A New Approach to Advanced ESL Placement Testing

Ilona Leki

Like 84% of college English programs for native speakers (Gordon 29), nearly all colleges and universities that enroll non-native speakers administer some type of commercially available or locally developed English placement exam to incoming non-native students. Despite the widespread use of placement exams, dissatisfied program administrators and teachers complain that those exams that can be evaluated quickly and "objectively" do not test the appropriate skills, while those that do test appropriate skills require a great deal of time and expertise to evaluate. Although literature on assessment abounds, the body of literature on placement testing is surprisingly small, presumably reflecting the special nature of placement tests. (See, however, J. D. Brown for an excellent discussion of reading placement exams.) First, placement exams reflect the content of particular programs and are therefore less generalizable than other types of tests. Second, the results of student performance on a placement test are not as consequential as the results of achievement or proficiency tests which can, on the one hand, cause a student to fail a course or, on the other, stand in the way of a student's graduation. A placement exam merely matches a student with an appropriate course. Third, if a student is incorrectly placed, the teacher or the WPA can presumably remedy the mistake. Thus, placement exams do not serve the same kind of gate-keeping function as other types of writing assessments, holding "unprepared" students back in their progress towards a degree. In fact, despite their deep concern about the political, social, psychological, and philosophical implications of all other forms of writing assessment, assessment researchers are generally comfortable with placement tests (Weaver 43). It is perhaps in this privileged arena, then, that innovations in testing are most appropriately explored.

However "safe" the placement test environment may be, placement exam designers are nonetheless confronted with several key decisions. Where will the placement exam fall on the continuum between objective, standardized grammar test and holistically rated writing sample, between global-skill and local-skill testing, between academically-oriented and personally oriented writing? (See Perkins for further discussion of writing...
evaluation options.) This paper discusses these questions and one attempt
to find innovative answers for ESL (English as a Second Language) writing
placement testing.

Debate on Writing Assessment

Most central to the current debate on writing assessment is the question of
whether to test students by means of writing samples or of objective,
grahmat-type tests (White; Leonhardt; Gordon; Carlson and Bridgeman;
Homberg). This issue is of particular concern in the testing of ESL students
since not only writing ability but general proficiency in English must be
determined to place students in appropriate writing courses.

The argument against using objective, grammar-type tests to evaluate
writing has always been that these tests lack validity since they do not
directly test composing skills. Keppy and Adams have found, for example,
that STEL (Structure Tests of the English Language) scores are poor
predictors of students' writing abilities. Raimes describes the mismatch
between student scores on the Michigan Test of English Language Profi­
cency (testing grammar, vocabulary, and reading) and a holistically
evaluated writing sample, and comments: "The lack of correspondence
between language proficiency score and placement in a writing course
assessed according to a writing sample points to the limitations of stan­
dardized testing for the evaluation of writing" (448). An informal study at
the University of Tennessee revealed very little correlation between scores
on the TOEFL (Test of English as a Foreign Language) and ratings on an
inhouse placement exam. While many institutions rely on TOEFL for basic
admissions information on second-language students, few trust TOEFL
scores for placing of students in writing courses. Even if we disregard the
lack of correlation between standarized tests and writing ability, for many,
writing samples simply have a great deal more face validity as tests of
writing than do standardized objective tests; that is, writing samples
appear to be valid tests of writing ability.

However, as Gordon notes, "face validity is the least sophisticated
kind of validity and psychologists are not satisfied with a measure if it
meets this criterion alone" (31). Furthermore, several recent studies do
seem to show high correlation between results of objective tests and
holistically rated essay exams. Leonhardt found that composite scores on
TOEFL for a hundred locally tested students tested correlated well with
scores given essays written by these same students and rated holistically
using the Newbury House ESL Composition Profile (Hughey et al), a
commercially produced and widely used, holistic scoring guide. The
producers of TOEFL itself resisted developing their long-awaited writing
test, despite pressure from ESL writing teachers, because they too have
claimed a high correlation between TOEFL and holistically-scored writing
samples. Furthermore, arguing for a return to standardized placement
testing for native speakers, Gordon points out the many problems inherent
in one-shot essay exams. Appropriate topics are difficult to come up with;
students have varying degrees of familiarity with any given topic; students
have only one opportunity to show their skills; students who are heavy
revisers are at a disadvantage because of time limits; the wording of the
prompt affects students ability to comprehend and respond to what is
being asked of them; and finally, evaluating essays holistically requires a
great deal of time, care, and preparation of the raters.

The issue of whether or not to use a writing sample is further compli­
cated in the case of second-language students. For a variety of reasons,
including simple lack of motor fluency in writing English, second-
language students typically write much more slowly than native speakers.
Producing only 250 words in a hour may prove difficult even for students
fairly fluent in English and skilled in writing. Thus, the writing sample
obtained may be rather small. Furthermore, since the student is in control
of the language used in the writing sample, the student may simply avoid
those structures he/she is unsure of. (See Schacter, for a discussion of
avoiding problematic features of a second language.) An objective, gram­
mar-type test of sentence level skills might require students to show
familiarity with a greater variety of structures.

A second issue particularly important in testing non-native speakers of
English involves the distinction between global and discrete point testing.
A writing sample draws on global language skills, calling for students to
produce text by making use of everything they have acquired of English,
morphology, syntax, lexicon, discourse features, register, etc. Discrete
point testing, on the other hand, pulls one piece of language at a time out
of the whole fabric of a language and tests students' masteries of that
individual bit. To know the answer to a discrete point test item, the student
needs only to focus attention locally: a item or an item; believe in or believe
to; they was or they were. A language context no broader than one word is
necessary to make these local decisions. Discrete point testing implies a
view of language as consisting of discrete, bite-sized units, each of which
can be mastered individually and added to previously learned bits to
result in full mastery of the language. If such a view of language informs a curriculum, its language students know the rules but cannot use the language, score extremely well on discrete point grammar tests but have great difficulty writing English (and reading, speaking, and comprehending spoken English, as well), a phenomenon not at all uncommon among ESL students. While this kind of knowledge of rules is obviously not to be disdained, since it can be applied in editing tasks, this type of knowledge does not generate language (Krashen) and constitutes only a small part of the knowledge and skill students need to survive in a second-language academic environment.

If global skills—that is, those skills more central to success in a second-language environment—are to be tested, a writing sample seems an appropriate way to learn how well a student writes. However, a third issue test developers must consider is the nature of the writing prompt, not only the topic the students will write on but also the framing and wording of the topic. In order not to discriminate against any particular group of students, the topic should be one all students can write on with some familiarity. This issue is especially pertinent for ESL students, who may be graduates or undergraduates, with widely varying ages and world experience, coming from all parts of the world, and representing a variety of academic disciplines. The question also arises whether the topic should be personal or academic. Personal subjects may be relatively easy to write on for students who have been educated in the U.S. school system with its emphasis on individual introspection. However, writing on personal topics may be unfamiliar to some ESL students and may cause discomfort. Furthermore, more academic topics better reflect what these students’ actual writing tasks will be at the university. (See Carlson and Bridgeman, and Horowitz for testing and teaching academic English).

Once general topics are determined, the question of framing and wording the essay-exam prompt arises. This question has been the subject of several recent investigations. First-language researchers (Hoetker and Brossell) and second language reading researchers (Carrell and Eisterhold) cite the current perception that meaning does not reside in a text but rather is constructed by the reader. Even in written texts, meaning is negotiated and therefore not stable but open to interpretation. Conlan (in Gordon), for example, reports that native speakers misinterpreted a prompt mentioning “waste” in American society (that is, wastefulness) to mean bodily waste. To prevent misinterpretation Hoetker and Brossell advise minimizing “opportunities for miscommunication by keeping the topics as brief as possible, as simply phrased as possible” (328). However, Hi-

rokawa and Swales find that ESL graduate students prefer longer, more academic prompts to prompts worded simply and personally. Presumably, these graduate students, being experienced writers in their native languages, are able to transfer some of their first-language writing strategies into the second-language context and are, therefore, able to profit from greater rhetorical specificity. Going even further, Hult argues for including a rhetorical frame in the prompt that provides students with a first sentence (a theorem) and a choice of follow-up second sentences, each of which indicates a different direction for developing the essay, and all of which serve to provide a focus for the essay. To further orient students in writing the exam, some exam prompts have been written specifying the audience and purpose. However, several researchers note that students writing essay exams tend to ignore these designated audiences and purposes and instead consistently construe the writing context as exactly what it is, an essay exam written for an audience of exam evaluators, regardless of what the prompt may indicate (Hult, Greenberg).

For second-language writers, the question of misinterpretation of the prompt is again particularly at issue. If a prompt is potentially confusing to a native speaker, the problem is exacerbated when a non-native speaker interprets the prompt. But for second-language students, misinterpretation may be a clue to the students’ level of English proficiency. Although test designers warn against making a writing test into a test of reading the writing prompt (Carlson and Bridgeman), with second language students, an ability to read and interpret a prompt forms part of a student’s general ability to function in English. For this reason, a test of that general ability is arguably not inappropriate. Here again the special nature of placement exams comes into play. Trouble with reading may indicate the student’s need for additional help to function in an English language educational institution.

ESL Writing Placement Exam: Old and New

The placement exam formerly used to place ESL students in writing classes on this campus consisted of a writing sample and a grammar test. The first part was an essay exam on general topics, such as, compare the training you received at home to the training you received at school or describe a tradition in your country that is disappearing. The students had one hour to write about 250 words. Each essay was holistically rated. The second part of the exam consisted of a locally prepared sixty-item grammar test.
Like most standardized objective tests, each item tested a different grammatical structure, and none of the sentences was related in content. The test covered verb use (fill in the blank with the correct form of the verb); word order, question formation, adjective-clause formation through sentence combining, logical sentence completions using cohesion words like “therefore” and adverbial clause markers, noun clause formation through sentence combining, reported speech, and preposition use. The exam ended with a forty-item multiple-choice section in which students selected the correct sentence completion from among three possible completions.

We were unhappy with this exam for several reasons. While the essays did give us a look at some of our students’ global writing skills, the topics suffered from a flaw that plagues nearly all writing sample tests and yet which is rarely mentioned in the debate on essay exams. This test required students to write on a surprise subject. Not one they had been thinking about, but one they were coming to completely cold. The artificiality of this type of writing situation is obvious; it is only during these types of exams that anyone is ever required to write on a subject essentially pulled from a hat.

The objective parts of the old placement exam also seemed unfair. The grammar items tested appeared with no context and with no content continuity from sentence to sentence. This forced the student to re-determine the context for each item, an added, and presumably disruptive, burden.

The worst problem with the exam, however, was that students were misplaced. Our courses had long since de-emphasized grammar in favor of communicative skills, yet the emphasis on the placement exam was still on grammar. As a result, students whose grammar skills were strong, as is often the case with ESL students, but who had difficulty applying those skills in actual language use, would be placed into courses they were unable to handle (another example, if one is needed, of how little impact grammar skills can have on writing ability). We thus faced the need to create a new placement exam for incoming international students.

In constructing a new placement test we took into account the various issues surrounding writing assessment and hoped to create an exam that would correctly place students in courses, that would be quick and easy to score with a fair degree of inter-rater reliability, and that would reflect current views in the profession that de-emphasize passive, discrete point grammar exams and lean toward proficiency-based exams, that test active, global skills—students’ ability to use the language.

Attempting to compensate for the problems of both discrete point grammar tests and single-shot essay exams, we hoped to construct a new placement exam that would reflect the skills students would need in their work at the university. This meant more emphasis on reading and incorporating in the writing task information from printed texts. It also meant emphasizing grammar differently and less. Totally de-contextualized, grammar-exercise-type sentences would be avoided. To reflect better what students would be asked to do with English at the university, editing would replace such artificial activities as combining sentences and filling in blanks.

Although face validity, that is, the appearance of validity, may be the least sophisticated of validities, the intuitive appeal of a writing sample to test writing ability is a very powerful one for teachers and students alike. Because we felt that students would do better on an exam that seemed valid to them, and despite the problems inherent in one-shot writing samples, we felt committed to require an essay as part of the placement exam to give us an idea of the students’ ability to write a sustained piece of English. Since we have neither the time nor the staff to allow us to take several writing samples, we hoped to minimize the negative features of the single-shot writing sample by creating a writing context that would permit students to consider the topic before having to organize and express their own thoughts on the subject. The topic for the writing sample would be more academic than personal, and the prompt would require students to address the topic in a manner that reflects typical academic writing assignments (Horowitz).

The new exam consists of three parts. The first part tests reading, the second writing, and the third grammar. The test begins with a reading passage about a thousand words long on a subject related to education, in this case, home schooling as an alternative to classroom schooling. While possibly a new concept to some of these international students, the topic of home schooling would certainly draw on the students’ experience with their own educational systems. The text is an adaptation of an article from a campus publication, a magazine directed at college students and therefore presumably at a level and on a topic appropriate for students. The reading is followed by eight comprehension questions that call for inferring meanings of words and inferring information, as well as testing simple comprehension. This part takes twenty-five minutes. (See Appendix)
The second part of the exam is the essay. Students choose from one of two subjects. The first choice is directly related to the article the students have just read. Since they have spent almost half an hour reading and answering questions on the reading passage, ideas on the subject are still fresh in their minds. The prompt calls for a summary of the information in the article and an analysis of some kind. Students may, for example, discuss advantages and/or disadvantages of home schooling or they may describe instances of home schooling with which they are familiar and comment on the results of that form of education.

Fearing that students might grow tired of the subject of home schooling, we provide a second possible topic for the essay that approaches the general subject from a different point of view. In this instance, for example, since the article on home schooling briefly discusses the function of grades, the second choice of subjects for the placement exam deals with grades: should grades be eliminated because they cause students to study only for grades or, on the other hand, are grades important motivators for students?

Since the students may incorporate information from a written source in their own essay, a clearer picture of their academic skills emerges. We have found, in fact, that the generally more academically prepared students (though not necessarily more proficient in English)—for example, the graduate students—tend to select the topic directly related to the article they have just read, while the younger students seem to feel more comfortable with the second, broader subject, which makes less specific cognitive and academic demands. However, the point of giving students the option of writing on a subject they have just read about is not to test their ability to incorporate that material but rather to make available additional ideas, and academic demands. However, the point of giving students the option of writing on a subject they have just read about is not to test their ability to incorporate that material but rather to make available additional ideas, support, even vocabulary. The reading passage activates networks of memory and information in the students' minds which they may then draw upon for either essay, since both of these essay topics maintain continuity with the rest of the exam. (See Carrell and Eisterhold for a discussion of schema theory and ESL reading pedagogy.)

Evaluating this section of the exam is certainly more time-consuming than evaluating an objective exam, and this essay is still a single-shot writing sample, but since the content of the essay is prompted by the reading, the panic of finding no ideas to write on is virtually eliminated. Students can use in their own essays the ideas they have just read about. Instead of testing students' abilities to think fast on the spot, this approach provides the students an opportunity to show their ability to synthesize new information from a written source with any information or experience they may already have on the subject and to discuss this knowledge in a way which communicates effectively to English-speaking readers.

Since one of the language subskills ESL students will need in their work at the university is the ability to recognize and correct errors in a text, the last thirty minutes of the placement exam are devoted to grammar and editing. Every item on this part of the test has as its subject matter the topic of the reading. The section testing verb tenses, for example, consists of a cloze passage in which only verbs have been deleted and which repeats or elaborates on information from the reading. Although this cloze tests only one feature of English, it is unlike discrete point grammar tests since cloze tests draw on global rather than local skills, requiring students to examine the linguistic context in order to fill in the blanks (J. D. Brown "Cloze").

The cloze is followed by an editing test of about twenty-five items in individual sentences with one error per sentence. The sentences are grouped by errors; that is, there is a group of sentences which contains an error in the formation of the adjective clause in the sentence, another group with errors in reported speech patterns, and so on. The groups are labeled with traditional grammatical terms, like "adjective clause," to help students who know those terms focus quickly on the flawed part of the sentence, but students are directed to ignore those labels if they are not helpful. The topic of each sentence is again taken from the reading passage. Thus, the subject matter of the sentences is continuous, providing a context for the sentences, reducing the workload, and therefore allowing students to concentrate on the detailed concerns of editing. Furthermore, in this edition test, students are called upon, not to perform such artificial manoeuvres as combining two simple sentences into a complex sentence, but to comb sentences for editing errors, an important skill for college students to develop. Yet since there is only one error per sentence, marking this exam admits a negligible risk of ambiguity, requires no particular training, and is as fast and easy as marking multiple-choice items. (See Appendix)

Placement of Students

The writing sample of this exam gives us a holistically scored measure of student writing proficiency. Essay raters are all teachers in the ESL writing program who use their individual familiarity with the courses available in
the curriculum to rate the essays and place the students in the courses which will most benefit them. But these placements are also informed by objective measures of student ability in reading and grammar. In many cases, all three measures place the student in the same course level. In a significant number, however, the measures do not match, not a surprising or disturbing outcome since different measures are intended to measure different skills. For these cases, we have built into the evaluation a gray area, or an area of flexible placement. When the measures do not match—that is, when one measure places a student into one course and another measure places the student into another course, the raters reevaluate the writing sample, rereading the essay, but this time focusing primarily on the writer’s fluency and proficiency in English, the second-language writer’s whole language skills. We look at this writing, in a sense, as a picture of the students’ acquired abilities to produce English rather than as a picture of their abilities to construct an essay. If this reevaluation indicates, for example, that the writer’s fluency and proficiency in English are weak despite strength in the ability to construct an essay, the student is placed into a language proficiency course rather than into an academic writing course.

This type of placement exam combines the advantages of both a proficiency exam, displaying students’ global skills in using English, and an achievement test, focusing on more discrete, local formal features of English. In the essay, students have nearly total control of the language they will use, thus giving us a picture of their English writing proficiency as it is enhanced by the prompting of the reading passage. To help us evaluate more precisely the range of their abilities in English, the editing section tests students’ abilities to recognize and correct structures they might have avoided in their essays. Although it is difficult to document statistically the assertion that this placement exam serves students better than the previous one did, anecdotal evidence from both teachers and students indicates that students are now being more appropriately placed. We no longer have the problem of students in writing courses who can cite grammar rules and apply them in single sentences but who cannot use the language for their own purposes or who cannot use the rules they know to edit their writing.

Limitations

One of the problems in constructing an exam for ESL students is the enormous diversity of the population to be tested: both graduate students and undergraduates, students who have lived in the U.S. for some time and those who arrived the day before the placement exam, people with limited experience of the industrialized world, and people from sophisticated world capitals. Thus, the degree to which the topic of an essay exam is culture bound is always an issue in ESL testing. However, by providing the students with a reading that anchords the rest of the exam, we feel that the disparity in student experience with the world and with U.S. culture is reduced and that we are less likely merely to be identifying students who are able to “discourse about current events, ethical conundrums, American politics, literary quotations or whatever” (Hoetker and Brossell 328).

A second problem is related to the editing portion of the exam. As with any test focused on discrete features of language, it is possible that a student otherwise proficient in English retains fossilized errors in specific features of English grammar and will therefore not recognize as errors those same features if they appear in the editing portion of the exam. Nevertheless, since editing is a skill which is taught in our classes, and one which students will need in an academic context, a focus on editing skills does not seem out of place.

Conclusions

Many questions in writing assessment remain unresolved. Furthermore, in her review of three important books on writing assessment, Weaver warns that “WPAs hopeful of finding a reliable writing test to import into their programs will be advised by these authors to develop their own, involving local faculty and clarifying local criteria” (41). White encourages classroom teachers to involve themselves in the creation, administration, and evaluation of writing tests in order not only to maintain some degree of control over the fates of their students rather than leaving test construction to administrators but also to take back to their writing classes new insights on the problems of constructing writing assignments and evaluating written work.

An appropriate placement exam “typically includes a sampling of material to be covered in the curriculum” (H. D. Brown 216) of a given institution. In developing this placement exam, we tried to take into account current thinking on writing assessment, and we hoped the exam would reflect the work students are called upon to do in our particular
setting. Our solutions may not be feasible or even desirable in other settings, but we hope they offer WPAs another perspective from which to view their own placement requirements.

Notes

1. The Test of Written English (TWE), which TOEFL did develop to provide writing samples, admits to limited usefulness with advanced second language writers.

Works Cited


Leonhardt, N.L. “The Effect of Topic on ESL Composition Scores.” Southeast Regional TESOL, October 1988, Orlando, FL.


Appendix: Excerpts from New Placement Exam

READING

DIRECTIONS:
Read the following short article and then answer the questions which follow. You have 10 minutes to read the article.

The Home-Schooling Alternative

Dissatisfied with public schools, thousands of families are learning that home schools can do a better job.

3. Colfax and his three brothers, who were also educated at home, combined ranch chores with learning. The children helped their parents build a redwood house, and developed a first-hand understanding of geometry as a result. Grant learned economics by recording the financial transactions concerning the buying and selling of pedigreed dairy goats. He studied at the kitchen table at night with kerosene lamps, and kept up with such extra reading as The New York Review of Books and The New Republic.

QUESTIONS:
1. Which of the following best expresses the main idea of paragraph 3. Circle the correct answer.
   a. All four of the Colfax children were educated at home.
   b. Grant Colfax's education was combined with practical work around his home.
   c. Even though he might have been tired from doing ranch chores, Grant Colfax worked hard for his education.

GRAMMATICAL STRUCTURE TEST

You have 30 minutes to complete both Part I and Part II of this test. The test goes to page 11.

PART I: VERBS

DIRECTIONS:
Read the following passage through. Then fill in the blanks with the correct form of the verb in parentheses. Give only one answer per blank. Each blank is worth 1 point (11 points total).

. . . When the Parshleys began educating their son Jason at home, he had been attending school for only a year, yet he (16) (already, learn) to be passive about school and uninterested in learning. However, since leaving school, he (17) (become) excited about learning again.

PART II: EDITING

DIRECTIONS:
Each sentence below has one error. Your task is to find the one error and to correct it as in the example.

EXAMPLE: This exam Ke easy.

Do not rewrite the whole sentence. You will be given occasional hints about where the error is located. If you do not understand the hints ignore them. Each sentence is worth 1 point (23 points total).

(Hint for #12, #13, #14, #15: Adverb clause errors)

12. Although most children go to public schools, but the schools do not do a good job of educating them.

13. Good students are sometimes frustrated because of the pace of regular classes is too slow for them.
17. Despite Holt's criticisms of schools, most parents still sending their children to public schools.

19. Parents might be doing their children a favor by educate them at home.

23. The Parshley children are intelligents, and now they are also excited about learning.