Challenging the Task

Liz Hamp-Lyons
The University of Michigan

Dan Horowitz was well-known for his research into essay examination prompts, and was greatly respected for the intellectual clarity of his work and for his humanistic grounding of that work in the central questions facing practitioners in ESL classrooms in colleges and universities. In this paper I review some of the work that has been done on prompt effects in ESL writing at the college or college preparatory level, focusing on just one small aspect in an attempt to move our work in this area toward a better general understanding. While I do not make explicit reference to Dan's work in the text, the collegial dialogue we maintained is an important underpinning of the paper.

There has been a great deal of research into the question of whether topic types, topics, the linguistic structure of questions, and an array of what we may group together as "prompt effects" have a significant effect on the measured writing ability of native writers of English. While far less work has been done on those same issues in relation to the writing of nonnative English users, it seems likely that the effects of prompts, if such exist, will only be exacerbated when we look at nonnative writers rather than native writers.

An overview of the field (Hamp-Lyons, in press) suggests that in first language writing assessment the trend is to treat topics, and even topic types, as equivalent. In the major ESL/EFL writing assessment programs the same trend emerges (Hamp-Lyons, op. cit.). For example, the TOEFL Program uses two topic types at different administrations of its Test of Written English, although analyses of trial prompts of the two topic types in the TOEFL Test of Written English (TWE) (Carlson, Bridgeman, Camp, & Waanders, 1985) showed that they behaved rather differently (they correlated at around .70). Reid (1989) looked more closely than Carlson et al. at the four writing prompts in the TWE study, and found that the students' tests varied significantly from topic type to topic type, even when the differences did not result in differing
score patterns. The variation was most marked for strong writers. Perhaps weaker writers have less language flexibility, while those with higher scores seem able to adapt to different topics. In looking at the issue of whether there is a significant effect on student essay test writing from the prompts used, then, it would seem that the answer depends on one's research orientation and the questions one asks as much as it does on “hard” numbers.

I can illustrate this from two studies carried out in the University of Michigan Testing and Certification Division. Both studies looked at prompts on the MELAB (Michigan English Language Assessment Battery, a test for nonnative users of English filling the same function as the TOEFL, but including a composition as a basic component rather than as an occasional, optional, extra). In the first, Spaan (1989) describes an experimental study of two prompts from the MELAB, chosen because they appeared on the surface to be dramatically different. Spaan provides a linguistic analysis to show the linguistic, cognitive, and schematic differences of the prompts, and interprets her score data as suggesting that even these prompts in fact yield scores which are significantly related.

In the second study, Hamp-Lyons and Prochnow (1990), in a post-hoc study, looked at all sets of MELAB prompts used in the period 1986-89, including those Spaan had used. They found that expert judges and student writers felt able to recognize easy and difficult topics, and that their judgments of the relative ease/difficulty of prompts were generally confirmed by score levels on prompts. While they confirm Spaan’s assessment of her two topics as radically different in difficulty, considered a-contextually, by looking too at a general language proficiency measure they are able to suggest some reasons for essay scores that are less different than predicted. It seems that nonnative writers taking the MELAB writing test component assess which prompt is easier, and that students with weaker language proficiency choose the easier prompt while students with stronger language proficiency choose the harder prompt. Hypothesizing too that reader accommodation plays its part in pushing disparate prompts toward parity of treatment, they suggest that both weaker and stronger writers regress toward the mean in their writing score, relative to their scores on other language components. These two research studies
from the same testing program differ in many ways because of researchers’ orientations, but only the complexity and intangibility of the kind of data we are working with when we struggle to pin down prompt difficulty permit us to arrive at such differing conclusions.

**Challenging the Task**

The interaction between tasks and individual writers’ affective responses is so complex that it is unlikely we will ever untangle it. But to the best of my knowledge there has been no research on how a student’s perception of herself as a member of a disciplinary community affects her response to a required task in the peculiar context of a discipline-linked writing assessment. As a very small part of a much larger study of an international writing assessment taken by nonnative English users wishing to pursue graduate study in English, I looked at this question, not because I originally planned to, but because it intruded itself upon my awareness. In so doing, I think I have arrived at some interesting and useful generalizations about the problems we face in understanding what we do when we “prompt” someone to write on a formal test.

The assessment I studied, the British English Language Testing Service examination (ELTS), aims to be fairer to students than general tests providing a variety of tasks depending on the disciplinary area the test candidate is familiar with (for a full discussion see Ramp-Lyons, 1987). The assumption is that students, particularly at the graduate level, will write better when they can write about what they know than when they must write about subjects unrelated to their studies. To investigate the claims for the test, I looked at whether test candidates got higher scores on tasks targeted on their areas of disciplinary knowledge, broadly defined, than on general writing tasks. My quantitative data yielded ambiguous results (Hamp-Lyons, 1986), and as I attempted to interpret them, I was led to look more closely at, among other things, what the test candidates had actually written.

A writing test is a discourse exchange, although a discontinuous one, and this particular realization of the discourse exchange has the expected sequence:
CHALLENGING THE TASK

**Figure 1**
General writing task

<table>
<thead>
<tr>
<th>INITIATION</th>
<th>RESPONSE</th>
<th>FOLLOW-UP</th>
</tr>
</thead>
<tbody>
<tr>
<td>= the essay question</td>
<td>= the student’s essay</td>
<td>= the essay scoring</td>
</tr>
</tbody>
</table>

The writing assessment which uses a discipline-related task is a three-way exchange:

**Figure 2**
Discipline-related task

<table>
<thead>
<tr>
<th>INITIATION</th>
<th>RESPONSE</th>
<th>FOLLOW-UP</th>
</tr>
</thead>
<tbody>
<tr>
<td>= the essay question test designer [A]</td>
<td>= the student’s essay student [B]</td>
<td>= the essay scoring trained essay scorer [C]</td>
</tr>
</tbody>
</table>

This participant structure contrasts with that for essays written for academic courses, where exchange structure is typically two-way, with the professor setting and responding to the assignment, or perhaps supervising teaching assistants’ responses:

**Figure 3**
Academic course task

<table>
<thead>
<tr>
<th>INITIATION</th>
<th>RESPONSE</th>
<th>FOLLOW-UP</th>
</tr>
</thead>
<tbody>
<tr>
<td>academic faculty [A]</td>
<td>student [B]</td>
<td>academic faculty (TA) [A]</td>
</tr>
</tbody>
</table>
When I was at the University of Edinburgh, I carried out a small scale study of the design parameters which faculty have in mind when they prepare written tests in their own discipline, and of the criteria they use when scoring the writing produced by the graduate students in their courses. Only one of the 24 faculty members I asked said writing was never used in the course. Seventy percent of the faculty used writing for most or all of their examination requirements at the graduate level.

Faculty cited "avoidance of ambiguity" as the most common design feature they used for their essay examination questions, and the next most frequent design factor had to do with finding out what students know. Here faculty talked about allowing students to show the ability to apply knowledge to problems in the discipline; they talked about getting students to show discrimination; they talked about designing tasks that required students to discuss or argue, not just display knowledge.

Certainly, clarity was central, as was the ability to take information, think about it, integrate it, and reshape it into knowledge. But then I turned back to my data, to the international writing assessment context, and studied the essays students had written. I began to believe I was seeing some qualitative explanations for the statistically ambiguous answer to my question of whether questions located in a student's areas of knowledge allow her to look like a better writer than general questions.

The first and more obvious of these explanations is that any test with only six discipline-related choices does not offer enough options to give every candidate a task in an appropriate area. For example, in the technology area the task required the student to write about the properties of metals appropriate for one specific engineering use. Clearly, there are technical fields where individuals do not work with metals in the raw state. In attempting to design tasks which would "work" for as many candidates as possible, the test designers had been forced to choose topics mastered rather early in the learners' development of disciplinary knowledge. Some older candidates had forgotten the broader basic content of their freshman and sophomore years. But as I looked at essays and found textual evidence for lack of content knowledge or loss of content knowledge, I saw something else that began to fascinate me. Some of the discourse exchanges were "challenges," a
term first used by Labov and Fanshel (1977) in describing a discourse function in sociological, psychological terms. Instead of supplying the predicted response, some test takers challenged the question itself.

The addition of the challenge to the discourse between faculty member and student significantly expands its potential:

![Figure 4]

In a discourse exchange on a writing assessment, the respondent to a question, as a participant in a discourse, has a choice. She may give the predicted, what I call the “unmarked,” response, or she may “challenge”—take issue with the validity of the question. The challenge is a “marked” response, what in some fields might be called an “abnormal” response, and here as in many other fields of the human sciences we can learn a good deal about normal behavior or performance from the abnormal.

Because many writers, sensible of the pragmatic function of the genre of the writing test before them, edit themselves out of their texts—or, and perhaps more commonly, with new arrivals into our academic community, edit a new self into the text, much of the interaction of the individual writer and the task that interests me
is not amenable to close investigation. Bartholomae (1985) looked at how new undergraduate writers use the written language to locate themselves within an academic discourse. Taking the position that to write at the university level involves establishing authority for oneself, Bartholomae saw some students taking an "outside" position—creating a "naive" discourse—where they establish only the authority to be present and the discourse says simply: I am here. This is who I am. At a transitional level, students mimic the sound and texture of academic prose without there being any recognizable interpretive academic activity under way. This is the attempt to claim the authority of membership. Many writers taking our ESL writing assessments will fall into this middle category, will be at the stage of taking on the protective coloration of their discipline, and thus it will be difficult for us to see their personal imprint upon the texts they create. But, says Bartholomae, more advanced students are those who "establish their authority as writers, who claim authority by placing themselves at once within and against a discourse community, working consciously to claim a territory of knowledge as their own" (158). Looking at the writing that such students do on writing assessments such as the ELTS, where disciplinary knowledge is deliberately introduced as a factor by the testing agency, may help us understand both test and writer better.

When we look at how nonnative graduate students write about discipline-related content, we are at the intersection of two kinds of knowledge: knowing how to write, and knowing a discipline. Many of the texts I studied appeared to be by students who were "transitional" both in language and content knowledge. Constrained by the awesome implications of the test they were taking, it is unsurprising if they generally focussed on "psyching out" the test, the task, and the reader's expectations. They wrote to camouflage, not to communicate.

Attempting to write with the voice of disciplinary membership, handicapped too by writing in a language other than their own, they did not offer the reader their voice, their sense of self, their authority. These, then, constitute what I am calling here my "unmarked" responses. Less competent writers, however, who had not yet learned how to distance themselves through rhetorical,
stylistic, and lexical shaping, were able only to establish the authority of presence (in Bartholomae's terms). When the interweaving of who they were and what they knew with the demands of the task became impossible for them, they were able only to protest or apologize.

1. I have no idea about “Green Revolution” but I think it is how to increase food supply.

2. My field of work is a computer science, so I'm afraid I can not exactly and rightly discuss the advantages and disadvantages of nuclear energy.

3. I have to recognize I don't exactly know how a factory is organized and then I cannot give a well opinion.

These are “marked” responses in my terms, but their lack of competence does not permit them the flavor of challenge. It is the other kind of “marked” responses, the papers where the student, more or less skilfully, but in my reading always authoritatively, takes a critical rather than an imitative stance in taking and shaping the task to herself, that I refer to as challenging the task. Challenging the task is not a new concept: John Swales (personal communication) recalls his days at Cambridge in the late 1950s, where the tradition even at the undergraduate level was to push against a language that would render one like everyone else, and where whenever the student was aware of a possible deficiency in content knowledge, the traditional strategy was “Don't answer the question, question the question.” But I suggest that by focussing in on all challenges we can learn more about tasks as they affect the wider population of student writers.

In oral discourse a challenge is an interruption in the normal flow of the social conventions of conversation, and similarly in written discourse a challenge is an interruption in the normal flow the social conventions of textualization, in the making of meaning between writer and reader. Challenges do not conform to the discourse expectations, and as Figure 4 indicates, while unmarked responses are unlikely to lead to a challenge, a marked response is considerably more likely to do so. In challenging the task, then, in pushing against the conventions, especially in a situation where as
much depends on the outcome as this one, the writer is taking a risk. As Swales (in press) has shown, writers at the center of their disciplines have a firm understanding of the strength and reasonableness of the claims they make, that is, they have self-confidence in their powers of judgment, in their ability to see a problem and say what needs to be done about it. But they have, in addition, a strong sense of their location in that discipline, of how far they can go in transcending the “rules” of their community. To “work self-consciously against the common code,” as Bartholomae has referred to it, is a difficult and dangerous activity. Not only is it subversive, it can be self-destructive.

Myers’s (1985) study of two biologists writing grant proposals for work in somewhat new fields shows how they needed to establish a claim to authority in that field, transferring only some of their authority from their usual field. The rhetorical demands on them as they interpreted their work for a disciplinary community that had not yet awarded them membership were far greater than usual. Myers describes the tentativeness with which they moved beyond the realm of their established authority, hedging their claims and supporting their right to make them with “face establishing” moves (Brown & Levinson, 1977). Taking this risk, one was successful and one was unsuccessful. When a student, even a graduate student, takes this kind of risk on a test which will affect her future, we must think about whether she is displaying authority or masking incompetence. When a graduate student who is a nonnative speaker does so, we must ask what has compelled her to push against the test.

I will look at just two examples of challenges in my test data. Both examples come from the set of responses to a prompt in the field of physics (given in the Appendix). As a close look at that prompt will show, writers are instructed to take a specific perspective (Hamp-Lyons, 1987), that is, to argue from a pre-specified position, and further, the position specified for them is one contrary to that actually held by many people. The data showed that more writers challenged on this prompt than on any other. I find in this point a strong indication of the interweaving of influences of writer and prompt, as I later find evidence for the still more complex interweaving of writer, prompt, and reader.
In my first example (Appendix, A), the writer establishes authority with "I am a scientist" but at the same time resists the implication that is set up by the question, that as a scientist he will inevitably take a perspective which supports the position stated. This is a very difficult rhetorical move, and he manages it successfully, it seems to me, with the strong opening phrase "Even though" used as an extratextual reference to the prompt. The writer picks up on the part of the prompt which was intended to provide the position for him to counter, and instead of countering he argues in its favor. From the first clause on, however, he makes no further reference to his authority as a scientist, using arguments and style that would just as likely come from any other writer. Having established "face" as a scientist, the writer does not insist on that role in the rest of his essay. Further, from that point on his answer is not marked as a challenge except that the position he argues opposes that set up for him.

In the second example (Appendix, B), the writer makes no apology. Her answer is an argument. She establishes her definition of what it means to be a scientist, then uses her definition to support her position that the question is not one to be asked of the scientist. When she says "For me, the question above does not make sense," she speaks from the role she is adopting for the task. Throughout her essay she maintains the role of scientist, as we can see by following the chain of lexis relating to the objective nature of scientific inquiry ("can be measured"); "know"; "facts"; "proper information") and she maintains her sense of distance from the issue ("only can develop"); "only can give"). We can note a stress on knowing and knowledge, juxtaposed with awareness of the limits of knowledge and the need for a human responsibility beyond absolute facts. These are coupled with a more subtle expression of the complexity of the issues: first she uses two contrasting adjectives to suggest the balancing of beneficial and harmful effects, and second, two contrasting passive verb forms to suggest the contradiction of enjoying and being harmed at the same time. She is claiming her authority, her privilege to speak, by placing herself, as Bartholomae would say, "within and against" a discourse (1985, p. 158), and in so doing she is appropriating the task, remaking it as her own. Through much of the text her voice is confident, her
indignation is righteous. The move from “Therefore” in paragraph two to “To me” in paragraph three suggests, however, an awareness that what was “proved” in the foregoing was not really proved, but left something to be desired. From my experience of students from a range of cultures and differing exposures to the academic conventions implicit in the test, I would suggest that her unease betrays a sense that her forthright approach to the topic is not quite what the decision makers are looking for here, that perhaps something more, or different, is expected of her, although she doesn’t know what.

In the examples, the first writer takes a small risk in taking up an opposing position, but perhaps he is more aware of the risk he takes than we would think. Perhaps he is aware both of the task and its unreasonableness, and of the judges and their powers and predilections. If he sees them as liberal intellectuals, he is not far off the mark. If he thinks that the unreasonableness of the task and the probability that the essay scorers will be English teachers are related, he is even closer. The second writer, on the other hand, displays no rhetorical sleight-of-hand. She deals with the issue from within the discipline, and expects the reader to meet her within the discipline. If we do not do so, can we say it is she, and not we, who is at fault?

I would like to argue that neither of these writers displays incompetence as a writer, that both know who they are and what they are doing. The incompetent writing that we looked at earlier was marked particularly by the appeal to the reader, by an inability to stay within the “rules” of this game, which require that the writer writes as though unaware that this text is to be evaluatively, not collegially, read. Neither of these more able writers does this. But the first writer is aware of his reader, of his rhetorical context. He says in effect: “I am a scientist—but think of me as one of you.” His is an appeal to the emotions. The second writer says: “The initiation of this discourse is grounded in invalid assumptions, and I must draw your attention to that fact.” It is an appeal to reason. In this she is apparently unaware of her reader, her rhetorical context.

Here I must briefly call attention to the three-way nature of this discourse exchange. It is not only writer and prompt that interact to create the “response,” marked or marked, on an essay test: the
CHALLENGING THE TASK

reader, that human judge, is truly a discourse participant. I have argued elsewhere (Cooper & Hamp-Lyons, 1989) that in large measure readers create the discourse anew as they read. If readers experience the discourse as discontinuous, as disrupted, we may predict that their responses will be negative. This possibility is represented in Fig. 4 by the challenge arrow after the writer's challenge. I suggest that the reader's inclination to challenge the writer's challenge may take the form of a low score, if the challenge is one the reader does not value. In contrast, as is implicit in the behavior of those Cambridge undergraduates of the 1950s, the "smart" challenge, the ability to take on the professor on his own terms and hold one's own, is often highly-valued and well rewarded.

What are the characteristics of the first example that led essay evaluators, English teachers all, to assign scores of 8 and 9 on a 9-point scale? My discussions with English specialists have shown, over and over, that comments focus on its language fluency. The writer is persuasive and controls persuasive language. And what are the failings of the second sample that led those same evaluators to assign scores of 4 and 5 on the same scale? Again, my discussions with English specialists have shown repeatedly that comments focus on incorrect lexis and poor spelling. The piece is not persuasive, does not engage the reader—and in not engaging the reader it leaves her free to notice the language weaknesses.

What makes the first piece fully competent or very close to it, and the second piece halfway to competence at best, for these judges? I would suggest that the emotional quality of the challenge in the first plays a large part in the readers' judgments, and that the rational quality of the challenge in the second similarly plays a large part in readers' judgments. I would further suggest that the first essay is not twice as good as the second, and that this is a case where readers' judgments are little based in the actual strength of the ideas or arguments the writers are making. Several professors in disciplines outside English/ESL have looked at these two pieces together, and have found them roughly comparable in writing quality, where the definition of "quality" was left open for them to make for themselves.
Conclusion

I believe that the focus on challenges, as exemplified by the two cases discussed above, has opened for me a whole range of insights and issues which can inform the study of prompts more generally. It has reiterated for me the critical significance of the selection and wording of prompts, not only in general but also in relation to the needs and background knowledge of the anticipated test audience. It has also led me back to basic issues such as the impact of readers on texts, and judgments of text quality. Once again I realize how improbable it is that we will ever pin down prompt difficulty, since so much of the difficulty is not in the prompt or the writer, but in the reader.

Without accompanying studies of students as they encounter written examinations we shall remain unsure of much that is important to an understanding of essay tests: how (indeed if) students read the prompt, how they establish salience for elements of the prompt, what makes them choose one prompt over another where a choice is given, how they decide what persona to present to the reader—all of this lies behind our numerical data. Since the many studies of topic effect have not led to consensus in the writing assessment community, a shift of focus seems called for. The study of challenges suggests that for a genuine understanding of what writers do on essay tests, and why, we will have to go much further. One possibility is the case study: an examination of the candidate’s writing development prior to the test. Such work is long overdue.

References


Appendix

Prompt Question and Two Challenge Responses

1. *New scientific processes often meet with opposition because of the pollution they cause to the environment, an example of which ...*  
As a scientist, how would you defend the continued use of such potentially harmful processes?

*Example A*

Even though I am a scientist, I strongly consider the opposition to new scientific processes as a healthy action. I think that man's new experiments should compromise their experiences and the environment, in order to avoid damages to the present and future nature. Basically, no man or nation has the right to, in the name of a scientific progress, destroy their own habitat—the only one we have now.

Fortunately, though, I believe that man can find better ways to guarantee the environment and mankind preserved. Indeed, I also believe that new scientific processes can be done just to improve the quality of life on Earth, in despite of all economic interests involving this man's action all around the world. Even though someone can find this thesis completely utopic, I really trust it.

*Example B*

The physical facts of pollution can be measured by using scientific equipments, and scientists know the process of the facts. Engineers who know the scientific knowledge only can develop facilities which reduces this harmful processes. On the other hand, politicians and executives of companies have a force to decide the use of the beneficial but harmful processes. The decision must be or reflect the will of people who are enjoying and are harmed by the process.

Therefore, the scientists only can give people proper information about the process, and the engineers only can give people proper information about the technology and the cost of preventing the harmful effects.

For me, the question above does not make sence. The choice of continual use of such potentially harmful process or cutting off the use of the process does not depend on the scientist. Scientists want to know everything in a rational way. The knowledge obtained by this way is so repeatable and testable, or reliable, that this knowledge have a power. The way how we use the power is not on the responsibility of the scientists.
Cross Currents is a semi-annual publication of the Language Institute of Japan (LIOJ) which provides a forum for the interdisciplinary exchange of ideas within the areas of cross-cultural communication and language skill acquisition and instruction. At Cross Currents, we are particularly committed to issues concerning both the theoretical and practical aspects of ESL/EFL instruction, cross-cultural training and learning, English language teaching as it applies to Japan, and English as an International Language.

Recent contributors to Cross Currents include:

Robert O'Neill
Braj Kachru
Alan Maley
Walter Hirtle
Richard Via
Larry Smith
Paul LaForge

<table>
<thead>
<tr>
<th>SUBSCRIPTIONS</th>
<th>Single issue</th>
<th>1 year</th>
<th>2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inside Japan*</td>
<td>¥1,300</td>
<td>¥2,600</td>
<td>¥5,150</td>
</tr>
<tr>
<td>Outside Japan — Individual Subscription</td>
<td>$8.00</td>
<td>$15.00</td>
<td>$25.00</td>
</tr>
<tr>
<td>Outside Japan — Institutional Rate</td>
<td>$19.50</td>
<td></td>
<td>$35.00</td>
</tr>
</tbody>
</table>

* Special rates available for JALT members. (See postal form in the Language Teacher.)

- Japan Subscriptions:
  School Book Service
  FI Building 4F
  1-26-5 Takadanobaba
  Shinjuku-ku, Tokyo 160
  03-200-4531
  (By postal transfer to Tokyo postal account No. 9-8616, or cash delivery—Genkin Kakitome.)

- Current Issues are also available in major Japanese bookstores through YOHAN in Japan.

- Back Issues are available from Cross Currents.

- Outside Japan:
  Cross Currents
  Language Institute of Japan
  4-14-1 Shiroyama
  Odawara, Kanagawa 250
  Japan
  (Make checks payable to Cross Currents (LIOJ) drawn on a U.S. bank.)