format did not always follow form.
- They were probably well integrated into the course itself and so we were very aware that, lacking that context, we couldn't genuinely score them by the abstract criteria of the outcomes goals. They were probably pretty good assignments, however.
I have a little difficulty understanding the reasoning here. If the papers did not demonstrate the GE outcomes then it's because it was really a good assignment? Maybe that's why critical thinking scores are so low??
- Describe and "compare and contrast" prompts, without additional articulated expectations to analyze, led to "book reports."
- They were open-ended. Expectations were not clear.
- Their assignments asked for "book report" responses; very, very little analysis or critical thinking. Students were clearly in Sgt. Friday mode—just the facts...
- Unclear instructions.
- In a number of cases the written assignments were too vague. I would assume that the students received more concrete information and clarification in class (verbal) that we were not given. However, this did provide insight in that, maybe, our instructors should convey their instructions clearly, in a written format.
Yep!
- The assignments that did not work well had the following in common: papers were poorly written; papers did not correspond to the assignment; assignment was not specific.
Seems like the poor papers emanated from poor assignments, not the other way around.
- Break down the instructions to specifics and give some illustrations of a good paper.
- Instructions that were not very clear, or very vague.
- Some assignments did not emphasize critical thinking skills or did not even require them. A term paper or report which primarily asks student to provide factual information does not require the use of many key critical thinking skills, such as reasoning, avoiding fallacies, using credible sources to support their claims (bibliographic formats such as APA or Chicago pretty much presume all sources are equal—give readers no way to determine whether authors are credible and unbiased), recognizing both sides of an issue, etc.
- They did not incorporate area goals and objectives within the requirements. The less successful assignments involved only reading and writing. Some assignments made absolutely no attempt to incorporate GE goals.
- The only exception to the above was one course that had virtually no science in the course content or writing assignment.
- A lot! Too narrow, prescriptive, and literal. Closed out student reflection or application, and assessment of knowledge. Called for too much summary. Too form-driven. Non-specific, contradictory instruction.

- Just asking students to write a report following just a general rubric format is not as valuable.
 - Vague instructions.
 - Some professors did not adequately prepare the students (on paper). Students should not only be instructed about the appropriate components of an assignment, but also the inappropriate elements.
- Good point. Something I don't think most of us think much about..."non-examples" of what is wanted.

What recommendations do you have for IB/IC faculty regarding the design of assignments for this (or a similar) scoring process?

- Integrate the rubrics into grading schema of the class.
- Slavishly give the administration the numerically quantifiable definition of learning they seem to seek? Teach to the test?
- Faculty should be aware of the rubrics themselves. Faculty should make sure their assignments will meet the GE goals.
- Students must be guided out of the fact regurgitation mode if the IB outcomes are to be achieved.
- See the scoring rubric first and become very familiar with it.
- Faculty need to provide clear instruction to the students. It is difficult to assess the written products if the assignment instructions are vague.
- Offer a course syllabus with elaborate, clear, specific, and concise assignments. Offer examples of an “A” paper.
- If we are to evaluate student on critical thinking, the assignments must require students to do the types of critical thinking that are emphasized on the rubric.
- Many IB/IC faculty need to go to a friendly seminar about paper assignment construction, which will enable them to incorporate GE goals comfortably. Often, people not in their discipline can be the best guides to this.
- The process showed quite clearly that nobody in the IB courses teaches science. They cover the knowledge generated by science.
- Need classes for professors. Need a WAC person on this campus.
- We need to provide students a way for formulating questions on which to direct discussion of a topic. Comparative papers would work the best.
- Establish objectives that are consistent with outcomes established for rubric. Tailor assignments consistent with desired outcomes.
- I wish I could take a few of the samples! Scorers could mark the excellent examples of assignment design and instructions and templates could be created for mass distribution to faculty (like the excellent sample syllabi).

4. Some faculty have suggested that we develop and use a single writing prompt to be given to all students (or to all students within an area). What is your reaction to this suggestion?

- Not a good idea unless the prompt was not directly related to their grade in the class. Disciplines are not homogenous which would make a single writing prompt difficult to develop.
- This is probably no more objective or gradeable, but it would be more uniform.
- Difficult given diverse courses unless you can get a real global prompt.
- No! However, examples of clear and successful writing prompts would be very useful to faculty. A point person trained in writing prompts might be made available to faculty to help edit their prompts. Alternatively, a standard prompt with certain elements (similar to the syllabus template) might be developed.
The idea of a prompt template is an intriguing one, I think. It might provide a sort of happy medium between a single, required prompt and chaos.
- It is worth a try. Before this grading experience, I would not have supported this idea, but now I like the idea.
- I don't agree because of the variability in topics and interest of each class.
- I would be hesitant to support this; each area (field) may have written communication and patterns of thought that are unique. While we attempt to assess the written product, there should still be flexibility in the presentation/structure of their work.
- Try it.
- That's a good idea.
- I don't like that. Individual instructors should have flexibility to tailor assignment to their pedagogy and learning objectives. However, the instructors need to be aware of what criteria will be used to test the students' writing, critical thinking, etc. They should then tailor the details of the assignment to enable the students to demonstrate the skills we are evaluating. A single prompt also facilitates diverse forms of academic dishonesty.
The plagiarism thing is something I hadn't thought of. I wonder if it would inspire increased dishonesty? Don't think it would be the same problem with a template.
- Certainly such a prompt should be given to all faculty. Then with a friendly seminar (as described above), I predict faculty would be willing to incorporate area goals and find the results rewarding.
- No strong feeling one way or the other.
- Not a good idea. Better to teach teachers how to use writing in their classes. Much can be done.
- I agree! [That there should be a common prompt.]
- I am not sure a single prompt is appropriate—but several examples of excellent work would be appropriate—especially by discipline.

5. If there's anything else you'd like to suggest, please do so below.

- Thanks.
- The writing rubric was generally clear. However, I would like to see someplace where plagiarism is addressed. Several papers were perfectly written (warranting a 4) but were almost certainly plagiarized. Thus, the scoring will not tell us very much in these cases.
- I would suggest doing a pilot testing phase in which faculty use the rubric in classes to determine gaps in the process during a real application.
- Include curriculum faculty (School of Education) as scorers.
Not my suggestion!
- See that students who are taking GE courses understand the goals/expectations of these courses. Some sort of a workshop or orientation to incoming students will certainly help.
Does this happen in University 1? Really seems like this is something the faculty teaching GE courses should take responsibility for.
- Departmental leadership is really needed in making individual faculty accountable to GE goals in a collegial and cooperative way. If even one senior faculty in a department would be available to others and especially to new faculty to provide and clarify the goals of their GE courses, then there would be some progress.
She goes on to provide lengthy specific suggestions for improving the IC rubric and for “educating” faculty, specifically lamenting the fact that many/most faculty still don't understand the difference between grading and assessment.
- This process is vital and needs to be replicated at many levels and they have scant (?) impact on the classroom. Need a university journal that responds to faculty ideas and recommendations about teaching.
This was hard to read but I think that's the gist of what it said.
- Work in a quieter location—how about library where students are expected to keep voices quiet!
Seemed really quiet most of the time.
- It may be productive to randomly select courses or students and evaluate whether they feel these outcomes have been met. Encouraging excellence in faculty is a tough job—many of us deeply care about the future of our students, and sincerely appreciate the process of evaluation. If one professor each year is encouraged, it is worth it!!
I sure wish it could go a little faster than that!